AI AND CRYPTOGRAPHY:

NAVIGATING THE RISKS

Team Members

Shivam Bhosale (SUID - **268919718**) Contributions: Designed the algorithm and documentation.

Divya Kapil (SUID - **611012880**) Contribution: Data creation, cleaning and processing along with the documentation.

Prerna Shinde (SUID - **768132215**) Contribution: Model Development and documentation.

AI AND CRYPTOGRAPHY: NAVIGATING THE RISKS

Shivam Bhosale
Computer of Engineering and Computer Science
Syracuse University
Syracuse, United States
shbhosal@syr.edu

Prerna Shinde
Computer of Engineering and Computer Science
Syracuse University
Syracuse, United States
pgshinde@syr.edu

Divya Kapil
Computer of Engineering and Computer Science
Syracuse University
Syracuse, United States
dikapil@syr.edu

***Abstract*— Continuous development in artificial intelligence (AI) have posed challenges to the discipline of cryptography, which has historically served as the foundation for secure communication and data protection. This study looks into the negative effects of AI on cryptography systems, including its role in improving cryptanalysis and enabling advanced assaults such side-channel exploits. This study advances our understanding of the complex interaction between AI and cryptography by studying real-world use cases and providing novel solutions, as well as contributing to the creation of resilient security frameworks.**

I. INTRODUCTION

As AI technologies evolve, cryptography which once represented secure communication and data protection is facing new challenges. AI has the potential to enhance cryptanalysis, break encryption schemes which poses a significant threat to modern cryptography. This project focuses on analysing the negative impact of AI on cryptography in detail and proposes practical solutions to tackle these issues. To analyse the threats posed by AI, we would be conducting research on the current usage of AI in the cryptography industry, including case studies of AI based cryptanalysis and side channel attacks. The research will help us to identify areas where AI poses a significant threat to the encrypted information. We would also be evaluating a machine learning model with the ability to recognize cryptographic patterns and break them. However, we would also be looking at how we can use AI to boost the security of the cryptographic systems. Overall, we would be developing a detailed report that not only highlights the negative impact of AI but also provides solutions to safeguard the cryptographic systems which would help in securing data.

II. LITERATURE REVIEW

A. Title: How Artificial Intelligence becomes a threat to cryptography-A systematic literature review

The paper [1] "How Artificial Intelligence becomes a threat to cryptography-A systematic literature review" provides a detailed explanation of the evolving relationship between AI and cryptography. This paper mainly focuses on how AI, Maintaining the Integrity of the Specifications particularly machine learning (ML) and deep learning (DL), can prove to be a threat to the traditional cryptographic systems. The authors emphasize that while AI has the potential to automate and optimize cryptanalysis by analysing large datasets and learning from patterns, the technology is still in the early development stages. We still have a lot to learn about AI before we can effectively deploy it to compromise cryptographic systems. One of the main limitations currently in the cryptography industry is the lack of robust AI-resistant mechanisms. By reviewing this paper which highlights the negative impacts of AI, we aim to provide a balanced view of how AI, despite the threats, can offer new methods to improve cryptography. In our project, we would further explore how AI can contribute in securing the current cryptographic systems.

B. Title: Cryptography: Against AI and Odds

This paper [2] "Cryptography: Against AI and QAI Odds" provides a comprehensive analysis of the risks posed by AI and quantum AI(QAI) to traditional cryptographic systems. Currently, cryptography does not take in account the ability of the AI to break encryptions. Most systems, including RSA, AES and Diffie-Hellman, depend on the hardness of mathematical problems. These systems are effective against traditional brute-force attacks. However, AI with its advanced pattern recognition abilities and high computational power, can exploit the weakness in the encryption algorithms making it easier to gain unauthorized access. Another limitation of the current practice is reliance on black-box models for cryptographic purposes. These models, while powerful, are not transparent, making it difficult to identify how truly secure these systems are. Considering the limitations of the current cryptographic systems, we plan to incorporate AI resistant techniques like introducing randomness and pattern devoid encryption strategies that will create highly unpredictable ciphertexts, in our project.

