NEURAL NETWORKS for Smart IAM (Identity and Access Management)

Apurva P. Keskar[1], Poonam R. Faldu[2]
[1]*Assistant Professor, Parul Polytechnic Institute, Parul University, Vadodara*
[2]*Head of Department, Parul Polytechnic Institute, Parul University, Vadodara*

***Abstract* - Cloud computing enables information to be accessed from a variety of devices. It makes it possible to connect with your business at any time and from any place. There are several perks to using the cloud, including reduced IT expenditures, infrastructure costs, adaptability, and continuity planning. Workplace flexibility, collaborative efficiency, and automated upgrades are all factors to take into account. These advantages can only be realized with solid identity and access control for cloud apps and services. Our system challenges include user password fatigue, manual provisioning and an act that is prone to failure, visibility of compliance, isolated user directories for each application, access management across an explosion of browsers and devices, updated application integration, the variation of administrative models for different applications, and a lack of knowledge of quality standards for optimal cloud service resource usage. To address these concerns, this analysis gives a neural network-trained intelligent identity and access management framework.**

***Index Terms* - Cloud computing, Neural network, Smart identity and access controls.**

INTRODUCTION

Connectivity to cloud resources and resource extraction necessitates the use of an authentication and authorization portal, which causes an unauthorized party to attempt to breach the system and get unauthorized access resources. This reason is especially vulnerable in the cloud compared to traditional infrastructure where only a few administrators have access to features as well as geographical security To provide successful identity and access, four important roles are employed.These are identity provisioning and processioning, authentication, and authorization, Confederation, authentication, and user profile management, as well as compliance support [1]. Due to the obvious substantial concerns associated with confidential data theft and disclosure, enterprises are cautious about embedding their resources in the cloud environment. Study conducted recognizes the critical relevance of access control security.[3].

According to International Data Corporation, security remains a hurdle for cloud customers. Actual security events, such as outages at Amazon Web Services and Email, are a clear illustration of high-level insecurity. Identity, infrastructure, and information are the three most important aspects of security.[34]

LITERATURE REVIEW

Identity and Access Management in cloud computing was proposed by Khandakar Entenam Unayes Ahmed and Vassil Alexandrov.[4] According to an International Data Corporation survey, 87.5 percent of customers are hesitant to employ cloud computing for future project deployments. The reason for this reluctance is the security of data stored on the cloud. Identity and Access Management is one such solution presented by the authors, which offers increased data security in the cloud. This approach comprises Kerberos authentication and authorization based on an RBAC processor that implements authorization policies for allowing user access using Java. [6] The services provided by this model include authentication server service, ticket-granting server service, RBAC processor service, and edge node service. This strategy is vulnerable to losing personal identifying information owing to the risk of a man in the middle attack between the cloud service provider and a trusted third party.

Ruediger Schulze offered identity and access management for payment card industry cloud services [17]. Its data integrity rule specifies that all customer

’s data environments implicated in the credit card payment process adhere to all applicable requirements. Identity management in the cloud was presented by Alina Madalina Lonea, Huaglory Tianfield, and Daniela Elena Popescu [28]. This paradigm is mostly concerned with web application security and virtualization security concerns. There is discussion of mitigation approaches that offer proper identification and management architecture. Umme Habiba, Rahat Masood, Muhammad Awais Shibli, and Muaz A Niazi suggested assessment criteria for assessing existing and prospective cloud-based identity and access management systems [7]. This study looks into assaults on identity and access management systems, as well as responses, as well as the feature mechanism relationship utilized to assess cloud-based identity and access management systems.[9] Surya Majumdar, Taous Madi, Yushun Wang, Yosr Jarraya, Makan Pourzandi, Lingyu Wang, and Mourad Debbabi provided a framework for cloud security compliance auditing using OpenStack in "Security Compile Auditing of Identity and Access Management in the Cloud: Application to OpenStack" [5]. In "Identity-based Access Control for Digital Material based on Ciphertext Policy Attribute-Based Encryption," [13] Win-Bin Huang and Wei-Tsung Su developed a method to digital content. The solution is identity-based access control, which is further supported by ciphertext-policy attribute-based encryption (iDAC) [21]. In "User-Centrist Trust-Based Identity as a Service for Federated Cloud Environment," Samlinson.E, M.Usha suggested a service that attempts to create trust among Cloud Service Providers (CSPs) [12]. A user-centrist trust agent identification service is what this service is called. The authors addressed different standards such as SAML [12], OAuth [22], XACML [23], and SPML in order to provide trust and safe access to cloud services.In "A Paradigm for Identity Management with Privacy in the Cloud," [30] Jorge Werner and Carla Merkle Westphall suggested a model that will handle privacy concerns linked to Personally Identifiable Information (PII) [27]. Using Open ID Connect (OIDC), a prototype with dynamic scopes, federation agreements, and security policies was created [19].

