P-ATI Encryption Model Based on E-commerce Platform Data: A Explorative Study

Yutao An

School of Artificial Intelligence, Hebei University of Technology, Tianjin 300401, P.R.China

ABSTRACT

E-commerce platform plays an extremely important role in people's life today, but the Data Leak it produces will also bring great losses to users. In order to protect the vital interests of users, many encryption algorithms have been continuously proposed. This paper innovatively proposes a P-ATI encryption model based on existing algorithms. It combines four encryption algorithms, namely, Vigenère cipher, AES, DES, and RSA, and generates an encryption model that can apply different algorithms for different data according to their types and degrees of protection. The experimental results show that the algorithm achieves an improvement in encryption efficiency while ensuring the encryption effect.

KEYWORDS

Data analysis; Layered encryption model; E-commerce; Privacy protection; Information security

DOI: 10.47297/taposatWSP2633-456916.20230402

1 Introduction

In the current era of the Internet, people's reliance on e-commerce platforms is becoming increasingly heavy, and the frequency of use is also increasing. A large amount of data is stored in the cloud, therefore once this data is attacked and leaked, it will inevitably cause significant losses. Prior to this article, many experts and scholars had conducted influential research on the topic of data security and proposed many effective methods, such as New Lightweight Cryptographic Algorithm (NLCA) [1], Content based Double Encryption Algorithm [2], Two phase Hybrid Algorithm (THCA) [3]. However, these studies mostly focus on data security in cloud computing, wireless devices, or only use one or two algorithms, which may not have a good effect on security protection for data with diverse types and varying degrees of protection on e-commerce platforms. To address this issue, this paper analyzes user data based on e-commerce platforms, and designs a 4-layer layered encryption model P-ATI, taking into account the importance level of individual privacy and different functional divisions.

The top layer is defined as the User Privacy Layer, which includes bank card information and payment password bound to the platform. This type of information is closely related to the user's personal interests, and leakage can easily cause property damage to users. Therefore, the RSA algorithm of high protection level is used. The second and third layers are defined as the User Application Layer and User Trace Layer, both of which contain data generated by users when browsing and purchasing products, indicating their purchasing behavior and preferences, as well as their shipping information. They can be shared as public information under certain conditions, but those parts of information should be considered as personal privacy, so AES and DES algorithms of a relatively high level of protection are applied. The bottom layer is defined as the User Information Layer, which describes the user's name, gender, age and other basic information. Data indicates that

the basic information of this layer is public in most e-commerce platforms (Taobao, JD, Amazon), so the degree of privacy is low. Therefore, this paper uses Vigenère cipher of a relatively low protection level to encrypt, while ensuring the user's use efficiency.

The hierarchical encryption method this paper proposes enables the model to take different processing approaches according to different characteristics of the data, realizing dynamic equilibrium between encryption costs and data security.

Table1. Information leakage incidents on different online platforms from 2019 to 2022

Time	Incident
2019	Uniqlo was attacked by hackers. and more than 460000 customers' information was leaked. Personal information such as name, address, telephone number and e-mail address was disclosed.
2020	Estee Lauder's official server was attacked by hackers, leading to the data breach of its unencrypted cloud database, which contains sensitive data of more than 400 million customers, including their records, internal logs, e-mail addresses and other important private information.
2021	The online store of Delta Airlines was attacked by hackers, and the customer's name, address, e-mail, credit card information and test results of COVID-19 were leaked.
2022	Samsung Electronics' US system has been hacked, and some data in the system, such as name, contact information, demographic information, birth date, and product registration details, may be leaked.

2 Related Work

Fursan Thabit et al. [1], have proposed a New Lightweight Cryptographic Algorithm (NLCA) for improving data security in cloud computing which encrypts data based on symmetric cryptography. They concluded that the new algorithm has a high security level with low computation costs, compared to the frequently-used algorithm such as DES, AES.

Sourabh Chandra et al. [2], have proposed a Content-based Double Encryption Algorithm which follows the symmetric key cryptography method. The algorithm applies a binary addition operation on the content of the plain text twice and gives special attention to the security of the secret keys.

Sherief H. Murad et al. [3], showed the AES-RSA model is the fastest and most efficient hybrid model and the Double AES-RSA model has the superiority among three-tier models. They also found that using the same algorithm, the efficiency of the three-tier model is less than two-tier by an average of 51.04%.

