Intrusion Detection Techniques Using Data Mining: Analysis

Rishi Kishan[1], Suraj Pal[2]
[1,2] *Department of computer science, GCET Gurdaspur, India*

***Abstract*- Advance computing is the most extreme basic part utilized component that has aimed to reserve the information by the machines which does not have adequate capacity abilities. The advance computing will be the system which enables machines to reserve the information over the capacity abilities of the specific machine. The security of user identity is on risk because many users interact through it. To handle data mining data security various types of intrusion detection techniques has been utilized. But these techniques utilized more space and less security, so in this paper we analysis various techniques used for intrusion detection.**

***Index terms*- Security, Intrusion detection, Byte level, space**

INTRODUCTION

In recent year the fast development of media transmission and web, data security turns out to be increasingly huge. Cryptography is way for ensuring secure data. Cryptosystems can be separated into two kinds, secret key cryptosystem, and public-key cryptosystem. The main sort (Secret key cryptosystem), utilizes a similar encipher key to encipher the original text and decode the figure content. For this reason, this kind is additionally called as a symmetric cryptosystem. In spite of the fact that secret key cryptosystem is effortlessly to actualize because of less calculation, a few disadvantages of this system is, excessively numerous keys, key conveyance issue, verification, and no repudiation issue. The imperative kind which is the public-key cryptosystem is created to take care of the issues of a cryptosystem, and RSA cryptosystem prevalent approach. As of late, information security is essential issue for public, private and guard associations as a result of the extensive misfortunes of illicit information get to. To shield secret information or data from unapproved access, illegal changes and propagation, different sorts of cryptographic strategies are utilized. One of these critical strategies is cryptography that is investigation of writing in secret frame and it is isolated into two sorts: symmetric and asymmetric cryptography.

Advance computing becomes the need of the hour nowadays but many experts argue about it [4]. Highly scalable services are given by the data mining. Users can utilize the services on a pay per use basis. Advance computing theoretically provides infinite resources but due to the growing number of users, practically services and resources becomes limited. The services and resources required to be distinguished on the basis of the scale of utilization along with cost. Although energy consumption and starvation problems nowadays, associated with advance computing but still improvement in services could lead to the better framework for concurrent users to access resources more than capable of the machine user hold and hence leads to more popularity and user community attracted towards the data mining.[5].

In the proposed system security mechanism along with redundancy handling mechanism is enforced for ensuring the quality of service. The attributes considered for evaluation are described as under

1.1 Attributes

Before some of the attributes will be defined, the term data mining should be explained. Advance computing used widely from a long time and provides an opaque framework where services are visible to the user but internal working is hidden[6]. Key attributes in advance computing are described in this section:

- Service-Based: Data mining main objective is to provide a service-oriented framework by hiding details and showing only the necessary features to the user. This mechanism is also termed as an abstraction.
- Scalable and Elastic: services associated with the data mining are not fixed. Services and be added

as and when required depending upon mass usage of services. In other words, the scalable environment is provided by advance computing[7]. Elasticity in the framework indicates resources are provided on different platforms accessible by multiple users at a time. In other words, concurrency is supported through the use of advance computing framework.

- Shared:[8] the resources are provided by the use of advanced computing environment in shared manner. Resource if free can be accessed by any number of resources provided resource is not exclusive in nature. The exclusive resource cannot be shared and that resource accessing required queue to be maintained.
- Metered by Use: multiple payment modes are supported by data mining infrastructure [9]. Services are accessed on a pay per use basis. Service provider and clients are bound by the service level agreement. The user need to pay for accessing the services mentioned within SLA. The problem however is, even if service is down for the period of time, still user is required to pay for that service.
- Uses Internet Technologies: Services are delivered to the user through the use of the internet. Protocol such as hypertext transfer protocol (HTTP), file transfer protocol (FTP), Terminal network (Telnet) etc. are used for this purpose[10].

Symmetric calculations are commonly viewed as quick and they are appropriate for preparing an expansive stream of information. A portion of the popular and proficient symmetric calculations incorporate two fish, Serpent, AES, Blowfish and IDEA. Also, there are non-specific calculations which offer an elective system for encipher. When all is said in done, hereditary calculations contain three essential administrators: multiplication, hybrid, and change. Then again, there are distinctive well known and productive deviated calculations including RSA, NTRU, and Elliptic curve cryptography.

