A Review on Machine Learning Approaches for Cryptography

Vinod Shanataram Mahajan[1], Hitendra D. Patil[2]

Kavayitri Bahinabai Chaudhari North Maharashtra University Jalgaon[1]
SSVPS's B.S. Deore College of Engineering, Dhule, India[2]

Corresponding author: Vinod Shanataram Mahajan, Email: vinodsm83@gmail.com

The exponential growth of technologies and online applications in the last decade attracts researchers toward using Machine Learning (ML) to provide information security. In today's world, ML has proven itself by performing outstandingly on various issues and challenges of real-world applications. It has a vital domain of applications, one of which is cryptography. Cryptography is used to build a secure cryptosystem to provide information security. Although, the use of ML in cryptography is not novel. In the last decade, many notable machine-learning approaches have been applied in cryptography to achieve significant information security. This review paper looks at emerging research into applying ML approaches to cryptography. Moreover, we discuss the key concepts of ML and cryptography, summarize them, and list the future research directions.

Keywords: Cryptography, Cryptosystem, Machine Learning, Security

2023. In Satyasai Jagannath Nanda & Rajendra Prasad Yadav (eds.), *Data Science and Intelligent Computing Techniques*, 301–306. Computing & Intelligent Systems, SCRS, India. https://doi.org/10.56155/978-81-955020-2-8-26

1 Introduction

The swift advancements in web technology and the widespread adoption of these tools are speedy, producing an enormous amount of valuable data. One of the numerous real-world issues and applications arising from this is protecting sensitive information. In the past, a wide variety of symmetric and asymmetric cryptographic methods were developed in order to address this concern. Similarities between cryptography and machine learning drive the adoption of machine learning techniques, as described in [1]. In addition to these areas of study, machine learning has many applications, some of which are listed in [2], including learning association, classification, and regression, amongst others. In recent years, there has been a tremendous expansion in the application fields of machine learning. Machine learning can be divided into five categories: supervised machine learning, unsupervised machine learning, semi-super-vised or semi-unsupervised machine learning, and reinforcement learning [3, 4].

This paper reviewed machine learning approaches used in cryptography in recent years and investigated where cryptography research is headed. The remaining parts of this article are structured as follows. Section II covers the machine learning approaches to cryptography. The findings of this work are outlined in detail in Section III. The final section concludes this work and also suggests new directions for investigation.

2 Machine Learning Approaches For Cryptography

This section presents the recent literature review on machine learning approaches to build secure cryptographic systems.

Mishra et al. [5] presented the application of a Genetic Algorithm in cryptography. In this paper, 192-bit public and private keys were generated with the help of proposed PKC using Genetic algorithms. The keys generated are random, enhancing the key strength and security. They used C++ and Microsoft Excel for implementation and analysis.

A novel symmetric key technique that is based on a counter-propagation neural network was proposed by Kumar et al. [6]. The Input, Khonen, and Grossberg layers are the fundamental components of the CPN. While encrypting a message, the plaintext is first transformed into ASCII, then binary, and finally fed into the Khonen layer to produce the cypher text and the target value. The cypher text and the target value are entered into the Grossberg layer of the decryption algorithm. The binary output of the layer is then translated to ASCII, and finally, plaintext is produced. Experiments were carried out in MATLAB, and the results reveal that the performance is unaffected by either a low or a high bandwidth.

The neural key exchange approach Sahana et al. [7] presented was based on synchronizing the tree parity machines the sender and receiver used. They synchronized the TPM with the help of the Anti-Hebbian Learning Rule by adjusting the weights and encrypting the email using the produced secret key. They utilized the formula (2L+1) (K*N) to determine the length of the key. Where N is the value of the input vectors, K is the number of concealed neurons, and L is the depth of the synaptic connection. Moreover, an increase in the values of N and L can increase the key's strength. The suggested approach may be vulnerable to geometric and genetic attacks.

Atee et al. [8] developed a neural network-based sub-key generation scheme for a secure cryptosystem. It is based on an extreme learning machine for one hidden neural network layer with 100 neurons. This scheme in ELM training parallels generates 10 independent sub-keys using 10 internal rounds and then makes a final key of length 120. The experimental results showed that >99% sensitivity was achieved with key space 2120. In future, this scheme can be compared with Advanced Encryption Standards (AES) and other sub-key generation algorithms.

