American Journal of Electronics & Communication

Ankita Saha et. al., American Journal of Electronics & Communication, Vol. I (4), 22-26

# A Study of Machine Learning Techniques in Cryptography for Cybersecurity

Ankita Saha[1]
Chanda Pathak[1]
Sourav Saha[2]

[1]Department of Computer Science & Engineering, Institute of Engineering and Management, Kolkata
[2]Department of Computer Science & Engineering, Kalyani University
ankitasaha.ac@gmail.com
chandapathak34@gmail.com
souravsaha1977@gmail.com

**Abstract**

The importance of cybersecurity is on the rise as we have become more technologically dependent on the internet than ever before. Cybersecurity implies the process of protecting and recovering computer systems, networks, devices, and programs from any cyber attack. Cyber attacks are an increasingly sophisticated and evolving danger to our sensitive data, as attackers employ new methods to circumvent traditional security controls. Cryptanalysis is mainly used to crack cryptographic security systems and gain access to the contents of the encrypted messages, even if the key is unknown. It focuses on deciphering the encrypted data as it works with ciphertext, ciphers, and cryptosystems to understand how they work and find techniques for weakening them. For classical cryptanalysis, the recovery of ciphertext is difficult as the time complexity is exponential. The traditional cryptanalysis requires a significant amount of time, known plaintexts, and memory. Machine learning may reduce the computational complexity in cryptanalysis. Machine learning techniques have recently been applied in cryptanalysis, steganography, and other data-security-related applications. Deep learning is an advanced field of machine learning which mainly uses deep neural network architecture. Nowadays, deep learning techniques are usually explored extensively to solve many challenging problems of artificial intelligence. But not much work has been done on deep learning-based cryptanalysis. This paper attempts to summarize various machine learning based approaches for cryptanalysis along with discussions on the scope of application of deep learning techniques in cryptography.

*Key words:* Cryptanalysis, machine learning, deep learning, ciphertext, cryptography, cybersecurity.

## 1. Introduction

Cryptography is the process that involves encryption and decryption of text using various mechanisms or algorithms. A cryptographic algorithm is a mathematical function that can be used in the process of encryption and decryption. Classical cryptography is the process of transforming the plain text into the ciphertext so that the data can be transmitted through some communication channels. But these communication channels are mainly insecure mediums. A data string is being used as the key which helps in the transformation of the data from plain text to cipher text. It depends on the computational complexity of the factorization of the large numbers. As the classical cryptography is solely based on mathematics, so for the manual purpose it is easy to use. The plaintexts are being protected from the casual snooping using the classical cryptography. However, the Machine Learning techniques can also be applied in cryptography. Machine learning techniques have had a long list of applications in recent years. But the use of machine learning in information and network security is not new. Machine learning and cryptography have many things in common. The most apparent is the processing of large amounts of data and large search spaces. In its varying techniques, machine learning has been an interesting field of study with massive potential for application. In the past three decades, machine learning techniques, whether supervised or unsupervised, have been applied in cryptographic algorithms, cryptanalysis, steganography, among other data-security-related applications. This paper presents an updated survey of applications of machine learning techniques in cryptography and cryptanalysis.

## 2. Traditional cryptography and its limitations

Traditional cryptography focuses on the sender and receiver of a known message and it uses the same secret key. In this secret key cryptography method, the sender uses the secret key to encrypt the message, and the same secret key is being used by the receiver to decrypt the message at the receiver end. Here the sender and receiver encounter the problem of agreeing on the same secret key without anyone else intercepting it. As secret-key cryptography was having difficulty regarding secure key management, public-key cryptography was invented to solve the key management problem. In public-key cryptography, a pair of keys, called the public key and the private key is provided to both the sender and the receiver. Here each person's public key is known while the private key is kept secret from both the sender and receiver. Public-key cryptography can be used for authentication as well as for encryption. Rosen-Zvi et. al.[1] proposed that the common learning in a tree parity machine can also be used as a public-key cryptosystem. It also explains that the tree parity machine has significant potential in being used as a public-key cryptosystem. But the traditional cryptography has a disadvantage of being slower. There are popular secret-key encryption methods which are significantly faster than any currently available public-key encryption method. But traditional cryptography comes at a cost

