Advances and Applications in Mathematical Sciences
Volume 22, Issue 4, February 2023, Pages 901-912
© 2023 Mili Publications, India

NAAS: AN AI-BASED AMALGAMATED APPROACH FOR SECURING E-COMMERCE TRANSACTIONS

NISHANT AGNIHOTRI[1] and AMAN KUMAR SHARMA[2]

[1]PhD Research Scholar
[2]Professor
Department of Computer Science
Himachal Pradesh University, Shimla, India
E-mail: er.nishantagnihotri@gmail.com
sharmaas1@gmail.com

Abstract

The security of the data is highly important as far as external threats to online transactional data are concerned. There require some suitable measures to protect the data. The level of the threat measure depends upon the level of the threat to the data. These threats have been found quite extensive during the Covid on e-Commerce Transactions. Therefore, a secure mechanism is required for the such online transaction. Keeping in view, the above challenges, an AI-based mechanism will be used for selecting the level of complexity of the data based on a few parameters. One major parameter will be the field to which data belongs. The data relating to financial institutions will be considered to be of the highest level of complexity and data of general use will be considered of the lowest complexity. If the AI system will predict data to be of higher or medium level complexity then the Enhanced K anonymity technique will be applied to the data and if data is identified as of the lowest complexity, then the Enhanced-Symmetric Key Cryptography Algorithm (E-SKA) technique will be applied. This proposed model is named NAAS. The proposed automated model has attained a 14.03 per cent improvement over the manual model.

1. Introduction

E-commerce was one of the prominently used mediums while covid for society to fulfil their purchasing needs. That also goes without saying that e-commerce has also faced a plethora of challenges during that era. Now, it has become of utmost importance for researchers to overcome those issues, primarily the issue of secure financial transactions.

2020 Mathematics Subject Classification: 68T07.

Keywords: Cryptography, Machine Learning, Classification, Security, Anonymity, Security.
Received January 23, 2023; Accepted February 1, 2023

Post covid, people have more evolved towards offline shopping, one of the major reasons behind this change of trend was the security concerns behind e-commerce transactions. Therefore, to bring back the era of e-commerce to its peak along with security, certain better security solutions are required. To secure the data flowing over any medium, cryptography is used as a primary technique. The two main categories of cryptographic algorithms are symmetric encryption algorithms and asymmetric cryptographic algorithms. (Simmons, [18]), which are also known as public key cryptographic algorithms. The earliest and most prominent encryption method is symmetric encryption, which involves applying a special hidden key to a message to upend the information in an unique way. The secret key might be a number, a phrase, or just a string of random characters. (Hammad et al., [8]). It may be as easy as moving each letter a few positions in the alphabet. All messages that utilize this key can be encrypted and decrypted as long as the sender and recipient are both aware of it. Secret key exchange over the Internet or a big network presents a challenge since it is difficult to keep them out of the wrong hands. Any individual with the necessary secret key can decode the communication. (Simmons, [18]). So, the distribution of keys securely over a channel is one of the important issues related to symmetric key cryptography.

Furthermore, we may utilize a different technique called public key cryptography to resolve such a problem. (Chandra et al., [5]) in which there are two related keys, a key pair, and sender gets access to the public key without charge. The other individual is the only one who has access to a second, private key. The only way to decode a communication (whether it is text, a binary file, or a document) that has been encrypted using the public key is to use the same method and the corresponding private key. Only the matching public key may be used to decode any communications gets encrypted using the private key. As a result, you can transmit public keys over the Internet without concern. However, asymmetric encryption has the drawback of being slower than symmetric encryption. The computing power needed to encrypt and decode the message’s content is significantly higher (Chandra et al., [5]; Hammad et al., [8]).

It is necessary to learn other public keys in asymmetric encryption, and there must be a method to learn other public keys. Apart from this, anonymization (Wallace, [25]) of data is another prominent method to secure the exploitation of data (Tina Coffelt, [23]), belonging to any individual. This method changes the form of data instead of cyphering it (Wiles et al., [26]). The form is made in such a way that the prediction of data being transferred, becomes quite tedious for the intruder to interpret the data. This technique in other words keeps the data limited to the knowledge of the one who is entitled to see or understand it.

2. Literature Survey

(Patro et al., [12]) highlighted that although e-commerce generated a plethora of options for industry and at the same time it also creates new dangers and weaknesses, such as hacking and security concerns. As a result, it is a crucial technological and organizational need for a successful and effective digital payment transaction activity. A match of algorithmic and technological solutions is necessary due to the constant technological and economic development. The following subjects were examined in this paper: An overview of e-commerce security, various processes for placing an order, Security purposes in E-commerce, significant security challenges in E-commerce, suggestions for secure online shopping, and so on.

