International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/IJSRCSEIT

Machine Learning Models for Analyzing Different Cryptosystems' Security Levels

S. Sandhya[1] , S. Swetha[2]
[1](M.Tech Student) [2](Associate Professor)
Department of Computer Science and Engineering, Sree Rama Engineering College, Tirupati, India

Article Info

Publication Issue :
Volume 8, Issue 5
September-October-2022

Page Number : 136-143

Article History
Accepted: 15 Sep 2022
Published: 27 Sep 2022

ABSTRACT

Recent developments in multimedia technology have made the security of digital data a vital concern. To address the shortcomings of the current security mechanisms, researchers frequently concentrate their efforts on altering the existing protocols. However, during the past few decades, a number of suggested encryption algorithms have been shown to be unsafe, posing a major security risk to sensitive data. The optimal encryption technique must be used in order to protect against these attacks, but the type of data being secured will determine which algorithm is ideal in a particular circumstance. However, comparing several cryptosystems one at a time to find the best one might take a lot of processing time. We provide a support vector machine-based method for quickly and accurately choosing the appropriate encryption algorithms for photo encryption techniques (SVM). As part of this endeavour, we also generate a dataset using widely used security standards for encryption, including entropy, contrast, homogeneity, peak signal-to-noise ratio, mean square error, energy, and correlation. These factors are utilized as traits that were extracted from different cypher images. Dataset labels are divided into three categories based on their security level: strong, acceptable, and weak. Our recommended model's performance was examined for accuracy, and the findings demonstrate the effectiveness of our SVM-supported system.

Keywords : Support vector machine (SVM), security analysis, image encryption, cryptosystem.

I. INTRODUCTION

Due to the exponential increase in multimedia data transfers through unsecure networks, security has become a prominent study area (most notably the Internet). To shield data from prying eyes and unauthorized users, several academics have turned to creating new encryption techniques. When encrypting digital photos, dispersion and misunderstanding are two essential components (also known as scrambling). According to a hypothesis put forward by Claud Shannon, a cryptosystem containing confusion and diffusion techniques can be regarded as safe. On digital photos, the scrambling process can be applied directly to the pixels or to the rows and columns, whereas diffusion modifies the original pixel values. In other words, the replacement process replaces each distinct pixel value with the value of the S-unique box. The privacy of the data cannot be fully protected by transmission in an encrypted manner. Despite being encrypted for transmission, the information can still be viewed by unauthorized users due to the encryption algorithms lax security. The resilience of the picture is significantly impacted by the security level of the encryption technique used to encode it. The plain picture will be completely encrypted using a very powerful encryption technique, making it resistant to assaults on its availability, secrecy, and integrity. When selecting an encryption technique, temporal complexity is another important factor to take into account in addition to security. The type of application that has to be encrypted affects the choice of cryptosystem since various forms of data have different security requirements. We provide a security level detection method for picture encryption techniques by adding a support vector machine since the image encryption algorithm is crucial (SVM).

II. RELATED WORKS

Automated detection and classification of cryptographic algorithms in binary programs through machine learning by Diane Duros Hosfelt: Threats from the internet, particularly malicious software (i.e., malware) often employ cryptographic algorithms to hide their operations and even to seize control of a victim's machine (as in the instance of ransom ware). Malware and other threats proliferate too quickly for the time-consuming traditional methods of binary analysis to be effective. By automating identification and categorization of cryptographic methods, we can expedite programme analysis and more efficiently tackle malware. This thesis will give multiple approaches for automatically identifying and categorizing cryptographic algorithms in compiled binary code using machine learning. The findings in this work suggest that machine learning may be used to discover and identify cryptographic primitives in built code, while further research is required to thoroughly verify these approaches on real-world binary coders. The discovery and classification of cryptographic algorithms in small, single-purpose programmers is now accomplished using these techniques, and further research is being advocated in order to apply these methods to practical settings. This thesis explored the process of collecting characteristics and training machine learning models for the identification and classification of cryptographic methods in compiled code. Utilizing four distinct learning methods, three various model types were assessed on four different feature sets. Despite the fact that decision tree models were shown to perform best on these data, it is likely that an SVM with a linear kernel will generalize to real-world data more effectively. Cross-validation findings imply that algorithm classification and detection will be around 95% accurate, provided a reasonably small and homogenous.