Related Work

[11] Portrays the IBE procedure with re-appropriating calculation and furthermore offloads the key age activities to Key Update Data mining specialist organization. It additionally centers around the basic issues of character renouncement. It achieves reliable profitability for both count at PKG and private key size at the customer, User needs not to contact with PKG in the midst of key-update, so to speak, PKG is allowed to be separated resulting to sending the denial summary to KU-CSP, No ensured channel or customer confirmation is required in the midst of key-update among customer and KU-CSP.

[12]proposed the main MCL-PKE scheme without blending operations and gave its formal security. Our MCL-PKE takes care of the key escrow issue and disavowal issue. Utilizing the MCL-PKE conspire as a key building block, it proposed an enhanced way to deal with safely share sensitive information out in the public data mining. This approach supports quick denial and guarantees the classification of the information put away in an un trusted open data mining while authorizing the entrance control strategies of the information proprietor. The exploratory outcomes demonstrate the productivity of fundamental MCL-PKE scheme and enhanced approach for people in general data mining. Further, for various clients fulfilling similar access control arrangements, the enhanced approach perform just a solitary encipher of every datum thing and lessens the general overhead at the information owner.

[13] Proposed a variation of CP-ABE to effectively share the various hierarchical documents in distributed computing. The hierarchical documents are scrambled with an incorporated access structure and the cipher text parts identified with characteristics could be shared by the records. Thus both cipher text storage and time cost of encipher is saved. The proposed system has benefits that clients can decode all approval documents by figuring secret key once. Therefore, the time cost of decoding is also saved if the client needs to decode various documents.

[14]design a virtual encipher card framework that gives encipher card usefulness in virtual machines. It additionally settled a trusted chain amongst clients and enciphers cards in light of the composed protocols. The design of the virtual encipher card empowers the security and productivity of the encipher benefit. A usage examination shows that the effectiveness of the framework is similar to that of the native mode. Later on, it proceeds with the examination, trying to plan a virtual encipher cards bunch to help higher encipher speed and more reasonable similarity with virtualization.

[15]proposed a safe billing protocol for smart applications in distributed computing. It utilized homomorphism encipher by adjusting the Domingo-Ferrier's plan, which can perform different number arithmetic operations to fulfill smart grid billing necessities in a safe way. This plan keeps up the exchange off amongst security and versatility contrasted and other homomorphism plans that depend on either secure, yet inelastic in terms of arithmetic operations assortment. Additionally, it proposed an instrument that guarantees both security and integrity during correspondence between substances.

[16]propose a CP-ABE scheme that performs tedious matching operation can be outsourced to the data mining specialist organization, while the slight operations should be possible by clients. In this way, the calculation cost of the two clients and trusted specialist sides is limited. Besides, the proposed plot supports the capacity of keywords look which can enormously enhance correspondence effectiveness and further ensure the security and protection of clients. It is difficult to stretch out given KSF-OABE plan to help get to the structure represented by the tree in. [17]In this paper, based on contingent intermediary communicate re-encipher technology, an encrypted information sharing plan for secure distributed storage is proposed. The plan not just accomplishes communicate information sharing by exploiting communicate encipher, yet in addition accomplishes dynamic sharing that enables adding a client to and expelling a client from sharing gatherings dynamically without the need to change encipher open keys. Besides, by utilizing intermediary re-encipher innovation, this scheme empowers the intermediary (data mining server) to specifically share encoded information to the objective clients without the intercession of information owner while keeping information security, so significantly enhances the sharing execution. In the meantime, the rightness and the security are demonstrated; the execution is broke down, and the test comes about are appeared to confirm the possibility and the productivity of the proposed plot.

[18]proposed diagram encipher scheme just makes utilization of lightweight cryptographic natives, for example, pseudo-arbitrary capacity and symmetric-key encipher, instead of moderate homomorphism enciphers. Accordingly, the proposed graph encipher scheme is well disposed to a wide arrangement of graph information based distributed computing and edge registering applications, for example, interpersonal organizations, e-maps, criminal investigations, and so on. Contrast with graph anonymisation comes nearer from database group, the proposed system achieves higher security level as the chart itself is encoded and it doesn't make any suspicions on the sorts of attacks.