(Ji, [9]) figured out that the introduction of e-commerce has given Chinese manufacturing development as well as social and economic growth a new driving force. Even though e-commerce has dramatically widened its horizon, many risks have been brought about by information security problems of cyber exchanges which have not been strongly guaranteed in terms of legislation or technology, eroding consumers' trust and limiting the development of China's e-commerce. They maintained that to guarantee the security of information, a secure system was necessary.

(Smith and Shao, [20]) have looked at how this core principle has changed over time, emphasizing how social and technological developments have made the right to secrecy more difficult to uphold, and have investigated current computer systems that do so for e-commerce. Our study explores the privacy protections issues for prevailing e-commerce implementations, illustrates that historically, confidentiality has been primarily driven by our comprehension of privacy as well as technical progress, and recognizes orientations for the prospective creation of successful privacy-enhancing technologies.

(Chinnasamy et al., [6]) Have stated that Cryptography is used to solve the security problem and attain CIA property (confidentiality, integrity, and availability). The most effective Cryptography is a technique for maintaining high degrees of data storage and transmission security. Traditional symmetrical and asymmetrical patterns have certain drawbacks. To address this, a novel hybrid approach with high levels of data security and secrecy will be introduced. In this article, a hybrid method is constructed by combining ECC with Blowfish. The suggested technique offers excellent security and confidentiality of patient data when the efficiency of the hybrid system is evaluated in comparison to the present hybrid approach. Hybrid cryptography is employed to overcome the drawbacks of both symmetric and asymmetric encryption.

(Sidi et al., [17]) have asserted in their study that steganography and encryption, although having certain differences, are both used to help protect data and prevent the revelation of private conversations. Therefore, in this context, a hybrid shortcut approach centred on the LSB method and the RSA and Caesar Cipher encryption techniques is proposed. The trial's outcomes show how secure and excellent the system's cover image is and how it is virtually indistinguishable from the original image.

(Kumar et al., [10]) a multilayer security approach that was using cryptography for cloud computing was developed by the authors. The model combines symmetric and asymmetric key cryptographic methods. By providing several levels of encryption and decryption at the transmitter and receiver sides, respectively, the employment of the Data Encryption Standard (DES) and RSA in this scenario strengthens the security of cloud storage. The recommended approach was put into practice using Java and the cloud simulator application cloudsim. This methodology improves text file upload and download speeds while boosting data security to the maximum level achievable in contrast to the present system. Also, (Madaan et al., [11]; Sharma et al., [16]) have given a hybrid technique for bluffing intruders.

(Ahmad and Garko, [1]) Stated that in the form of Electronic Health Records (EHR), which are always accessible online, medically related data may be preserved. It includes information on the patient as well as ray x pictures, scan images, therapeutic procedures, and prescriptions for medications. The issue with all of these sensitive documents is how to keep them safe and who may read and access the data. They created a safe cloud storage system for medical data using a hybrid cryptographic method to address this issue. Keys are encrypted using an asymmetric method, while the data are encrypted using a symmetric technique. The outcomes unequivocally demonstrate that their approach offers more security than any existing hybrid algorithms.

3. Motivation

The most important and susceptible component in today's world is data, whose security is given top emphasis. The hazards involved in this statement are the main justification. Data may be personal, meaning it may be connected to a specific person, and illegal access to such data is never imaginable. As a result, data might include information about a company, bids, transactions, and national secrecy, to mention a few. Unauthorized access to such data can result in many serious problems, including impersonation, bullying, identity theft, and a danger to the integrity of a country. There is a requirement for an efficient data security solution that could protect data from such problems. All the operations on data in the twenty-first-century use technology for a variety of purposes. Over the open network, they are sharing data of different severity. The information provided over a public channel is extremely susceptible to both internal and external dangers (Singh and Tiwari, [19]) (Sasubilli and R, [15]). To safeguard protected data over a public network using an effective approach, some major efforts are needed. It will be necessary to choose a method that can adequately protect the privacy and security of the data within a limited system resource. In light of this, a higher-efficiency recommended model will be applied to the data. The security of data is of utmost priority and cannot be compromised (Sudha and Monica, [21]). In a plethora of instances, it has been observed that a breach in data occurs (Radhika Parashar, [13]) (Reuters, 2021). Data is compromised which leads to financial losses and depression for the concerned individual. Hence, there is a need of having a better security mechanism (Tayal et al., [22]).