Applications in Security and Evasions in Machine Learning: Machine learning (ML) has emerged in recent years as a crucial component to produce security and privacy in a variety of applications. Serious problems including real-time attack detection, data leaking vulnerability assessments, and many more are addressed with ML. In a variety of areas, including real-time decision-making, huge data processing, shortened learning cycle times, cost-efficiency, and error-free processing, ML comprehensively meets the demanding needs of the present security and privacy situation. As a result, in this work, we examine the cutting-edge techniques that make better use of machine learning (ML) to address current security-related real-world needs. We look at several security applications where ML models are crucial and contrast the accuracy outcomes using many available metrics. An outline for an interdisciplinary study field is provided by the analysis of ML algorithms in security applications. Attackers can circumvent the ML models by engaging in adversarial assaults, even with the deployment of modern, sophisticated technology and techniques. As a result, it becomes necessary to evaluate the ML models' susceptibility to adversarial assaults at the time of creation. To further support this notion, we also examine the many adversarial attacks that may be made against ML models. We have modelled the threat model and protection tactics against adversarial attack techniques to provide accurate representation of security features. Additionally, we addressed the model point at which potential assaults may be carried out and illustrated the adversarial attacks depending on the attackers' knowledge of the model. Finally, we also research various adversarial attack characteristics.

Machine learning approaches to IoT security: Attacks against IoT applications are continuously increasing as a result of the IoT applications' constant development and innovation. Our objective in writing this systematic literature review (SLR) report is to give academics a resource on current research trends in IoT security. The primary focus of our SLR article was a set of six research questions on machine learning and IoT security. A few important research trends that will guide future study in this area were discovered by this thorough literature review of the most current papers in IoT security. It's crucial to create models that can incorporate cutting-edge methods and technologies from big data and machine learning given the increase in large-scale IoT threats.

In order to identify the best algorithms and models to detect IoT threats in real or almost real-time, accuracy and efficiency are crucial quality aspects. This study looks into current research directions for machine learning applications in IoT security. We extracted the most pertinent and academic material published in the previous two years in order to evaluate recent research and anticipated developments in IoT security (2019 and 2020). The desire to combine three popular study areas—IoT, machine learning, and information security—drives our goals. Six research topics are extracted from the integration of those three areas and are provided in Section 3.1. The thorough but organized approach to the literature evaluation described in this work demonstrates the particular standards utilized to weed out studies that didn't advance the research objective. It assisted us in finding more recent and specialised research on IoT and machine learning approaches that are suggested to guard against widespread assaults on IoT devices.

While numerous survey studies were conducted in the field of IoT security research, they were either not systematic or did not concentrate on methods based on machine learning and deep learning to identify large-scale threats. The specifics of the available IoT security literature studies are presented in Section 5. The primary goal of this research project was to protect IoT devices from distributed and widespread assaults like botnets and distributed denial of service (DDoS). Machine learning and deep learning algorithms show promise in detecting zero-day attacks when compared to conventional intrusion detection methods like firewalls, antivirus software, etc. There has been a lot of study done in this area; However, it can be difficult for intrusion detection systems to identify zero-day attacks due to the rapidly growing nature of IoT device types, network traffic patterns, and cyber-attacks. To give the reader the necessary background information to properly traverse the rest of the article, we included a thorough background section. Researchers' recommended approaches and their associated performance findings on various benchmark datasets are extracted from carefully chosen articles in section 6 through rigorous examination. Each of the six research topics is fully addressed by the review findings reported in section 7. Our objective is to give scholars a more specialised but thorough understanding of current approaches utilized recently to uncover recent trends, limits, and difficulties in order to contribute to the development of future successful intrusion detection systems.