Rawya Rizk and Yasmin Alkady [4], have proposed a Two-phase Hybrid Algorithm (THCA) which is applied to wireless sensor networks. The algorithm is a combination of both symmetric and asymmetric cryptographic techniques, using ECC and AES to provide encryption. Also, XOR-DUAL RSA is added for its robustness against attacks and MessageDigest-5 (MD5) is used to ensure the integrity of the original text during communications. The results of the experiment showed the better security with a shorter time and lower energy consumption.

Fursan Thabit et al. [5], have proposed a Novel Effective Lightweight Homomorphic Cryptographic Algorithm for data security in cloud computing. The algorithm contains two layers of encryption. The first layer uses the new effective, light-weight cryptographic algorithm depended on feistal and permutation/substitution architectural process with Shannon's theory of diffusion/confusion, improving the complexity of the encryption The second layer are multiplicative homomorphic schemes with property of the RSA, providing the algorithm advanced security data in cloud computing. The result of the proposed algorithm outperforms other light weight cryptographic algorithms and some of the symmetric and asymmetric algorithms The proposed encryption algorithm has also experimentally validated for brute force, encrypted text only, known-plaintext attacks and differential cryptanalysis attacks.

Shabana Urooj et al. [6], incorporated both techniques (AES and ECC) along with clustering through standard LEACH protocol for improving energy efficiency, data security, and network lifetime. The proposed algorithm is able to encounter a variety of security threats including side-channel attacks. The proposed algorithm has been proved to be better, compared to current techniques with respect to time complexity, encryption time and decryption time.

Jeffery S. Lee and Gerald B. Cleaver [7], have proposed a cryptographic algorithm with Cosmic Microwave Background (CMB) Radiation functioning as a Random Bit Generator. It is highly applicable for asymmetric and symmetric cryptography. They also concluded that when applied to a Vernam cipher and extended to key spaces in excess of the 256 ASCII character set, the CMB power spectrum offers the capacity for a key space that is too large for not only brute force attacks, but also can be too large for the encryption/decryption capacities of present computer systems.

3 Method

In order to address privacy issues and usage costs while using data while ensuring the encryption effect, this article proposes a P-ATI encryption model based on e-commerce platform data, which is designed as a 4-layer model. The first layer is the User Privacy Layer, which is represented by the P of P-ATI, the second layer is the User Application Layer, which is represented by the A of P-ATI, the third layer is the User Trace Layer, which is represented by the T of P-ATI, the fourth layer is the User Information Layer, which is represented by the I of P-ATI. The model achieves a certain dynamic equilibrium between encryption efficiency and data security. The layering of the model is based on the degree of privacy of the data, which is reduced from top to bottom, as seen in Fig.1:

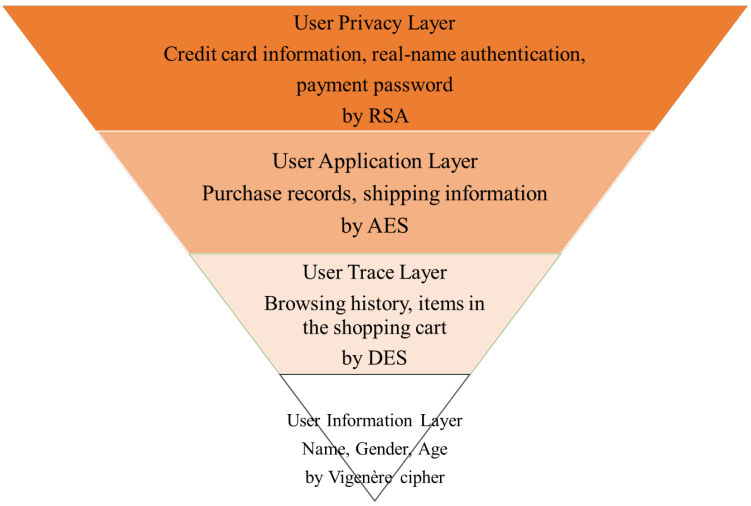

Figure 1. Model Structure

The first layer of the model is the User Privacy Layer, which includes the user's authentication information, bank card information, and payment cipher. This part of information is of extremely high privacy level and closely related to the user's personal interests. Once leaked, it is easy to cause great losses, so the RSA method of extremely high encryption level is adopted. RSA is an asymmetric encryption method that uses public and private keys for encryption. The receiver needs to generate a public and private key before the transmission of message, and then send the public key to the sender. After receiving the public key, the sender encrypts the data to be sent with the public key and sends it to the receiver. After receiving the ciphertext, the receiver can decrypt it with the private key. In this

process, the public key is responsible for encryption and the private key is responsible for decryption. Even if data is intercepted during transmission, attackers cannot crack it because they do not have the private key, which gives the algorithm extremely high security. The flowchart of RSA is shown in Fig.2.