4. Proposed Technique

The proposed technique is based on building a combined model (Udendhran, [24]) having few artificial intelligent components being used with AES (Akkar and Giraud, [2]) and K-anonymity (El Emam et al., [7]) in their enhanced version. These AI-based components will be able to decide on their own. The enhanced k-Anonymity technique (EK-A) and enhanced symmetric key cryptography algorithm are two algorithms. One has to be selected based on the level of complexity of the data supplied in the algorithm. There are three levels of complexity of the data: Type 1, Type 2, and Type 3. The AI-based mechanism that has used a supervised learning approach using supervised learning will be used for selecting the level of complexity of the data based on a few parameters. One major parameter will be the field to which data belongs. The data relating to financial institutions will be considered as of the highest level of complexity and data of general use will be considered of the lowest complexity. If the AI system will predict data to be of higher or medium level complexity based on machine learning (Bell, [4]) then the EK-A technique will be applied to the data and if data is identified as of the lowest complexity, then the E-AES technique will be applied. This proposed model is identified as highly efficient and effective for the data.

5. Flowchart of Proposed Model

Figure 1 has shown the model that has been adopted for the said work.

Terms Used

$T1$ - Security: Improved Optimal Lattice Anonymity with Binary Tree

$T2$ - Security: Enhanced Symmetric Key Cryptography Algorithm

$T3$ - Security: A hybrid Mode of $T1 + T2$

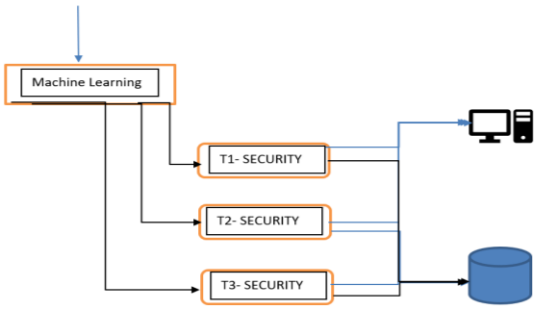

Figure 1. Proposed Model.

Figure 1 depicts that the system shall, first of all, segregate the data based on the level of severity, for the said purpose machine learning technique kNN (Balajee et al., 2020) has been opted based on several experiments. Also, kNN has been used for several clinical information management and is one the best classifier for data (Sharma et al., [16]). Thereafter, based on the level of severity, data shall go through improved optimal lattice *k*-anonymity with a binary tree. The second level of severity shall go through enhanced symmetric key cryptography and finally, the extreme severity of data shall pass through the hybridization of both techniques. This secured data can go to the user directly or can be stored in the local database.

6. Environment

Table 1 shows the experimental setup for executing an amalgamated AI-based data security system. The environment is suitable for running the proposed model on datasets of different sizes.

Table 1. Experimental Setup.

Name	Specification
Processor	I3 Processor
RAM	4 GB
Operating system	Windows 10
Simulator	Matlab 2015

Dataset: "Online retail", this dataset has been acquired from the free data repository source Kaggle on the link "https://www.kaggle.com/datasets/vijayuv/onlineretail?resource=download". The nature of this dataset is Alphanumeric and this contains 8 columns and the number of rows is 541910.

7. Results and Discussions

The result of the proposed model is evaluated in two parameters, one is the time and the second is the memory space requirement. The proposed AI-based system is using machine learning model to check the complexity of the dataset and select the most suitable security algorithm for a specific dataset. It will not only save time and memory but would also generate a highly efficient system in terms of security.

For the said purpose, a supervised machine learning technique has been chosen based on certain parameters like

- **Accuracy.** This is one of the metrics used to assess a machine learning model that allows us to determine how accurately the data is classified. Our formula for determining accuracy is as follows:

$$Accuracy= \frac{TP + TN}{P + N} \tag{1}$$

- **Precision.** Precision will be shown through precision. Formula for calculating accuracy is as follows:

$$\text{Precision}= \frac{TP}{TP + FP} \tag{2}$$

- **Recall.** It'll let us know how complete our model is. This formula is used to determine the recall:

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

- **F1-Score.** The harmonic mean of accuracy and memory yields the precision and recall results in balance. The following formula is used to determine the $f1$-score:

$$\text{Scoreof } f1 = \frac{2(\text{Precision Recall})}{(\text{Precision+ Recall})} \tag{4}$$

Table 2. Comparison of time taken by the techniques in seconds.