Secure, privacy-preserving and federated machine learning in medical imaging: Due to the lack of standardised electronic medical records, the limited dataset availability for algorithm training and validation, as well as the stringent legal and ethical requirements to protect patient privacy, are currently impeding the widespread application of artificial intelligence techniques in medicine. While electronic data storage and standardised, harmonised data sharing formats, such as Digital Imaging and Communication in Medicine, partially alleviate the first issue, the criteria for privacy protection are as stringent. The use of technology solutions that concurrently satisfy the needs for data security and usage is required to prevent patient privacy breach while supporting scientific research on massive datasets that aims to enhance patient care. With an emphasis on medical imaging applications, we offer a review of existing and emerging techniques for federated, secure, and privacy-preserving artificial intelligence, along with potential attack vectors and future prospects in medical imaging and beyond.

As shown, for instance, in medical imaging, where the use of computer vision techniques, conventional machine learning, and—more recently—deep neural networks have produced notable breakthroughs, artificial intelligence (AI) technologies have the potential to transform the field of medicine. This development can be attributed to the dissemination of massive, curated image corpora, with Image Net possibly being the best-known example. This development led to efficient pre-trained algorithms that facilitate transfer learning and increased publications in both oncology—with applications in tumour detection, genomic characterization, tumour subtyping, grading prediction, outcome risk assessment, or risk of relapse quantification—and non-oncologic applications, like chest X-ray analgesia.

Machine Learning and Cryptographic Algorithms: The use of AI, deep learning, and machine learning algorithms is expanding across all information technology application areas, including information security. The classic password management systems, autoprovisioning systems, and user information management systems are well known in the information security area. With ransomware, there is also growing worry about application and system level security. Cyber-attacks using ransom ware are becoming more common on systems currently in use. Malware classified as "ransom ware" aims to obtain data using an encryption method and then return it in exchange for payment. The majority of ransom ware attacks target weakly secured systems that are exposed to the network. Machine learning techniques may be used to analyze the pattern of assaults. Create or discuss a workaround that combines a machine learning model with a cryptographic technique to improve the system's ability to respond to potential threats. With the aid of intelligent system password management and intelligent account provisioning, the second difficult aspect of the problem is to develop intelligence for enterprises to stop ransom ware assaults. I elaborate on the examination of machine learning methods for the issue of intelligent ransom ware detection in this research; the subsequent section will deal with algorithm creation.

III. METHODS AND MATERIAL

Proposed system:

In recent years, a plethora of encryption algorithms, including chaotic and transformation-based methods, have been introduced. Some of the currently used encryption methods have been shown to be insecure and to offer insufficient security based on statistical analysis of their results. One method to assess the level of security of an encryption algorithm is to analyze the statistics of its security parameters. Traditional approaches often involve making these comparisons one at a time, which takes a lot of time.

We developed a machine learning model that combines SVM to assist us in more quickly selecting an appropriate encryption method.

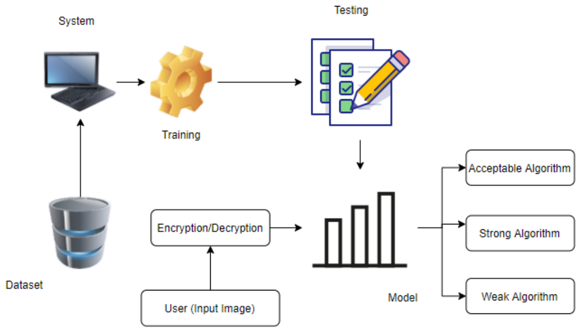


Figure 1: Block diagram

IV. IMPLEMENTATION

The project has implemented by using below listed algorithm.