PROPOSED METHODOLOGY

Cloud computing is a collection of various customization computing assets such as servers, networks, services, apps, and storage that enable cloud customers to get beneficial and on-demand access to cloud services. Individuals frequently refer to and use cloud computing in a variety of commercial industries.[10] However, managing these identities and regulating access by cloud clients and apps remains a major issue to this day. Implementing a reliable identity and access management (IAM) [14] system in the cloud is required to improve venture security. This research presents an intelligent and trustworthy IAM system based on neural networks. The goal is to develop a system capable of achieving an intelligent identification method for authenticating actual users and accessing data on the cloud server.

There are two parameters that make up access management:

(1) Efficient and simple file access

(2) File policy management

Files that are often accessed by users are retained on a separate active server based on learning and association rules in order to enable quicker access and minimize the latency of such files to the IAM (Identity and Access Management) system for further processing.[24] For this, the machine learning model anticipates data viewed by the consumer in the previous several days, as seen in Figure 1.

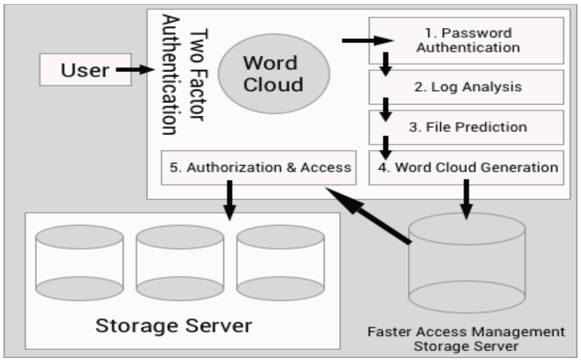
User
Two Factor Authentication
Word Cloud
1. Password Authentication
2. Log Analysis
3. File Prediction
4. Word Cloud Generation
5. Authorization & Access
Storage Server
Faster Access Management Storage Server

Fig 1. Machine Learning Model for File Prediction

The system will detect user logs in order to learn about a user's most frequently accessed files. [25] These user file logs can be produced by explicitly deploying Audit control services [29] on the cloud server. The resulting file logs will be in raw or unsupervised format. For subsequent processing, the acquired unsupervised data must be translated into supervised CSV (Comma Separated Values). [26]The data-set will be pre-processed [18] in order to clean and identify the string values so that the machine learning model can be built. The data is ready for analysis after the numerical and categorical data has been gathered following the pre-processed step. The arbitrary data will then be taken from the files specified by the Identification system for authentication of a certain user. The data will then be used to extract certain words, which will be blended into a word cloud [2], and the user will be required to identify the file name containing those words in order to gain access to the file server.

PROCESS MODEL FOR TRAINING AND EVALUATION

To identify the files (through filenames) based on the three fundamental properties listed above, a machine learning model must be chosen. Additionally, the IAM system will use these files for authentication. Without using any clustering methods, the machine learning model is primarily directly trained and assessed on pre-processed user logs contained in a CSV (Comma Separated Values) file. Figure 2 depicts the forecast using the Random Forest Algorithm.