[19]discussed security enhancement mechanisms including a symmetric, public key and homomorphism cryptosystems to enable experts to comprehend encipher plans for information on distributed storage. AES is utilized as a part of most secure applications for information on distributed storage. Completely homomorphism encipher plans are promising for data mining condition, however, a long way from being useful due to their execution rate. Homomorphism assessment of AES has fascinating applications as a reasonable encipher conspire for information on distributed storage.

[21]proposed a lightweight accessible open key encipher (LSPE) conspire with semantic security for CWSNs. LSPE decreases countless calculation escalated operations that are received in past works; along these lines, LSPE has sought execution near that of some useful accessible symmetric encipher schemes.[22] proposed a protected data mining data encipher framework, named the Circulated Ecological Key (DENK in short), with which all records are encoded by one encipher key got from numerous coordinating keys which are keys gotten from approved clients' secret key keys and a believed PC's natural key.[23] proposed to present an effective and unquestionable FHE in light of another mathematic structure that is without commotion.

[24]described various way which is used in advance computing for data security. The information is put away on to an incorporated area called data centers having a substantial size of information storage. In this way, the customers need to put stock in the supplier on the accessibility and additionally information security. Before moving information into general society data mining, issues of security gauges and similarity must be tended to. Advance computing guarantees to change the financial matters of the server farm, yet before sensing and managed information move.

Author / Year	Method Used	Encryption technique	Complexity	Storage	Security Strength	Performance
R. Chen, Y. Mu, G. Yang, and F. Guo[1] 2015	BL-MLE	RSA Technique	High, In terms of iterations required	Repeated bits present hence high storage requirements	Least since no security slandered are followed	Better Handling of redundancy
R. Miguel[2] 2013	Homomorphic Security Process	Homomorphic Encryption	High, in terms of length of code	Bit level duplication showing high storage	Key is used hence medium security is present	Better data Handling
G. Zhu, X. Zhang, L. Wang, Y. Zhu, and X. Dong[3] 2011	Intelligent Back Up system	RSA Technique	Length of code ensure complexity	Multiple copies of data ensures high storage	No security slandered established	Back is improved
H. Nagarajaiah, S. Upadhyaya, and V. Gopal [4]2012	Embedded Processor Dedupliation	AES Algo	Calculations in AES makes it complex	Repeated data elimination techniques ensures less storage requirements	Private and public keys ensure high security	Security is improved
S. C. Satapathy, P. S. Avadhani, S. K. Udgata, and S.Lakshminarayana [5] 2016	Critical Infratructure is considered	None	Length of code in terms of LOC is less	Redundancy is handled hence space requirement is low	No security standards are used	Infrastructure capabilities are considered
J. J. Park, A. Zomaya, H.-Y. Jeong, and M. Obaidat[6] 2014	Frontier Technique	None	Multiple algorithms ensure complexity	Length of code required maximum storage	Security standards ensure high security	A new Technique where storage compression is considered
K. He, C. Huang, H. Zhou, J. Shi, X. Wang, and F. Dan[7] 2016	Public auditing technique	RSA algorithm	Complexity in terms of calculations is high	Storage requirements is high since LOG is maintained	Security Mechanisms are used	Public auditing for encrypted data with client-side Security Process in data mining storage
X. Li, J. Li, and F. Huang[8] 2015	Fuzzy Security Process	Fuzzy Algorithm	Low complexity since logical values in terms of 0 and 1 is used	Low storage requirements since result is stored in terms of Boolean values	Fuzzy storage has least security associated with it	A secure data mining storage system supporting privacy-preserving fuzzy Security Process
N. Christin and R. Safavi-Naini, Eds[10] 2016	Financial cryptography	Cost Based Encryption	Complexity in terms of calculations is high	Redundancy is handled hence space requirement is low	Security Mechanisms are used	A cost is a factor on which Security Process is considered
F. Rashid, A. Miri, and I. Woungang[11] 2016	Secure Enterprise Data Security Process	AES Algorithm	Complexity in terms of calculations is high	Redundancy is handled hence space requirement is low	Security Mechanisms are used	High security in data Security Process
W. K. Ng, Y. Wen, and H. Zhu[12] 2009	Private data mining data Security Process	None	Length of code in terms of LOC is less	Storage requirements are high in terms of multiple data	Security is low since no security slandered s are used	Only private data mining is considered