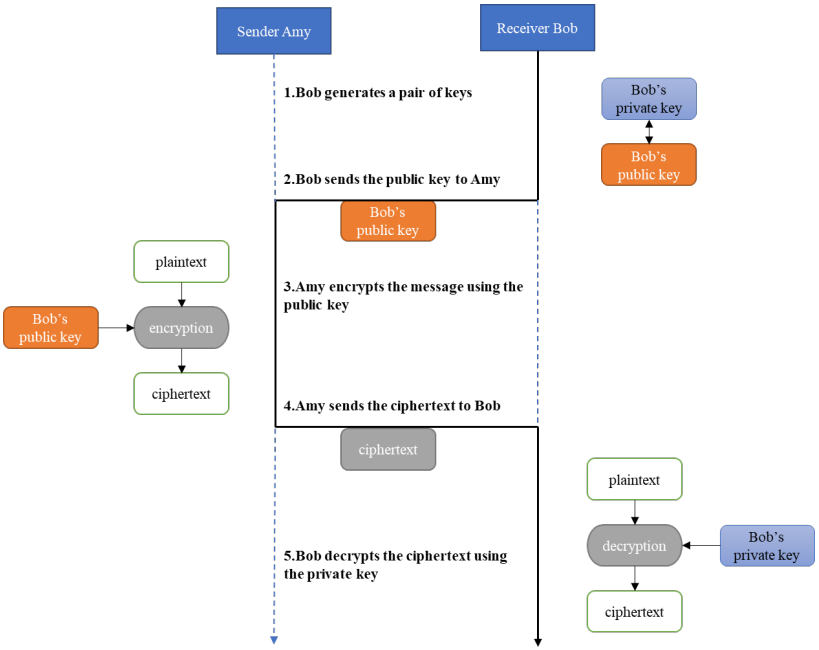

Figure 2. The flowchart of RSA encryption and decryption

The second layer of the model is the User Application Layer, which includes the user's purchase records and delivery information. The privacy level of this information is high, so it requires encryption of high security level, but considering that this part of data is of small volume, we decide to apply AES here. In the model, we adopted AES's CBC encryption mode, which requires an initialization vector (IV) for encryption and decryption. Before each encryption or after each decryption, the initialization vector is used to XOR plaintext or ciphertext. This means that even if there are two identical plaintext blocks, the encrypted ciphertext blocks obtained are different, which makes this mode less vulnerable to attacks and can greatly ensure the data security. The flowchart of AES is shown in Fig.3.

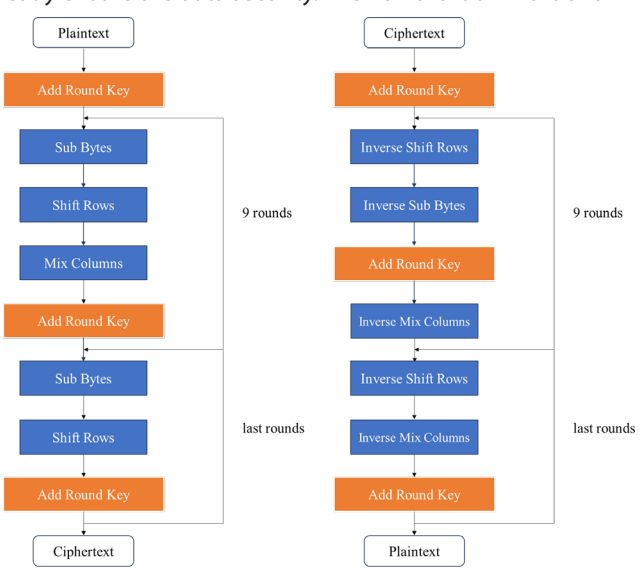

Figure 3. The flowchart of AES encryption and decryption

The third layer of the model is the User Trace Layer, which includes the user's browsing records and the products added to the shopping cart. With the user's permission, it can be shared as public information. Due to the large volume of data in this part, taking efficiency into consideration, DES with fast encryption and decryption speed but rather high security level is used. DES is the most common symmetric encryption method, with the main idea being the permutation and shift process of data bits. The final ciphertext is obtained through 16 iterations of encryption and the final inverse permutation. The decryption method only needs to be solved according to the inverse process of encryption. The flowchart of DES is shown in Fig.4.