1. Support-vector machine

Support-vector machines are supervised learning models using learning algorithms that evaluate data for classification and regression analysis in machine learning. SVMs, which are based on statistical learning frameworks, are one of the most reliable prediction approaches. An SVM training algorithm creates a model that assigns new examples to one of two categories, making it a non-probabilistic binary linear classifier, given a series of training examples, each marked as belonging to one of two categories. SVM maps training examples to points in space in order to widen the distance between the two categories as much as possible. Then, based on which side of the gap they fall, new samples are projected into that same area and predicted to belong to a category. A support-vector machine, in more technical terms, creates a hyper plane or set of hyper planes in a high- or infinite-dimensional space that can be used for classification, regression, or other tasks such as outlier detection. Intuitively, the hyper plane with the greatest distance to the nearest training-data point of any class (so-called functional margin) achieves a decent separation, because the larger the margin, the lower the classifier's generalisation error. The sets to discriminate are typically not linearly separable in that space, even though the beginning problem is described in a finite-dimensional space. In order to facilitate separation, it was proposed to move the original finite-dimensional space into a considerably higher-dimensional area. By specifying the mappings employed by SVM methods in terms of a kernel function, it is ensured that the dot products of pairs of input data vectors may be computed easily in terms of the variables in the original space.

2. DNA encoding

In place of conventional silicon-based computer technology, DNA computing makes use of molecular biology, biochemistry, and other biological processes. Bimolecular computing, often known as DNA computing, is a rapidly growing multidisciplinary field. Researchers have developed several biological operations and algebraic operations based on DNA sequence due to the quick growth of DNA computing [13]. Four nucleotides make up the single-strand DNA sequence: A, C, G, and T. A and T and C and G complement one another. The binary system is used to express all information in the contemporary theory of the electronic computer. However, DNA sequences are thought to represent information in the DNA coding hypothesis. Therefore, the four bases in the DNA sequence are expressed using binary numbers, with each base being represented by a two-bit binary integer. The binary system's theory states that 0 and 1 are complimentary, therefore we can determine that 01 and 10 are likewise complimentary, along with 00 and 11. Four bases may be expressed using the digits 00, 01, 10 and 11, and there are 4! = 24 different coding combination types. There are 24 different types of coding combinations, but only eight of them meet the complementary base pairing concept since DNA bases complement one another.

3. Logistic Map

The logistic map, a degree 2-polynomial mapping (also known as a recurrence relation), is commonly touted as a paradigmatic example of how complex, chaotic behaviour may emerge from remarkably simple non-linear dynamical equations. A discrete-time demographic model resembling Pierre François Verhulst's logistic equation, the map gained notoriety in a 1976 article by the biologist Robert May [1]. This nonlinear difference equation aims to represent two effects: reproduction, where the population will grow while the population is small at a rate proportionate to the existing population. Starvation (density-dependent mortality), in which the growth rate will decline at a rate inversely correlated to the value determined by subtracting the present population from the environment's theoretical "carrying capacity". However, the logistic map's fundamental flaw as a demographic model is that certain beginning conditions and parameter values (for instance, if r > 4) result in negative population numbers. The earlier Ricker model, which likewise displays chaotic dynamics, does not have this issue.

4. Rubik's Cube Image Encryption

To permute picture pixels, this approach is based on the Rubik's cube theory. Using a key, the XOR operator is used to odd rows and columns of a picture to muddle the link between the original and encrypted images. Even rows and columns of the picture are treated with the same key that has been flipped. The suggested algorithm's resistance to many forms of assaults, including statistical and differential attacks, has been tested experimentally with comprehensive numerical analysis (visual testing). Additionally, performance evaluation experiments show how highly secure the proposed picture encryption technique is. Additionally, it has quick encryption and decryption capabilities, making it appropriate for real-time Internet encryption and transmission applications.