C. Wang, Q. Wang, K. Ren, and W. J. Lou[13] 2012	Ensuring Data Storage Security in Data mining Computing,	AES Encryption	Complexity in terms of calculations is high	Redundancy is handled hence space requirement is low	Security Mechanisms are used	Data security is high
Y. Yuan, X. Wu, and Y. Lu, Eds[14]2016	*Trustworthy Computing and Services*	None	None	None	None	Trust parameter is considered
C.-I. Fan, S.-Y. Huang, and W.-C. Hsu[15] 2015	Hybrid data Security Process in data mining environment	Cost based Encryption	Complexity in terms of calculations is high	Redundancy is handled hence space requirement is low	Security based standards are used	Hybrid Security Process is considered
F. Rashid, A. Miri, and I. Woungang[16] 2016	A secure data Security Process framework for data mining environments	Encryption based on cost	Complexity is high in terms of LOC	Storage requirements are high since code is complex	Data security mechanism are required	A security is provided
X. Zhang and J. Zhang[18] 2014	Data Security Process Cluster Based on Similarity-Locality Approach	Cluster based encryption is used	Complexity is high since cluster of information is present	High storage requirements in terms of clusters	Encryption standards ensure high	Cluster based approch is used
P. Puzio, R. Molva, M. Onen, and S. Loureiro[19] 2013	Data miningedup: Secure Security Process with Encrypted Data for Data mining Storage	Encryption based Secure Security Process	Complexity is high in terms of LOC	Storage requirements are high since code is complex	Data security mechanism are required	Secure Security Process is used
W. Leesakul, P. Townend, and J. Xu[20] 2016	Dynamic Data Security Process in Data mining Storage	AES Encryption	Complexity in terms of calculations is high	Redundancy is handled hence space requirement is low	Security Mechanisms are used	Dynamic Security Process is used
T. Johansson and P. Q. Nguyen, Eds[21] 2015	Cryptography advancement is used	Cipher Text is used	Low complexity in terms of lease code utilized	Redundancy is handled hence low storage requirements	Security Mechanisms are used	Cryptography is used
V. Inukollu, S. Arsi, and S. Ravuri[22] 2015	Security Issues Associated With Big Data in Data mining Computing	Encryption based on DES	Complexity in terms of calculations is high	Redundancy is not handled hence storage requirement is high	Security mechanism prevent malicious attacks	Fast Encryption is used

Table 1: Comparative analysis of literature

To resolve the problem with the existing literature proposed literature present efficient solution. The encipher mechanism with the redundancy handling mechanism can be proposed.

CONCLUSION

Advance computing not only provides the resources to the users but also give a big challenge of security. There are securities requirements for both users and providers but sometimes it may conflict in some way. Security of user depends upon trusted computing and cryptography. In our review paper some issues

related to data location, security, storage, availability, and integrity. Establishing trust in Data mining security is the biggest requirement. The problems and corresponding solutions may require further investigation in terms of key size and complexity. The complexity of the key can be further enhanced by the use of a pseudo-random number generator within the key generation phase.

By incorporating complex key structure, data mining performance and user interaction can be further enhanced using complex keys and by reducing attacks.

REFERENCES

[1] F. Sabahi, "Advance computing Security Threats and Responses," pp. 245–249, 2011.

[2] X. Wu, R. Jiang, and B. Bhargava, "On the Security of Data Access Control for Multiauthority Data mining Storage Systems," pp. 1–14, 2015.

[3] J. Aikat et al., "Rethinking Security in the Era of Advance computing," no. June 2017.

[4] K. Hwang, X. Bai, Y. Shi, M. Li, W.-G. Chen, and Y. Wu, "Data mining Performance Modeling with Benchmark Evaluation of Elastic Scaling Strategies," IEEE Trans. Parallel Distrib. Syst., vol. 27, no. 1, pp. 130–143, Jan. 2016.

