ISSN 2278-3091
Volume 9, No.4, July – August 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse109942020.pdf
https://doi.org/10.30534/ijatcse/2020/109942020

A New Lightweight Proposed Cryptography Method for IoT

Auday H. AL-Wattar[1]
[1] University of Mosul, Iraq, Ahsa.alwattar@uomosul.edu.iq

ABSTRACT

The key challenge to information security is the requirement for using untraditional philosophies and alternative means, as concentrating on new aspects to achieve security. In this article, a new proposed method has been proposed, using the stored biological data on GenBank with its features. This article uses the new aspects for achieving the cryptography depending on this data. The proposed method can be used in many new fields, especially in IoT and smart cities. The limited resources of the smart devices that form the basis of the IoT, particularly in terms of power and storage, showed the need to find a new method to ensure confidentiality of these devices.A set of analysis based on calculations has been used to test the new technique. The results and the analysis indicate that the new technique is secure.

Key words : Cryptography, Security and GenBank, IoT.

1. INTRODUCTION

Information security has always been one of the most vital matters taken into consideration due to its significant role in life. This interest has grown significantly after the emergence and adoption of computers in most fields of modern life. Computer security is a comprehensive label that includes a set of behaviors, methods, means, and tools intended to preserve, defend, guard and protect information and data of computer systems by preventing hackers from accessing them illegally. A secure connection is essential to achieve trusted interchange of data among any partners. The internet has been the backbone of all electronic commerce transactions and banking. The emergence of the Internet of Things has added significant security challenges, which are based on finding appropriate ways to achieve security, especially that the IoT requires a special environment and specific conditions must be considered. In order to meet these security requirements, many techniques, and systems have been developed within traditional cryptographic methods, particularly in the mathematic domain of data encryption and decryption. Untraditional ways are used to overcome these techniques utilizing some new lightweight cryptography techniques and methods. Scholars have attempted to practice new methods based on bio□techniques cryptography and steganography. In this article, a new method is proposed through employing GenBank DNA data as lightweight encryption technique.t a section that you want to designate with a certain style, then select the appropriate name on the style menu.

1.1 Cryptography

Traditionally, encryption is known as methods enabling two or more parties to establish secret communication over an unconfident channel that is visible to eavesdropping. A large domain of security targets is accomplished by means of cryptographic techniques [1]. Cryptography was formerly limited to the dominions of government, military, and the academic world. Due to technological advances, cryptography has started to penetrate all aspects of everyday life.

All items, whether tiny chips or banking, rely heavily on cryptography to keep information safe.

Cryptography may be classified into three categories of cryptographic algorithms according to the type of encryption key used; Symmetric key encryption, keyless encryption and Asymmetric or public key encryption.

1.1.1 Symmetric key encryption

Symmetric key encryption was the only encryption technique used before the evolution of the public key in 1970 [2][3]. When same key is used for both encryption and decryption, it is called shared-key.

1.1.2 Keyless encryption

It is called keyless cryptographic function, and includes random number generation and cryptographic hash function.

1.1.3 Asymmetric (public) key encryption

Different keys are used for encryption and decryption where a pair of keys - public and private keys - is used for encryption and decryption. An asymmetric encryption has a single-direction function. Such a function is effortless to do in one way, but complex or impossible to reverse. Its security depends on an extremely great difference in complexity of computation between a mathematical calculating function and its inverse. any additional blank lines between paragraphs.

1.2 GenBank

According to [4], "GenBank® is an extensive public database that includes publicly available nucleotide sequences to support bibliographic and biological notations. GenBank is a portion of the International Nucleotide Sequence Database Collaboration, which includes the European Nucleotide Archive (ENA), the DNA Data Bank of Japan (DDBJ), and GenBank at NCBI.

It is constructed and publicized by the National Center for Biotechnology Information (NCBI), a section of the National Library of Medicine (NLM), situated in the campus of the US National Institutes of Health (NIH) in Bethesda, MD, USA.

NCBI generates GenBank mainly from authors' sequence information submissions and bulk submissions of whole genome shotgun (WGS) and other high-throughput information from sequencing centers. Sequences from issued patents are also provided by the US Patent and Marks Office. As a partner in the International Nucleotide Sequence Database Collaboration (INSDC) [5], GenBank works with the EMBL-EBI European Nucleotide Archives (ENA) [5] and DNA Data Bank of Japan (DDBJ) [7]. The INSDC partners exchange data on a regular basis to guarantee that the global sequence information collections are uniform and thorough. NCBI provides GenBank information free of charge over the Internet, FTP and a broad variety of web-based facilities for analyzing and recovering [8].

A number of DNA Cryptography algorithms have restrictions in that they still employ segmental arithmetic cryptography at some of their stages, or they are biological workroom research-based which is not appropriate for the digital computing background. To get over this gap, we propose a new, secure, encryption and decryption method with an analysis of its performance. This cryptography uses the data segments available on GenBank for encryption and decryption.

This Paper is divided into the following four sections: Section One (the introduction), Section Two (the proposed method), Section Three (discussion), and Section Four (the conclusion).

1.3 Public key principle

Encryption and decryption are carried out using two different keys. The two keys in such a key pair are referred to as the public key and the private key. This approach allows publicity of the encryption key; anyone can access the public key for encrypting a plaintext to the recipient. The scenario is as follows:

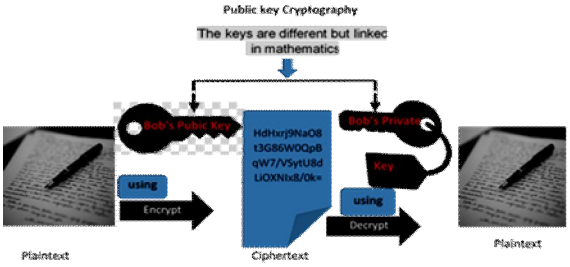

Figure 1: Public key Cryptography Principle

Formally:

f(x) is a one-way function from a set X → set Y, so that f(x) is easily computed by all x ϵ X, but it is "computationally ineffective," to locate any x ϵ X, such as f(x) = y for "fundamentally all" elements y ϵ Y.