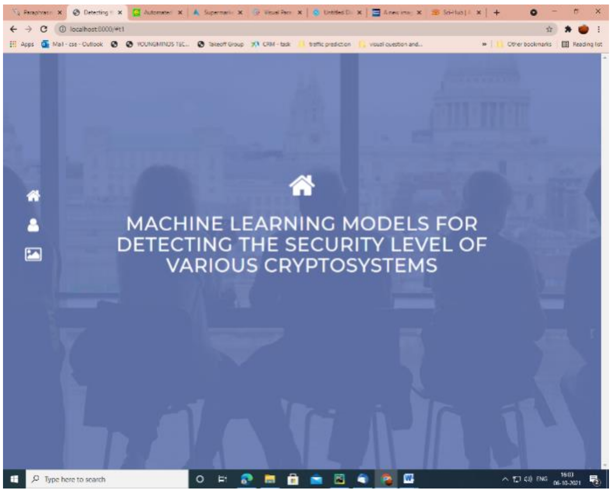
5. Lorenz Image Encryption

A model of thermally induced fluid convection in the atmosphere is all that the Lorenz equation is. E.N. Lorenz documented and published the model for the first time in 1963 [11–13]. Due to the attractor having two wings like butterflies, it is considered to be one of the classical chaotic systems and is implicated in scientific investigations as the origin of the "butterfly effect" [12, 14]. As a result, it has received much research in the fields of chaotic control, chaotic theory, and synchronization phenomena. A 3D dynamical system called the Lorenz chaotic equation is defined by x, y, and z. In relation to the original system parameters, the equation system exhibits chaotic behaviour. The Lorenz system exhibits a chaotic behaviour that is far more complex than any 1D or 2D chaotic systems. We encrypt the photos using the Lorenz equation.

V. RESULTS AND DISCUSSION

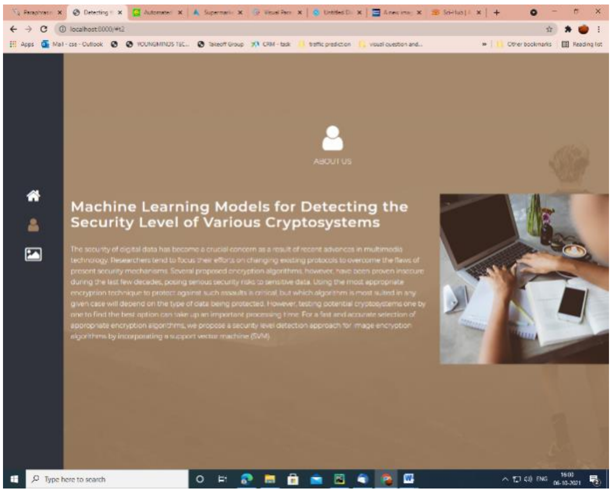
The following screenshots are depicted the flow and working process of project.

Machine Learning Models for detecting the security level of various cryptosystems:



MACHINE LEARNING MODELS FOR DETECTING THE SECURITY LEVEL OF VARIOUS CRYPTOSYSTEMS

This page describes the concept behind this project.

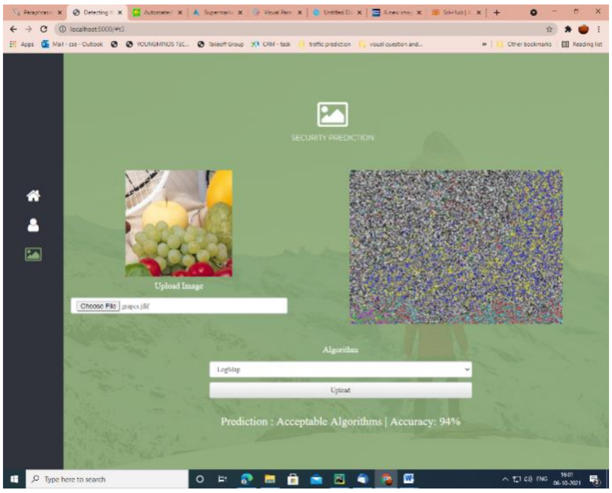


ABOUT US

Machine Learning Models for Detecting the Security Level of Various Cryptosystems