which includes time as well as money. Addition of cryptographic techniques in the information processing leads to delay. The use of public key cryptography requires setting up and maintenance of public key infrastructure requiring the handsome financial budget. Machine learning can drastically reduce the number of computational complexity regarding cryptanalysis for block ciphers and also machine learning can produce very powerful distinguishers. Elliptic curve cryptography is an alternative technique to RSA which is a powerful cryptography approach. It generates security between key pairs for public key encryption by using the mathematics of elliptic curves. Elliptic curve cryptography has gradually been growing in popularity recently due to its smaller key size and ability to maintain security. This trend will probably continue as the demand on devices to remain secure increases due to the size of keys growing, drawing on scarce mobile resources . Hashing is a method of cryptography that converts any form of data into a unique string of text. Any piece of data can be hashed irrespective of its size or type. In traditional hashing, regardless of the data's size, type, or length, the hash that any data produces is always the same length. The average user encounters hashing daily in the context of passwords. Hashing is a mathematical operation that is easy to perform, but extremely difficult to reverse. The most widely used hashing functions are MD5, SHA1 and SHA-256.

## 3. Machine Learning in Cryptography

Machine learning has the capability to reduce the number of computational complexity in cryptanalysis. Machine learning is basically the field of study which gives the computers the capability to learn without being explicitly trained. The process of learning begins with observations or data such as, direct experience, examples or instructions, in order to look for patterns in the data and make better decisions in the future, based on the examples that we provide. The primary aim of machine learning is to allow the computers to learn automatically without human assistance or intervention and adjust actions according to that. The basic difference between machine learning and traditional programming is, the data input and the program logic has to be fed in and then run it on the machine to get the desired output. But in machine learning, the data input and the output has been fed in and it has to be run on the machine and the machine creates its own program logic which can be evaluated while testing. Feature extraction is an important part of machine learning. It is a process of reduction of the data by selecting and combining variables into features.

R.L Rivest[2] suggested, Machine learning and cryptanalysis share many similarities. In a typical cryptanalytic situation, the cryptanalyst tries to break the complete cryptosystem or at least some parts of it. So the cryptanalyst tries to find the secret key used by the users of the cryptosystem, where the general system is already known. The main aim of cryptanalysts was to exactly identify the decryption function which was being used. This problem was described as the problem of learning an unknown function on the basis of the input/output behavior and the prior knowledge about the class of possible functions.

M.M. Alani[3] proposed different types of attacks in cryptography as well as cryptanalysis, which also showed that machine learning and cryptography share many common properties. After examining the security attacks on machine learning techniques and machine-learning-based systems, we can say that the use of machine learning techniques can be done in cryptanalysis to extract decryption keys from ciphertext blocks and also to improve their efficiency in finding solutions in the search space.

Ramani Sagar [4] proposed an overview on how AI can be applied for encrypting data and undertaking cryptanalysis of such data and other data types in order to assess the cryptographic strength of an encryption algorithm. Jonathan Blackledge et. al.[5] proposed some recent advances in the field of Cryptography using Artificial Intelligence. It specifically considered the applications of Machine Learning and Evolutionary Computing to analyze and encrypt data. It considered the implementation of Evolutionary Computing and Artificial Neural Networks for generating unique and unclonable ciphers. Al-Shammari et. al.[6] proposed a classification technique which was based on machine learning, to classify encrypted traffic. It was done to assess the robustness of machine learning classification of encrypted traffic.