C. Title: Privacy and Security Concerns in Generative AI: A comprehensive survey

This paper [3] provides the detail regarding the analysis made regarding current use of generative AI technologies. Nowadays, generative AI is using its advanced skills and dominating in various applications like natural language processing, image processing, text processing, image recognition etc. Generative AI can be used to solve complex problems and algorithms such as VAEs and GANs that are trained with big datasets to create novel content. But this advancement is also not good as it leads to security threats.

This paper highlights the issues caused by the AI models and focuses on the risks that can be caused like data leaking. This is because the data that is fed into the models are remembered which can lead to privacy issues and the data gets saved throughout which can be stolen. Furthermore, the approaches used currently do lack the robust mechanisms which can secure the data leading to challenges of tracking the usage of data and if it ensures privacy regulations. Additionally, researchers also found that many generative AI use black box, which increases difficulties for developers and individuals to understand as to how decisions were made. Because of this it becomes more complicated as it is no longer transparent and outputs are unfair and biased.

D. Title A Survey of Privacy Risks and Mitigation Strategies in Artificial Intelligence Life Cycle

This research [4] analyzes the methods that are used to handle the privacy risks that are present during the development of AI models. Organizations generally employ methods such as encryption, access controls, and anonymization to protect the personal information. Anonymization is used for removing information that can be identified from the datasets but has limitations and it can be misused. In encryption, we cannot address the risk that is associated within AI models even if it's a way of securing the data. There are different privacy regulations like GDPR and CCPA that are used more in organizations but the current practices usually have placed their focus towards adhering to laws. However, current procedures frequently limit the use of privacy impact assessments (PIAs) to the very beginning of development, which may provide room for changes in data usage over time. Organizations have exposed themselves due to use of AI techniques that develop and use fresh data updates because of reactive strategy. Many privacy measures don't even consider the implications of AI on social equity. Overall, even if the privacy concerns in AI have been solved to a good extent still the current methods are insufficient and expect more proactive approaches that should be implemented in privacy management.

E. Title: Advances and challenges in cryptography using Artificial intelligence

The paper [5] "Advances and challenges in cryptography using Artificial intelligence." provides an overview of how AI is transforming the field of cryptography. AI offers new methods for improving encryption techniques and securing communication channels as it continues to advance. However, it also introduces challenges, as AI can be used by attackers to break cryptographic algorithms more efficiently. This paper discusses positive and negative effects of AI on cryptography, explaining how AI is applied to strengthen cryptographic systems and conversely how it can be exploited to undermine them. It highlights current advances in AI driven cryptography, like enhancing key generation, encryption algorithms, and real-time cryptanalysis. This paper even addresses the challenges like cryptographic security, including the potential for AI to accelerate attacks on current encryption methods posed on by AI. The paper also explores future research directions and suggests ways to mitigate the risks AI introduces to cryptographic systems while maximizing its benefits.

F. Title: AI resistant (AIR) Cryptography

This paper [6] "AI resistance (AIR) Cryptography" provides an overview of threats AI and quantum AI pose to the traditional cryptography systems. The authors of this paper propose solutions on how AI as it becomes more advanced, can speed up the process of breaking codes by predicting likely messages. The aim of these solutions is to make it difficult for AI to derive meaningful insights from ciphertexts by introducing unpredictability into the encryption process. The idea is to make solutions so random and complex that even advanced AI will not be able to find patterns or make useful guesses. The paper argues that with AI becoming powerful, cryptographers need to rethink how to design encryption systems to stay ahead of attacks. The solution proposed by the authors aims to make AI-driven attacks less effective by making encrypted data hard to guess no matter how much AI advances.