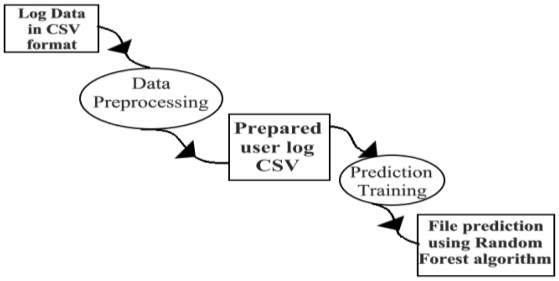

Fig 2. Prediction using Random Forest Algorithm

To identify the user files, another model employing Artificial Neural Network (ANN) is implemented on pre-processed CSV (Comma Separated Values) data without clustering. Figure 3 depicts the forecast made by an Artificial Neural Network.[31]

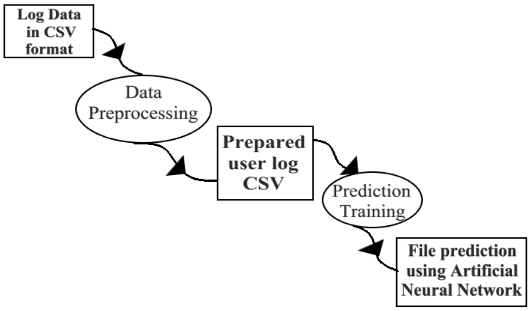

Fig 3. Prediction using Artificial Neural Network

When deployed directly without creating any clusters, the Random Forest [11] algorithm outclasses the Artificial Neural Network [15] in terms of accuracy. However, the logarithmic loss of the Random Forest algorithm is significantly higher than that of the Artificial Neural Network.[32]

A novel strategy is used, in which a clustering algorithm is applied to the pre-processed data before deploying the Artificial Neural Network algorithm on the clustered data. Several Clustering algorithms were used, and it was discovered that the Gaussian Mixture Model (GMM) [20] produces the best results with a Silhouette score of 0.4334 and the lowest Inertia score in segregating the files into two clusters, one for the files that are useful for the Identification process and another for the files that are not.

EXPERIMENTAL ANALYSIS

To begin appropriate data analysis, the model initially accepts audit files including log records as input and transforms the raw data provided into a CSV (Comma Separated Value) file. The data will be split among 16 columns in the CSV (Comma Separated Value) file.[33]

Upon exporting our information to a CSV (Comma Separated Value) file, the model pre - processes it to make it easier to examine. The model then applies multiple Machine Learning algorithms to evaluate which ones offer us with the best evaluation criteria.

Name	Train time	Accuracy	Precision	Recall	Log loss	ROC AUC
Random forest	7s	0.76	0.37	0.44	0.62	0.98
ANN	0s	0.76	0.37	0.44	1.46	0.95
GBT	2s	0.71	0.42	0.49	1.44	0.94
Logistic Regression	9s	0.57	0.18	0.22	1.95	0.93

The acquired findings show that the Artificial Neural Network outperforms the Random Forest and, when combined with the Gaussian Mixture Clustering model, achieves around 99 percent accuracy. [35] The model can successfully forecast the ideal file for word cloud production using the upgraded artificial neural network modified with the Gaussian mixture. Textract was used to create the word cloud. Multiple word clouds will be provided to a user for second-factor authentication, and if an acceptable selection is made, full file access and authorization will be granted.

CONCLUSION

An efficient model is suggested in this study to extract user's file access logs, learn from them, and give a word-cloud based on an intelligent identification system for user identification, authentication, and authorization of files, ensuring efficient and speedier file access. Model pre-processed the access logs by converting them to a CSV (Comma Separated Value) file. The data is then examined using several machine learning methods, and experimentally, Artificial Neural Network was shown to have an ideal performance measure of 99 percent accuracy.

REFERENCES

[1] Ahmed K.E.U., Alexandrov V. (2011): Identity and Access Management in Cloud Computing. In: Mahmood Z., Hill R. (eds) Cloud Computing for Enterprise Architectures. Computer Communications and Networks. Springer, London. https://doi.org/10.1007/978-1- 4471-2236-4_6.