## 4. Deep Learning in Cryptography

Deep learning is a field that is based on learning and improving on its own by examining the computer algorithms while machine learning uses simpler concepts. Deep learning works with artificial neural networks which are designed to imitate how humans think and learn. Deep learning has aided language translation, image classification and also in speech recognition. Any pattern recognition problem can be solved without human intervention using deep learning. As deep learning is a subset of machine learning, it uses a hierarchical level of artificial neural networks to carry out the process of machine learning. The hierarchical function of deep learning helps the machines to process data with a nonlinear approach. If the size of the data increases then deep learning can keep on improving. Deep learning has provided great improvements on a number of difficult tasks. Deep learning is providing great improvements in different fields day by day. Here we have studied some aspects of deep learning in different fields. The importance of cyber security is on the rise as we have become more technologically dependent than ever before. Cyber security is the state or process of protecting and recovering computer systems, networks, devices, and programs from any types of cyber attack. Cyber attacks are an increasingly sophisticated and evolving danger to our sensitive data, as attackers employ new methods powered by social engineering and artificial intelligence to circumvent traditional security controls. This paper summarizes the research done in these areas and provides suggestions for future directions in research.

Jaewoo So[7] proposed a generic deep learning based cryptanalysis model that finds the key from known plaintext-ciphertext pairs and shows the feasibility of the deep learning based cryptanalysis by applying it to lightweight block ciphers. The deep neural networks can also be used to find the key from known plaintexts. A generic and automated cryptanalysis model based on deep learning was developed and it was used for checking the safety of the lightweight block ciphers, such as S-DES, Simon, and Speck. But the drawback of the proposed DL-based cryptanalysis is that the keyspace is restricted to the text-based key. Modern cryptographic functions are designed to be very random looking and to be very complex, and therefore, machine learning is quite difficult for finding the meaningful relationships between the inputs and the outputs when the keyspace is not restricted. If the keyspace is not limited, the deep learning based cryptanalysis failed to attack the block ciphers.

## Conclusion

In this paper, we have presented the survey of applications of machine learning as well as deep learning techniques in cryptography and cryptanalysis. Several studies have accounted that machine learning can significantly reduce the computational complexity regarding cryptanalysis. We have also conducted surveys on the uses of deep learning techniques on different types of cryptographic implementations. For producing more powerful cryptographic distinguisher, machine learning can be used and the resultant data and graphs from the surveys showed that the neural distinguisher achieved better accuracy and also the success rate is higher. There are lots of scopes of machine learning. It can be further used in the investment sectors, security purposes, self driving car etc. Nowadays Deep learning can also be used in these sections. These works opened ways for further research on deep learning techniques in order to get better results while challenging cryptographic implementations.

## References

*[1] M. Rosen-Zvi, E. Klein, I. Kanter, and W. Kinzel, "Mutual learning in a tree parity machine and its application to cryptography," Physical Review E, vol. 66, no. 6, p. 066135, 2002.*

*[2] "Cryptography and Machine Learning", Ronald L. Rivest, Laboratory for Computer Science Massachusetts Institute of Technology Cambridge, MA 02139.*

*[3] "Applications of Machine Learning in Cryptography: A Survey", MOHAMMED M. ALANI, Khawarizmi International College, United Arab Emirates.*

*[4] "Applications in Security and Evasions in Machine Learning: A Survey", Ramani Sagar, Rutvij Jhaveri and Carlos Borrego.*

*[5] "Applications of Artificial Intelligence to Cryptography", Jonathan Blackledge, Napo Mosola.*

*[6] R. Alshammari and A. N. Zincir-Heywood, "Machine learning based encrypted traffic classification: Identifying ssh and skype," in Computational Intelligence for Security and Defense Applications, 2009. CISDA 2009. IEEE Symposium on, pp. 1–8, IEEE, 2009.*

*[7] "Deep Learning-Based Cryptanalysis of Lightweight Block Ciphers", Jaewoo So Department of Electronic Engineering, Sogang University, Seoul 04107.*

*[8]*