III. DATASET

We have developed a dataset that performs similar to real-world scenarios, ensuring both normal and attacked systems are categorized. The dataset has total of 1,000 samples, each classified into three attributes: *load, rate, and integrity_flag*. The *load* column represents the system's workload, while the *rate* shows its activity level. A server is considered to be operate under normal conditions if the *load* value is below 160 and the *rate* is below 15. But, if *load* exceeds the limit of 160 or the *rate* goes beyond 15, the server is will classify as being under attack. The *integrity_flag* indicates the server's status, where 1 signifies an attack and 0 represents normal operation.

IV. EXECUTION

The first step we executed was to develop a machine learning model capable of determining whether the network is operating under normal conditions or is under attack. For this purpose, we employed a Random Forest Classifier, trained on a dataset comprising network parameters such as *load*, *rate*, and *integrity_flag*. Once the prediction model is trained and ready with the accuracy of 93.5%, the system accepts text input from the user that they wish to encrypt. At the same time, the system collects current parameters, such as *load* and *rate*, to assess the state of the network using the trained model. Based on these network parameters, the machine learning model predicts whether the network is operating normally or is under attack. If the *load* exceeds 160 or the *rate* exceeds 15, the model flags the network as under attack; otherwise, it classifies the network as functioning normally. This prediction guides the selection of the subsequent encryption approach. When the network is predicted to be operating under normal conditions, the system employs AES-CBC (Cipher Block Chaining) mode for encryption. To begin, a random 256-bit key is generated, along with a random 16-byte Initialization Vector (IV), ensuring that the encryption process is both secure and non-repetitive. The plaintext message provided by the user is then padded using the PKCS7 scheme to ensure its length aligns with the block size required by AES. Once padded, the plaintext is encrypted using the AES-CBC mode, producing

a secure ciphertext. This method offers an optimal balance of security and efficiency, making it suitable for networks functioning without any detected anomalies. If the network is predicted under attack, the system switches to a more secure dynamic encryption algorithm which we implemented This mode incorporates several additional security measures to ensure robust encryption. A random 256-bit key and a 12-byte Initialization Vector (IV) are generated as the foundation of the encryption process. To further strengthen security, SALT is introduced, adding random bytes to the encryption key. These bytes are derived using the secure Key Derivation Function, PBKDF2HMAC (Password-Based Key Derivation Function [2] PBKDF2HMAC enhances security by generating strong, unique encryption keys that are highly resistant to cryptographic attacks, such as brute-force attacks, ensuring that even if the base key is known, the derived key remains unique. To increase randomness and reduce predictability, random padding is applied to the plaintext. The padding length is randomly selected between 1 and 16 bytes, making the encrypted output less susceptible to analysis. Additionally, an authentication tag is generated during encryption to ensure both data integrity and authenticity. This tag is verified during decryption to confirm that the ciphertext has not been altered and that the correct key and IV are in use. Together, these measures provide a highly secure encryption process tailored for networks under attack. During decryption, the system applies the same encryption mode that was used during the initial encryption process. For the dynamic encryption mode, the SALT generated during encryption is included to derive the same encryption key. The authentication tag is used to confirm the data integrity and prevent unauthorized access. The ciphertext is later decrypted, and the padding is removed to recover the original plaintext message. The system outputs the predicted network state (Normal or Under Attack) along with the encrypted ciphertext. In the case of dynamic encryption mode, the SALT and authentication tag are also provided to ensure that only authorized users can decrypt the ciphertext. The decrypted plaintext is displayed to verify the accuracy of the encryption and decryption processes.