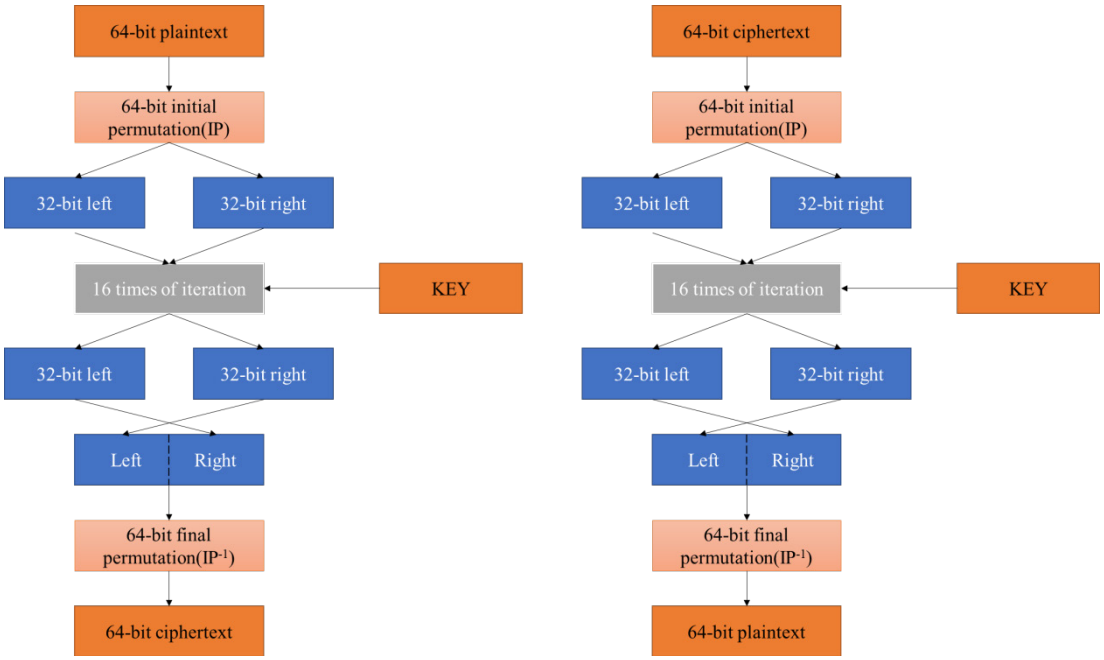

Figure 4. The flowchart of DES encryption and decryption

The bottom layer of the model is the User Information Layer, including the user's name, gender and age. Generally, it is public, so the Vigenère cipher of relatively low protection level is used. The principle of encryption and decryption is to encrypt different letters in the alphabet (as shown in Fig.5) with different displacements. For example; if the original text is *love* and the key is *ok*, the program will change the key by padding it to *okok*. Then, the original text and key will be matched one by one, using both as row and column values in the password table for searching. The resulting letters are ciphertext, and the decryption process is the opposite. If we define n as the Key size, k as the Key, p as the Plaintext, and c as the Ciphertext, the encryption and decryption process can be represented by the following formula:

$$En(p) = (p_1 + k_1, p_2 + k_2, \cdots\cdots, p_n + k_n) mod 26$$
$$De(c) = (c_1 - k_1, c_2 - k_2, \cdots\cdots, c_n - k_n) mod 26$$

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

Figure 5. The Alphabet used in Vigenère cipher

4 Experiment

In this section, we conducted experiments based on the encryption level from low to high.

(1) The bottom layer using Vigenère cipher

In the experiment, we used the information left by user Amy on the e-commerce platform, including public information such as name, gender, age, etc. The data is presented in JSON format as follows:

{Name: Amy

Sex: Female

Age: 25}

When the key is defined as *vig*, the ciphertext output is *Iisz: Ist Aks: Nkhirz Imz: 25*, and the ciphertext can be restored to plaintext through reverse operations. From the experimental result (seen in Fig.6), it can be seen that after inputting information, plaintext will be encrypted into ciphertext, and ciphertext can also be decrypted into plaintext when needed. However, this type of encryption method is relatively simple and only suitable for information with lower levels of protection.