Complexity	Datasets	With an Enhanced symmetric key cryptographic algorithm	With Optimal Lattice Anonymization with a binary tree	Manual Selection of hybridization	Proposed Automated Model	% Improvement
High	Retail (whole dataset)	131684.13	271.3343	153979.98	131955.46	14.30

Table 2 has displayed the execution time taken by the system when dealing with the whole dataset. The comparison has been made in the time that shall be taken by the enhanced symmetric key cryptography is 131684.13 seconds, the time that shall be taken by improved optimal lattice k-anonymity with a binary tree is 271.33 seconds, the time taken at the time of opting manual selection of hybridization is 153979.98 seconds and finally the time, the proposed method will take is 131955.46 seconds. The same has shown a 14.30% improvement.

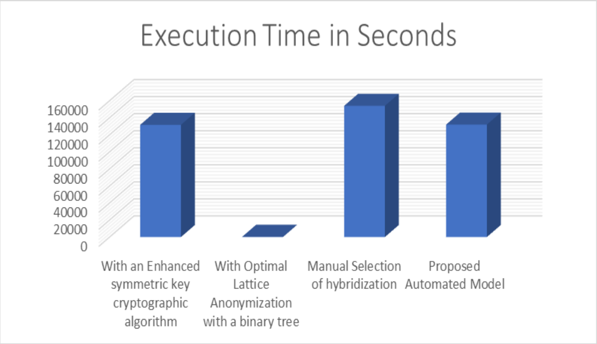

Figure 2. Comparison of Execution time.

Figure 2 shows the comparison of time for the existing and proposed algorithms. The proposed algorithm is highly efficient compared to the existing algorithm. The system has attained 14.30%.

8. Conclusion and Future Work

The proposed AI-based security scenario is highly efficient and useful for providing security to the data shared over the public channel. It provides security to the data such that automatic selection of the specific algorithm based on the complexity of the data. The proposed hybrid AI-based model is having higher efficiency in terms of the lower time of processing and also lower memory space requirements. The proposed algorithm is useful for those applications where the security of the data is primarily important. The proposed technique is having the basic aim to identify various types of E-commerce data while flowing over the web and then ensuring the appropriate level of security. However, as of now, this system has got supervised learning approach to classify the labelled data, but in future, this system can be evolved better with a semi-supervised learning approach and align it for securing E-Commerce Transactions from quantum computing attacks based on the level of the threats.

References

[1] S. A. Ahmad and A. B. Garko, Hybrid Cryptography Algorithms in Cloud Computing: A Review, 2019 15th International Conference on Electronics, Computer and Computation (ICECCO), (2019), 1-6. https://doi.org/10.1109/ICECCO48375.2019.9043254

[2] M. L. Akkar and C. Giraud, An Implementation of DES and AES, Secure against Some Attacks (2001), 309-318. https://doi.org/10.1007/3-540-44709-1_26

[3] Balajee Maram, G. Padmapriya and Aravapalli Rama Satish, A framework for performance analysis on machine learning algorithms using Covid-19 dataset, Advances in Mathematics: Scientific Journal 9(10) (2020), 8207-8215. https://doi.org/10.37418/amsj.9.10.50

[4] J. Bell, What Is Machine Learning? In Machine Learning and the City, Wiley (2022), (207-216). https://doi.org/10.1002/9781119815075.ch18

[5] S. Chandra, S. Paira, S. S. Alam and G. Sanyal, A comparative survey of Symmetric and Asymmetric Key Cryptography, 2014 International Conference on Electronics, Communication and Computational Engineering (ICECCE) (2014), 83-93. https://doi.org/10.1109/ICECCE.2014.7086640

[6] P. Chinnasamy, S. Padmavathi, R. Swathy and S. Rakesh, Efficient Data Security Using Hybrid Cryptography on Cloud Computing, Inventive Communication and Computational Technologies: Proceedings of ICICCT (2021), 537-547. https://doi.org/10.1007/978-981-15-7345-3_46

[7] K. El Emam, F. K. Dankar, R. Issa, E. Jonker, D. Amyot, E. Cogo, J. P. Corriveau, M. Walker, S. Chowdhury, R. Vaillancourt, T. Roffey and J. Bottomley, A globally optimal k-anonymity method for the de-identification of health data, Journal of the American Medical Informatics Association 16(5) (2009), 670-682. https://doi.org/10.1197/jamia.M3144