V. RESULTS AND EVALUATION

As mentioned in sections III and IV, we have described the dataset and method which we used to test our application, Initially the dataset which was generated had too much noise and unnecessary contents which got emulated in a csv file covering both the scenarios i.e. normal and attack scenarios

Classification Report				
Class	Precision	Recall	F1-Score	Support
0 (Normal Network)	0.95	0.98	0.96	179
1 (Under Attack)	0.75	0.57	0.65	21
Overall Accuracy	93.5%			

Figure 1: Classification Report of the data in two different scenarios

The data was stored in a csv file. The Random Forest classifier was trained on this dataset and it produced excellent results. We have evaluated the results of the model and created following classification report:

Figure 1 displays the classification report of both the classes, 0 for normal scenario and 1 for under attack scenario. When the model predicted “Normal Attack”, result was classified correctly 95% of time, actual model instance was also identified with 98% value. For “Under Attack”, when the model predicted this scenario it had precision of 75%, f1 score of 65% and the overall accuracy turned out to be balanced with both scenarios as 93.5%. The encryption that was applied has been tested with numerous conditions which was predicted by Random Forest model. The application automatically switched between the two modes AES CBC and Dynamic Encryption. In Normal Network mode, encryption scheme used is AES CBC with 128bit random initialization vector (IV). For example, if we send a message for encryption, the model first returns the status of the network, and then it encrypts message of the 32-bit length and then the decoder decodes the message with its own key and gets the same message back. Hence, it can be concluded that the message was successfully encrypted and decrypted. In attack scenario, we use Dynamic Encryption scheme which has advanced security features and operates in GCM mode, firstly we perform key derivation using [2] PBKDF2HMAC using unique 16-bit salt that enhances the security in the network. In the later step, padding is added of random keywords of length between 1 and 16 bytes that is added to plain text so that the ciphertext doesn’t get predicted easily. These features add an additional security and makes the system more efficient.

VI. TECHNICAL CHALLENGES

As we navigate the integration of AI and cryptography, several new and complex challenges await us. One of the major challenges is to manage the computational needs of an AI driven system. Thus, it becomes important to manage our resources so that we can run these systems without any problems. AI can be unpredictable which can lead to incorrect decisions being taken which can be quite harmful in a field like cryptography where everything from the first step to the last step needs to be on point. Therefore, we need to ensure that AI systems behave predictably under various conditions. New threats emerge every day in the field of cryptography, so we would need to make sure that our algorithm stays up to date with all the threats and is resistant to them. One major disadvantage of AI is the ethical and privacy issues that come with it. We can not expose too much confidential data to our model; however, we would also need to expose enough data to help the model to predict the attack. Thus, finding the right balance between what to expose and what not to is the key to success. In conclusion, addressing these challenges is important to ensure the integration of AI enhances cryptographic security without compromising user integrity and security.

VII. CONCLUSION

Evolution of artificial intelligence has bought challenges in the field of cryptography. This research paper focuses on the negative impact of AI in cryptographic system. Our model analyzes various types of attacks and provides additional security to the messages sent over the network by improving the cryptanalysis and side channel exploits. The research

focuses on developing a machine learning model which is trained to apply dynamic encryption schemes for additional security of messages. Furthermore, this study provides us the understanding of relationship between the ai and cryptography.

REFERENCES

[1] Chethiya, P., 2023. “How Artificial Intelligence becomes a threat to cryptography– A systematic literature review”.

[2] Harris, Sheetal, Hassan Jalil Hadi, and Umer Zukaib. "Cryptography: Against AI and QAI Odds." arXiv preprint arXiv:2309.07022 (2023).

[3] Abenzer, G., Mekonen, K., Pandey, A., Singh, A., Hassija, V., Chamola, V., & Sikdar, B. (n.d.). Privacy and security concerns in generative AI: A comprehensive survey. IEEE

[4] [4] Shahriar, S., Allana, S., Hazratifard, S. M., & Dara, R. (n.d.). A survey of privacy risks and mitigation strategies in the artificial intelligence life cycle. IEEE

[5] B. Sharma, P. Goel and J. K. Grewal, "Advances and Challenges in Cryptography using Artificial Intelligence," 2023 IEEE

[6] G. Samid, "AI Resistant (AIR) Cryptography," 2023 Congress in Computer Science, Computer Engineering, & Applied Computing (CSCE), Las Vegas, NV, USA, 2023. IEER.