[2] Andrei, T. (May 21, 2011): Cloud Computing Challenges and Related Security Issuess. A Survey Paper (online), http://www1.cse.wustl.edu/~jain/cse571-09/ftp/cloud/index.html.

[3] Ahmad, I., Basheri, M., Iqbal, M. J., & Rahim, A. (2018): Performance Comparison of Support Vector Machine, Random Forest, and Extreme Learning Machine for Intrusion Detection. IEEE Access, 33789–33795. doi:10.1109/access.2018.2841987.

[4] Anjana, P. S., Badiwal, P., Wankar, R., Kallakuri, S., & Rao, C. R. (2019): Cloud Service Provider Evaluation System Using Fuzzy Rough Set Technique. IEEE International Conference on Service-Oriented System Engineering (SOSE). doi:10.1109/sose.2019.00033.

[5] Bethencourt, J., Sahai, A., & Waters, B. (2007): Ciphertext-Policy Attribute-Based Encryption. IEEE Symposium on Security and Privacy. doi:10.1109/sp.2007.11.

[6] Brown, K. P., Hayes, M. A., Allison, D. S., Capretz, M. A. M., & Mann, R. (2012): Fine-Grained Filtering of Data Providing Web Services with XACML. 2012 IEEE 21st International Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises. doi:10.1109/wetice.2012.41.

[7] CSA, (October 14, 2010): Security Guidance for Critical Areas of Focus in Cloud Computing V2.1 (online) Cloud Security Alliance (2009), http://www.cloudsecurityalliance.org/csaguide.pdf.

[8] CPNI, (October 2, 2010): Information Security Briefing Cloud Computing (online) Centre for the Protection of National Infrastructure, http://www.cpni.gov.uk/Docs/cloud-computing-briefing.pdf.

[9] Chi, M.-T., Lin, S.-S., Chen, S.-Y., Lin, C.-H., & Lee, T.-Y. (2015): Morphable Word Clouds for Time-Varying Text Data Visualization. IEEE Transactions on Visualization and Computer Graphics, 21(12), 1415–1426. doi:10.1109/tvcg.2015.2440241.

[10] Dwivedi, S. K., & Rawat, B. (2015): A review paper on data preprocessing: A critical phase in web usage mining process. 2015 International Conference on Green Computing and Internet of Things (ICGCIoT).

[11] Fett, D., Kusters, R., & Schmitz, G. (2017): The Web SSO Standard OpenID Connect: In-depth Formal Security Analysis and Security Guidelines. IEEE 30th Computer Security Foundations Symposium (CSF). doi:10.1109/csf.2017.20.

[12] Gruschka, N., Iacono, L .L. (2009): Vulnerable Cloud: SOAP Message Security Validation Revisited. In: IEEE International Conference on Web Services, ICWS, Los Angeles, pp. 625–631.

[13] Gruschka, N., Jensen, M. (October 20, 2010): Attack Surfaces: A taxonomy for Attacks on Cloud, vol. 5(7), (32) (2010), http://download.hakin9.org/en/Securing_the_Cloud_hakin9_07_2010.pdf.

[14] Gross, T. (n.d.). (2003): Security analysis of the SAML single sign-on browser/artifact profile. 19th Annual Computer Security Applications Conference, Proceedings. doi:10.1109/csac.2003.1254334.

[15] Habiba, U., Masood, R., Shibli, M.A. et al. (2014): Cloud identity management security issues & solutions: a taxonomy. Complex Adapt Syst Model 2, https://doi.org/10.1186/s40294-014-0005-9.

[16] Huang, Y., Englehart, K. B., Hudgins, B., & Chan, A. D. C. (2005): A Gaussian Mixture Model Based Classification Scheme for Myoelectric Control of Powered Upper Limb Prostheses. IEEE Transactions on Biomedical Engineering, 52(11), 1801– 1811. doi:10.1109/tbme.2005.856295.

[17] IBM, (October 2, 2010): IBM Point of View: Security and Cloud Computing (online), ftp://public.dhe.ibm.com/common/ssi/ecm/en/tiw14045usen/TIW14045USEN_HR.PDF.