[8] B. T. Hammad, A. M. Sagheer, I. T. Ahmed and N. Jamil, A comparative review on symmetric and asymmetric DNA-based cryptography, Bulletin of Electrical Engineering and Informatics 9(6) (2020), 2484-2491. https://doi.org/10.11591/eei.v9i6.2470

[9] Q. Ji, Study on information security issues of e-commerce, IOP Conference Series: Materials Science and Engineering 452 (2018), 032050. https://doi.org/10.1088/1757-899X/452/3/032050

[10] S. Kumar, G. Karnani, M. S. Gaur and A. Mishra, Cloud security using hybrid cryptography algorithms, 2021 2nd International Conference on Intelligent Engineering and Management (ICIEM) (2021), 599-604. https://doi.org/10.1109/ICIEM51511.2021.9445377

[11] V. Madaan, D. Sethi, P. Agrawal, L. Jain and R. Kaur, Public Network Security by Bluffing the Intruders Through Encryption Over Encryption Using Public Key Cryptography Method, Advanced Informatics for Computing Research: First International Conference (2017), 249-257. https://doi.org/10.1007/978-981-10-5780-9_23

[12] S. P. Patro, N. Padhy and R. Panigrahi, Security Issues over E-Commerce and their Solutions, International Journal of Advanced Research in Computer and Communication Engineering 5(12) (2016), 81-85. https://doi.org/10.17148/IJARCCE.2016.51216

[13] M. K. Sasubilli and R. Venkateswarlu, Cloud Computing Security Challenges, Threats and Vulnerabilities, 2021 6th International Conference on Inventive Computation Technologies (ICICT) (2021), 476-480. https://doi.org/10.1109/ICICT50816.2021.9358709

[14] M. Sharma, S. Kumar Singh, P. Agrawal and V. Madaan, Classification of clinical dataset of cervical cancer using KNN, Indian Journal of Science and Technology 9(28) (2016). https://doi.org/10.17485/ijst/2016/v9i28/98380

[15] E. V. Sidi, I. Diop and K. Tall, A New hybrid approach of Data Hiding Using LSB Steganography and Caesar cipher and RSA algorithm (S-ccr), 2022 International Conference on Computer Communication and Informatics (ICCCI), (2022), 1-4. https://doi.org/10.1109/ICCCI54379.2022.9740979

[16] G. J. Simmons, Symmetric and Asymmetric Encryption, ACM Computing Surveys 11(4) (1979), 305-330. https://doi.org/10.1145/356789.356793

[17] N. Singh and P. Tiwari, SQL Injection Attacks, Detection Techniques on Web Application Databases (2022), 387-394. https://doi.org/10.1007/978-981-19-1122-4_41

[18] R. Smith and J. Shao, Privacy and e-commerce: a consumer-centric perspective, Electronic Commerce Research 7(2) (2007), 89-116. https://doi.org/10.1007/s10660-007-9002-9

[19] M. Sudha and M. Monica, Enhanced security framework to ensure data security in cloud computing using cryptography, In Advances in Computer Science and its Applications 1(1) (2012), 32-37. www.worldsciencepublisher.org

[20] N. Tayal, R. Bansal, S. Dhal and S. Gupta, A novel hybrid security mechanism for data communication networks, Multimedia Tools and Applications 76(22) (2017), 24063-24090. https://doi.org/10.1007/s11042-016-4111-x

[21] Tina Coffelt, Confidentiality and Anonymity of Participants, In the SAGE Encyclopedia of Communication Research Methods, SAGE Publications, Inc. (2017). https://doi.org/10.4135/9781483381411.n86

[22] R. Udendhran, A hybrid approach to enhance data security in cloud storage, Proceedings of the Second International Conference on Internet of Things, Data and Cloud Computing (2017), 1-6. https://doi.org/10.1145/3018896.3025138

[23] K. A. Wallace, Anonymity, Ethics and Information Technology 1(1) (1999), 21-31. https://doi.org/10.1023/A:1010066509278

[24] R. Wiles, G. Crow, S. Heath and V. Charles, The management of confidentiality and anonymity in social research, International Journal of Social Research Methodology 11(5) (2008), 417-428. https://doi.org/10.1080/13645570701622231

[25] Radhika Parashar, Data of Over 3 Million CoinMarketCap Users Put Up for Sale on Forums, NDTV Gadgets 360 (2021, October).

[26] Reuters, (2021, September 23), UK Suspends Official For Breach Exposing 250 Afghan Interpreters' Details, NDTV.