Sharma et al. [9] used a technique based on a deep neural network in steganography. It has four components; the first two components, the Prep Layer and the Hiding Layer are used to hide the image and act as a sender. The next two components reveal the layer and Decrypt layer used to retrieve the original image and act as the receiver block. They used an Adaptive Moment Estimation optimizer to minimize the network loss. To perform experiments, the Flickr30k dataset, ML Engine from Google Cloud Platform and libraries like Tensor-Flow, Keras, Matplotlib, NumPy, SciPy, etc., were used and achieved 90% accuracy. In future, Modern cryptography techniques can be used with this scheme to perform encryption if lossless neural networks are achieved. Other steganography techniques will also be implemented using NN.

V. Naveena et al. [10] proposed generating secret keys using tree parity machine neural networks. The concept is mutual learning for synchronization between sender and receiver; then, during synchronization, the same tree parity machine and PRNG are used by both parties to generate and exchange secret keys. In this paper, no specific method to check the strength of the key is not mentioned in the paper.

An image encryption technique known as ChaosNet was presented by Thoms et al. [11], and it was developed based on the Lorenz chaotic system and chained finite field transformation layers. In order to generate an encryption algorithm, this approach is paired with the whole-picture permutation method. Images of highway traffic were obtained from SERSU and used for encryption. The experiments' findings demonstrated that the proposed algorithm's performance is superior to that of other chaos-based algorithms in most evaluation metrics, including information entropy, differential attack analysis, low entropy encryption analysis, key sensitivity, pixel correlation, and occlusion attack analysis. The author notes that additional work might be done to improve the efficiency and performance of the algorithm and the hardware implementation.

Li et al. [12] introduced a technique of symmetric encryption for text encryption called TEDL based on deep learning. The proposed approach is broken down into two stages: the communication preparation stage and the communication process stage. In communication preparation, two public corpora, the Chinese Wikipedia corpus and the English Wikipedia corpus, were used. The sender and receiver modify these corpora with the help of a key and get a synthetic corpus. Word vector tables were constructed using deep learning models trained on the synthetic corpus. Time-varying codebooks were obtained by processing word vector tables and the SHA 256 function.

During the communication process, both the sender and receiver use a codebook to encrypt and decode messages, respectively. Following the transmission of a message, both parties will update their respective copies of the code book. The author did an in-depth examination of the suggested approach by employing a variety of factors, such as security, recovery, the amount of time consumed by brute force, frequency analysis, correlation, sensitivity analysis, efficiency analysis, and generality analysis.

They also claimed that in the not-too-distant future, Objects that come in various formats, such as binary numbers, words, photos, videos, or even multimodal information, could be encrypted using TEDL. Expansion is possible for encrypted items as well as models.

Table 1 summarises the literature reviewed in cryptography and machine learning regarding machine learning approaches, types of cryptosystems, evaluation metrics, and future scope.

Table 1. Summary of reviewed literature in the field of cryptography and machine learning

Ref. No	ML Approach	Type of Cryptosystem	Purpose	Evaluation matrices	Future Scope
[5]	GA	Asymmetric Cryptography	Key Generation	Randomness and fitness of key	Not Mentioned
[6]	CPNN	Symmetric Cryptography	Text Encryption and Decryption	Not Mentioned	Not Mentioned

[7]	TPM NN	Symmetric	Key Generation	Key strength	Not mentioned
[8]	ELM and NN	All	Key Generation	Key Space and Sensitivity	It can be compared with AES and other sub-key generation algorithms.
[9]	NN	Steganography	Image Hiding	Accuracy	AES and DES can be used if a lossless neural network is achieved. Also, Other steganography encryption can be integrated with NN
[10]	TPM NN	Symmetric Cryptography	Key Generation	Not Mentioned	Not Mentioned
[11]	NN	Steganography	Encryption and Decryption	key sensitivity, Entropy, low entropy encryption, differential attack, pixel correlation, occlusion attack analysis	optimization of algorithm and hardware implementation
[12]	DL	Symmetric Cryptography	Text Encryption and Decryption	Safety, revival, Measures of frequency, correlation, sensitivity, and efficiency, Generality analysis	Expansion is possible for both encrypted objects and models.

Table 2 summarizes the different datasets and tools used and refereed by authors. As shown in below table dataset is not mentioned in [5, 6, 7, 8] and tools or programming language is not mentioned in [10, 11, 12].