The security of digital data has become a crucial concern as a result of recent advances in multimedia technology. Researchers tend to focus their efforts on changing existing protocols to overcome the flaws of present security mechanisms. Several proposed encryption algorithms, however, have been proven insecure during the last few decades, posing serious security risks to sensitive data. Using the most appropriate encryption technique to protect against such assaults is critical, but which algorithm is most suited in any given case will depend on the type of data being protected. However, testing potential cryptosystems one by one to find the best option can take up an important processing time. For a fast and accurate selection of appropriate encryption algorithms, we propose a security level detection approach for image encryption algorithms by incorporating a support vector machine (SVM).

This Page describes the accuracy of models.



localhost:8000/#t3

SECURITY PREDICTION

Upload Image

Choose File grapes.jfif

Algorithm

LogMap

Upload

Prediction : Acceptable Algorithms | Accuracy: 94%

Accuracy with svm: 94%

VI.CONCLUSION

In this post, we built and suggested a model that can rapidly and accurately determine the security level of various encryption systems. We began by creating a dataset and adding traits that indicated the security criteria that were shared by various encryption methods. To create a dataset, we divided the values of all attributes into three ranges, strong, acceptable, and weak, which stand for the resultant security levels. The level of security that each of the different encryption algorithms provides is then assessed using our recommended methodology. The level of security of various encryption algorithms may be manually determined by computing the statistical statistics of each one. Standard testing methods take a lengthy time to finish this process; however, with our recommended method, testing might be finished in a couple of seconds. Finally, we assessed the performance of our proposed model using a variety of tests and found that it produces predictions 94% accurately and much more quickly than other models already in use.

VII. REFERENCES

[1]. I. Hussain, A. Anees, A. H. Alkhaldi, M. Aslam, N. Siddiqui, and R. Ahmed, "Image encryption based on Chebyshev chaotic map and S8 S-boxes.

[2]. A. Anees, I. Hussain, A. Algarni, and M. Aslam, "A robust watermarking scheme for online multimedia copyright protection using new chaotic map,".

[3]. A. Shafique and J. Ahmed, "Dynamic substitution based encryption algorithm for highly correlated data, ".

[4]. F. Ahmed, A. Anees, V. U. Abbas, and M. Y. Siyal, "A noisy channel tolerant image encryption scheme,".

[5]. M. A. B. Farah, R. Guesmi, A. Kachouri, and M. Samet, "A novel chaos based optical image encryption using fractional Fourier transform and DNA sequence operation,".

[6]. C. E. Shannon, "Communication in the presence of noise,".

[7]. S. Heron, "Advanced encryption standard (AES),".

[8]. H. Liu, A. Kadir, and X. Sun, "Chaos-based fast colour image encryption scheme with true random number keys from environmental noise,".

[9]. Y.-L. Lee and W.-H. Tsai, "A new secure image transmission technique via secret-fragment-visible mosaic images by nearly reversible color transformations".

[10]. A. Anees, A. M. Siddiqui, and F. Ahmed, "Chaotic substitution for highly auto correlated data in encryption algorithm,".

[11]. L. Liu, Y. Lei, and D. Wang, "A fast chaotic image encryption scheme with simultaneous permutation-diffusion operation".

[12]. M. Khalili and D. Asatryan, "Colour spaces effects on improved discrete wavelet transform-based digital image watermarking using Arnold transform map,".

[13]. L. Zhang, J. Wu, and N. Zhou, "Image encryption with discrete fractional cosine transform and chaos.

Cite this article as :

S. Sandhya, S. Swetha, "Machine Learning Models for Analyzing Different Cryptosystems' Security Levels", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 8 Issue 5, pp. 136-143, September-October 2022.
Journal URL : https://ijsrcseit.com/CSEIT228521