[5] T. H. Noor, Q. Z. Sheng, L. Yao, S. Dustdar, and A. H. H. Ngu, "Data miningArmor: Supporting Reputation-Based Trust Management for Data mining Services," IEEE Trans. Parallel Distrib. Syst., vol. 27, no. 2, pp. 367–380, Feb. 2016.

[6] M. Armbrust et al., "A view of advance comsputing," Commun. ACM, vol. 53, no. 4, p. 50, 2010.

[7] R. Buyya, C. S. Yeo, and S. Venugopal, "Market-oriented advance computing: Vision, hype, and reality for delivering IT services as computing utilities," Proc. - 10th IEEE Int. Conf. High Perform. Comput. Commun. HPCC 2008, pp. 5–13, 2008.

[8] S. J. Nirmala, N. Tajunnisha, and S. M. S. Bhanu, "Service provisioning of flexible advance reservation leases in IaaS data minings," vol. 3, no. 3, pp. 154–162, 2016.

[9] M. Marwan, A. Kartik, and H. Ouahmane, "Secure Data mining-Based Medical Image Storage using Secret Share Scheme," 2016.

[10] D. V. Dimitrov, "Medical internet of things and big data in healthcare," Healthc. Inform. Res., vol. 22, no. 3, pp. 156–163, 2016.

[11] J. Li, J. Li, X. Chen, C. Jia, W. Lou, and S. Member, "Identity-based Encipher with Outsourced Revocation in Advance computing," pp. 1–12, 2013.

[12] S. Seo, M. Nabeel, and X. Ding, "An Ef fi client Certi fi cateless Encipher for Secure Data Sharing in Public Data minings," pp. 1–14, 2013.

[13] S. Wang, J. Zhou, J. K. Liu, J. Yu, and J. Chen, "An Efficient File Hierarchy Attribute-Based Encipher Scheme in Advance computing," vol. 6013, no. c, pp. 1–13, 2016.

[14] D. Xu, C. A. I. Fu, G. Li, and D. Zou, "Virtualization of the Encipher Card for Trust Access in Advance computing," vol. 5, 2017.

[15] A. Alabdulatif, H. Kumarage, I. Khalil, M. Atiquzzaman, and X. Yi, "Privacy-preserving data mining-based billing with lightweight homomorphic encipher for sensor-enabled smart grid infrastructure," IET Wirel. Sens. Syst., vol. 7, no. 6, pp. 182–190, 2017.

[16] J. Li, X. Lin, Y. Zhang, and J. Han, "KSF-OABE Outsourced Attribute-Based Encipher with Keyword Search Function for Data mining Storage," vol. 1374, no. c, pp. 1–12, 2016.

[17] L. Jiang, D. Guo, and S. Member, "Dynamic Encrypted Data Sharing Scheme Based on Conditional Proxy Broadcast Re-Encipher for Data mining Storage," vol. 5, 2017.

[18] C. Liu, S. Member, L. Zhu, J. Chen, and S. Member, "Graph Encipher for Top-K Nearest Keyword Search Queries on Data mining," vol. 3782, no. c, pp. 1–11, 2017.

[19] C. Song, Y. Park, J. Gao, S. K. Nanduri, and W. Zegers, "Favored Encipher Techniques for Data mining Storage," pp. 267–274, 2015.

[20] N.Veeraragavan, "Enhanced Encipher Algorithm (EEA) for Protecting Users' Credentials in Public Data mining."

[21] P. Xu, S. He, W. Wang, W. Susilo, and H. Jin, "Lightweight Searchable Public-key Encipher for Data mining-assisted Wireless Sensor Networks," IEEE Trans. Ind. Informatics, vol. XX, no. XX, pp. 1–12, 2017.

[22] K. L. Tsai et al., "Data mining encipher using distributed environmental keys," Proc. - 2016 10th Int. Conf. Innov. Mob. Internet Serv.

Ubiquitous Comput. IMIS 2016, pp. 476–481, 2016.

[23] A. El-yahyaoui, "A verifiable fully homomorphic encipher scheme to secure big data in advance computing," 2017.

[24] G. Thomas, "Advance computing security using encipher technique," pp. 1–7.