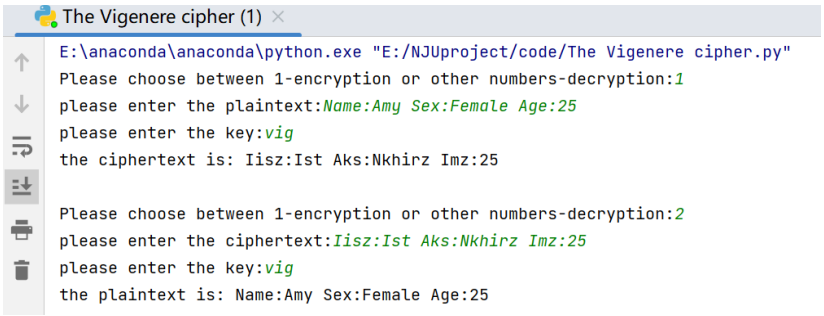

Figure 6. The Experimental Result of Vigenère cipher

(2) The third layer using DES

In the experiment, we used the information left by user Davis on the e-commerce platform, which includes browsing history and shopping cart information, including store name, brand, item and its number. This information can generally be shared publicly with user permission. The data is presented in JSON format as follows:

Browsing History:

{Store: Apple

Brand: Apple

Item: iPhone 13

Item Code: 246327866}

Shopping Cart Information:

{Store: NBA store

Brand: Adidas

Item: training jersey

Item Code: 394678554

Number: 2}

We use shopping cart information for encryption and decryption demonstration. From the experimental result (seen in Fig.7), it can be seen that after defining the key, the input plaintext information will be encrypted into a longer length of ciphertext, and the ciphertext can also be decrypted into plaintext under the same key. This is a characteristic of symmetric encryption methods. This type of encryption already has high security, but the encryption and decryption process consume less time and are of lower computational complexity, making it suitable for information that needs to be protected but has a large volume.

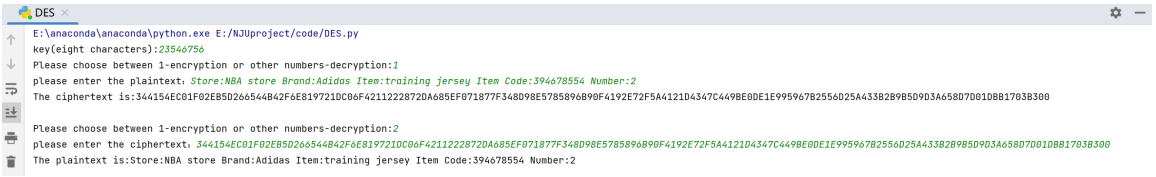

Figure 7. The Experimental Result of DES

(3) The second layer using AES

In the experiment, we used information left by user Belly on the e-commerce platform, including purchase records and delivery information, including store name, brand, item and its number, quantity of items, and user address. This type of information involves the user's private information and is generally not disclosed. The data is presented in JSON format as follows:

Purchase Records:

{Store: sneakers

Brand: Nike

Item: Air Force 1 white

Item Code: 365879346

Number: 1}

Delivery Information:

{Name: Belly

Address: 818 Ashford Ave, Queen, NY 11204}

We use delivery information for encryption and decryption demonstrations. In this section, we use AES's CBC encryption mode. According to the experimental result (seen in Fig.8), after defining the key and initialization vector (IV), the input plaintext information will be encrypted into a longer complex ciphertext. Moreover, since AES is also a symmetric encryption method, the ciphertext can also be decrypted into plaintext under the same key and initialization vector.

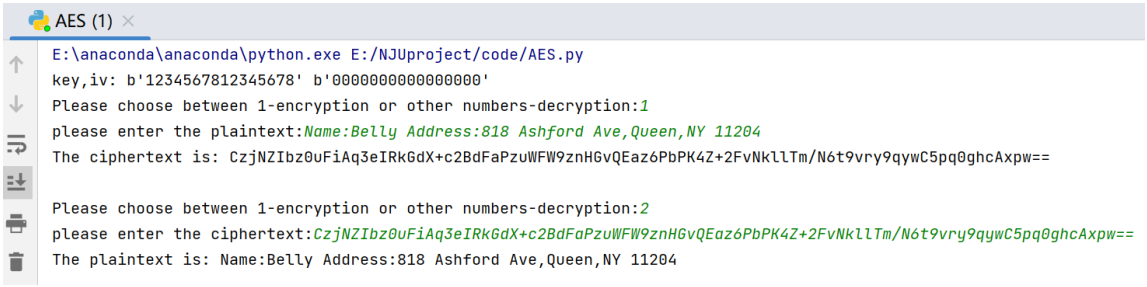

Figure 8. The Experimental Result of AES

(4) The top layer using RSA

In the experiment, we fabricated the information of user Ann, including the user's name, social security number, bank card information, and payment cipher. This type of information is not only related to the user's personal privacy, but also closely related to the user's property security. The data is presented in JSON format as follows:

{Name: Ann

SSN Numbers: 457823497

Credit Card Number: 2341543776456834

Payment Cipher: Ann123456#}

It can be seen from the experimental results (seen in Fig.9) that the ciphertext generated by RSA encryption method is long and complex, and because of the special characteristics of the private key of asymmetric encryption method, it is difficult to crack, thus extremely secure, suitable for storing sensitive private information such as ID number and password. Similarly, decryption of ciphertext can be completed under the same set of public and private keys.

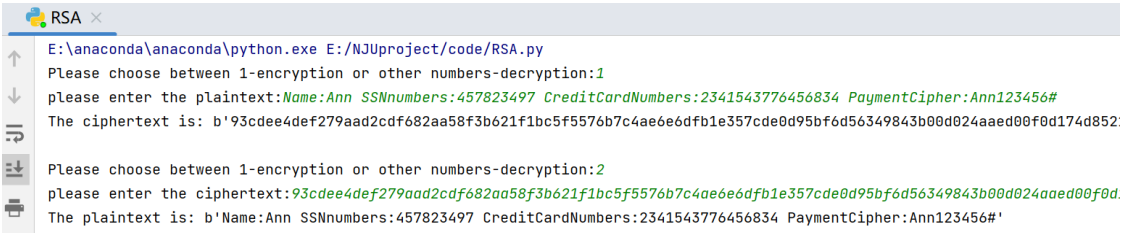

Figure 9. The Experimental Result of RSA

5 Conclusion

With the increasing frequency of people using e-commerce platforms, the leakage of user data may cause greater losses than ever, so the demand for data protection on e-commerce platforms is also increasing. Many platforms use encryption methods such as DES and AES, but their use is not differentiated based on different data types. Therefore, this paper proposes a layered encryption model P-ATI to distinguish different data types, which is divided into User Privacy Layer, User Application Layer, User Trace Layer and User Information Layer. RSA, AES, DES and Vigenère cipher are used for them respectively, achieving dynamic equilibrium between encryption cost and encryption effect. The feasibility of implementing hierarchical data encryption on e-commerce platforms has been verified through experiments, which indicates the further practical application of the P-ATI model.

About the Author

Yutao An is undergraduate of School of Artificial Intelligence, Hebei University of Technology, and his research field is Computer Science.

References

[1] Thabit, F., Alhomdy, S., Al-Ahdal, A. H., & Jagtap, S. (2021). A new lightweight cryptographic algorithm for enhancing data security in cloud computing. Global Transitions Proceedings, 2(1), 91-99.

[2] Chandra, S., Mandal, B., Alam, S. S., & Bhattacharyya, S. (2015). Content based double encryption algorithm using symmetric key cryptography. Procedia Computer Science, 57, 1228-34.

[3] Murad, S. H., & Rahouma, K. H. (2021). Implementation and performance analysis of hybrid cryptographic schemes applied in cloud computing environment. Procedia Computer Science, 194, 165-72.

[4] Rizk, R., & Alkady, Y. (2015). Two-phase hybrid cryptography algorithm for wireless sensor networks. Journal of Electrical Systems and Information Technology, 2(3), 296-313.

[5] Thabit, F., Can, O., Alhomdy, S., Al-Gaphari, G. H., & Jagtap, S. (2022). A Novel Effective Lightweight Homomorphic Cryptographic Algorithm for data security in cloud computing. International Journal of intelligent networks, 3, 16-30.

[6] Urooj, S., Lata, S., Ahmad, S., Mehfuz, S., & Kalathil, S. (2023). Cryptographic Data Security for Reliable Wireless Sensor Network. Alexandria Engineering Journal, 72, 37-50.

[7] Lee, J. S., & Cleaver, G. B. (2017). The cosmic microwave background radiation power spectrum as a random bit generator for symmetric-and asymmetric-key cryptography. Heliyon, 3(10).