[18] Jensen, M., Schwenk, J., Gruschka, N., Iacono, L. L. (2009): On technical Security Issues in Cloud Computing. In: IEEE International Conference on Cloud Computing, Bangalore, pp. 109–116.

[19] Jing, X., Jian-Jun, Z. (2010): A Brief Survey on the Security Model of Cloud Computing. In: Ninth International Symposium on Distributed Computing and Applications to Business Engineering and Science (DCABES), pp. 475–478.

[20] Kandukuri, B.R., Paturi, R.V., Rakshit, (2009): A.: Cloud Security Issues. In: IEEE International Conference on Services Computing, Bangalore, pp. 517–520.

[21] Katsuno, Y., Kundu, A., Das, K. K., Takahashi, H., Schloss, R., Dey, P., & Mohania, M. (2016): Security, Compliance, and Agile Deployment of Personal Identifiable Information Solutions on a Public Cloud. 2016 IEEE 9th International Conference on Cloud Computing (CLOUD). doi:10.1109/cloud.2016.0055.

[22] Lonea A.M., Tianfield H., Popescu D.E. (2013): Identity Management for Cloud Computing. In: Balas V., Fodor J., Várkonyi-Kóczy A. (eds) New Concepts and Applications in Soft Computing. Studies in Computational Intelligence, vol Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-28959-0_11.

[23] Liu, K., & Xu, K., O. Auth. (2012): Based Authentication and Authorization in Open Telco API. International Conference on Computer Science and Electronics Engineering. doi:10.1109/iccsee.2012.275.

[24] Majumdar, S., Madi, T., Wang, Y., Jarraya, Y., Pourzandi, M. Wang, L., & Debbabi, M. (2015). Security Compliance Auditing ofIdentity and Access Management in the Cloud: Application to OpenStack. doi:10.1109/cloudcom.2015.80.

[25] Rittinghouse, J.W., Ransome, J.F. (2010): Cloud Computing Implementation, Management and Security. CRC Press, Boca Raton.

[26] Ramgovind, S., Eloff, M.M., Smith, E. (2010): The management of security in Cloud Computing. In: Information Security for South Africa (ISSA), pp. 1-7.

[27] Schulze R. (2018): Identity and Access Management for Cloud Services Used by the Payment Card Industry. In: Luo M., Zhang LJ. (eds) Cloud Computing. Lecture Notes in Computer Science, vol 10967. Springer, Cham. https://doi.org/10.1007/978-3-319-94295- 7_14.

[28] Samlinson, E., & Usha, M. (2013): User-centric trust based identity as a service for federated cloud environment. 2013 doi:10.1109/icccnt. 2013.6726636.

[29] Sharma, A., Sharma, S., & Dave, M. (2015): Identity and access management- a comprehensive study. International Conference on Green Computing and Internet of Things (ICGCIoT). doi:10.1109/icgciot.2015.7380701.

[30] Thukaram, D., Khincha, H. P., & Vijaynarasimha, H. P. (2005): Artificial Neural Network and Support Vector Machine Approach for Locating Faults in Radial Distribution Systems. IEEE Transactions on Power Delivery, 20(2), 710– 721. doi:10.1109/tpwrd.2005.844307.

[31] Win-Bin Huang, & Wei-Tsung Su. (2015): Identity-based access control for digital content based on ciphertext-policy attribute-based encryption. 2015 International Conference on Information Networking (ICOIN). doi:10.1109/icoin.2015.7057862.

[32] Werner, J., & Westphall, C. M. (2016): A model for identity management with privacy in the cloud. 2016. doi:10.1109/iscc.2016.7543782.

[33] Wang, Y., Gou, Y., Guo, Y., & Wang, H. H. (2020): Construction of Audit Internal Control Intelligent System Based on Blockchain and Cloud Storage. 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI).

[34] Zhou, M., Zhang, R., Xie, W., Qian, W., Zhou, A. (2010): Security and Privacy in Cloud Computing: A Survey. In: Sixth International Conference on Semantics Knowledge and Grid (SKG), pp. 105–112.