Table 2. Summary of datasets and tools used in reviewed literature in this section

Ref. No.	Dataset	Tools/ Prog. Language
[5]	Not Mentioned	C++ and Microsoft Excel
[6]	Not Mentioned	MATLAB
[7]	Not Mentioned	JDK 1.7
[8]	Not Mentioned	Not Mentioned

[9]	31,783 images From Flickr30k public dataset	ML-Engine from Google Cloud Platform and Python 3.5.3
[10]	Not Mentioned	Not Mentioned
[11]	Highway traffic images from SERSU	Not Mentioned
[12]	Chinese (1.3 GB) and English (600MB) Wikipedia corpus	Not Mentioned

3 Discussion

In section II, we studied notable works in cryptography based on machine learning approaches. Now, in this section, we illustrate and discuss notable points.

- It is observed that most of the literature uses neural networks, genetic algorithms, and deep learning in cryptography for generating secret keys and/or to perform encryption and decryption.
- Neural network is more secure and robust to generate and exchange the key [5, 7, 8, 10].
- The cryptosystem produced by machine learning approaches achieves significant security.
- Modern cryptographic techniques can be integrated with a neural network if a lossless neural network is achieved [9].
- The key length of 64-bit provides high security against brute force attacks [8].

4 Conclusion

In this paper, we gave an overview of several cryptography techniques that use machine learning. According to the findings of the review, the area of cryptography that falls under the purview of machine learning has a significant potential for expansion and improvement. The future research opportunities in machine learning approaches to cryptography include the following but are not limited to these are

- Designing of improved symmetric cryptosystem using machine learning algorithms.
- Apply machine learning techniques in cryptanalysis to identify encryption algorithms, keys, or modes of encryption algorithms from cypher texts.

References

[1] Rivest, R.L. (1993). Cryptography and machine learning. In: Imai, H., Rivest, R.L., Matsumoto, T. (eds) Advances in Cryptology — ASIACRYPT '91. ASIACRYPT 1991. Lecture Notes in Computer Science, vol 739. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-57332-1_36

[2] Alani, M. (2019). Applications of Machine Learning in Cryptography: A Survey. In Proceedings of the 3rd International Conference on Cryptography, Security and Privacy (ICCSP '19). Association for Computing Machinery, New York, NY, USA, pp. 23–27. https://doi.org/ 10.1145/3309074.3309092.

[3] Tom Mitchell (1997). Machine Learning. First ed., McGraw Hill. pp. 414.

[4] Ethem Alpaydin (2010). Introduction to Machine Learning. 2nd ed., The MIT Press Cambridge. Massachusetts, London, England. pp. 579.

[5] Mishra, S., & Bali, S. (2013). Public key cryptography using genetic algorithm. International J Recent Technol. Eng.(IJRTE), 2(2), 150-154.

[6] Kumar, Krishan and Sagar, Vikas. (2014). A Symmetric Key Cryptographic Algorithm Using Counter Propagation Network (CPN). 10.1145/2677855.2677906.

[7] Sahana,S,K., and Mahanti,P.K. (2015). An Analysis of Email Encryption using Neural Cryptography. Journal of Multidisciplinary Engineering Science and Technology, Volume-2(1), pp. 83-87, ISSN: 3159-0040, www.jmest.org/wp-content/uploads/JMESTN42350307.pdf.

[8] Atee, Hayfaa and Ahmad, Robiah & Noor, Norliza & Yasari, Abidulkarim. (2016). Machine learning-based key generating for cryptography. 11. 1829-1834. 10.3923/jeasci.2016.1829.1834.

[9] Sharma Kartik; Aggarwal Ashutosh; Singhania Tanay; Gupta Deepak; Khanna Ashish (2019). Hiding Data in Images Using Cryptography and Deep Neural Network. Journal of Artificial Intelligence and Systems, 1, 143–162. https://doi.org/10.33969/AIS.2019.11009.

[10] Naveena, V., and Satyanarayana, D.S. (2019). Symmetric Cryptography using Neural Networks. ” International Research Journal of Engineering and Technology (IRJET), 6(8), 1556-1558.

[11] Thoms GRW, Muresan R, Al-Dweik A (2019). Chaotic encryption algorithm with key-controlled neural networks for intelligent transportation systems. IEEE Access. 7:158697–158709. doi: 10.1109/ACCESS.2019.2950007.

[12] Wang, Peng, and Xiang Li. (2021). "TEDL: A Text Encryption Method Based on Deep Learning" Applied Sciences 11, no. 4: 1781. https://doi.org/10.3390/app11041781