1.4 Internet of things (IoT)

The Internet of Things (IoT) is a network of physical items or "things" integrated with electronic products, devices, sensors, and network connectivity that allows these items to accumulate, send and receive data. The Internet of Things (IoT), in some cases it is called Internet of Everything (IOE), comprises of all the web-enabled equipment that assemble, transmit and exchange the data they gain from their nearby environments employing specific tools including hardware communication and embedded sensors and tiny CPU's work as processors [9].

They are processing, storing, and communicating delicate, sensitive, basic data, so security is a countless challenge to protect data from any vulnerability. It is a platform of connected physical objects that can be accessed via the Internet[10]. It is often considered as the Machine-to-Machine (M2M). The Internet of Things components rely on the sensors, communications, people, and processes used. It is currently facing serious challenges that can be summarized in terms of scalability, specifications that must be met, technical implementation, software complexity, and IoT security. A number of solutions are implemented to resolve the challenges. In the term of security IoT needs powerful encryption algorithms, authentication mechanisms, and a framework that can predict and control inconsistencies or discrepancies on the network. the receiver malicious tampering to the receiver.

1.5 Man in the middle attack (MIMT)

The attack by MITM specifies that the individual at the center interrupts the transmitter data, and after malicious manipulation, the data is transmitted to the receptor. This MITM attack sends the wrong information over the network and has significant implications.

1.6 Black hole attack

The black hole attack is usually a Service Denial (DoS) attack which is one of the most manifest attacks. the Black Node occurred during the route exploration process; the sender does not know the path to the receiver at first. A malicious node utilized its routing protocol to declare that the node gets the shortest path to the target node, while there is no route to the recipient of the black hole node. In that case, in the data route, the black hole node is present. If the paths have been set, the transmitter sends the packets to the black hole node and begins the packet drops without transmitting them to the destination node.

2. The Proposed Method

The proposed method can be represented using a protocol between two parties (Alice and Bob), as outlined in Figure 2.

The proposed method is composed of the following elements:

- Sender (Alice).
- Receiver (Bob).
- Message (m): (plaintext)
- Encrypted message (C): ciphertext
- GenBank's: DNA banks
- Si: ith DNA segments within GenBank's
- SBob: Private key of the receiver
- PBob: Public key of the receiver

The GenBank contains DNA Data bases. Each segment has a certain location and value (a number of bases with a specific length). Locations of values of segments will be used in the proposed method as private and public key of the receiver (Bob). The following section will describe the scenario of the work in detail.

2.1 Work Scenario

Generally, the sender (Alice) will encrypt the message m using the receiver's key (Bob public key) which is obtained from the GenBank's. The Receiver (Bob) will use his private key (SBob) to decrypt the message *m*. The public key of Bob (PBob) will be distributed to the trusted parties using the same protocol.

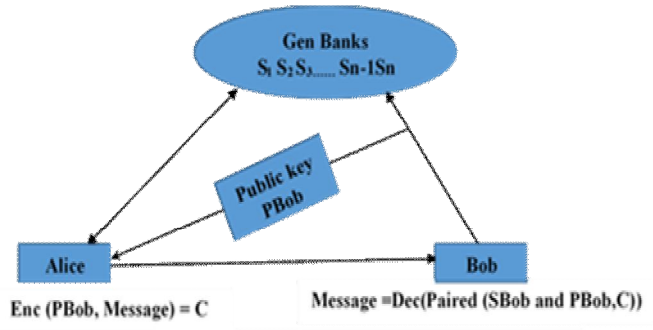

Figure 2: Scenario of the proposed method

The Bob's private key (SBob) can be one or more DNA segment (s1, s2,.., Sn) in the GenBank, i.e. the value of the chosen DNA segment with some DNA bases and specific length. While the receiver public key (PBob) will be the location of the chosen DNA segment. Thus, the pairs of keys will be as follows:

- The location of the segment within GenBank represents Public key of Bob (PBob).
- The value of the chosen segment itself represents the secrete key of the receiver Bob (SBob).
- Only the receiver (Bob), has the ability to use chosen DNA segments (its value) to calculate the private key (SBob).

The proposed structure for the keys (PBob) and (SBob) is shown in Figure 3.

- The SBob can be described as a specific DNA segment within GenBank with its value and length that can be any sequence of DNA bases with specific length.

As SBob → SiL.

Where:

Si → the specific DNA segment S within the GenBank, and, L → the segment length.

The key SBob can be obtained from the chosen DNA segment with their length. Many techniques done by the receiver can be used to get this key. For example, the receiver (Bob) may use an exact mathematic method to get a specific segment or a part of segment, a chaotic or any approach could provide a random access to specific bases of the segment. Generally, the receiver can use any technique to produce the key or to use it.

The receiver (Bob) knows this key, while the sender (Alice) does not.

The sender (Alice) knows the public key of the receiver which is the location of the DNA segment within the a specific GenBank. As MGN (Si).

Where: MGN(Si) → Location of the segment Si within the GenBank

So, the message m can be decrypted by using the paired public and private keys (the location of the DNA segment within a specific GenBank (PBob), and the specific DNA segment with its bases (SBob). The DNA segment could be of any size depending on the parties (sender and receiver).

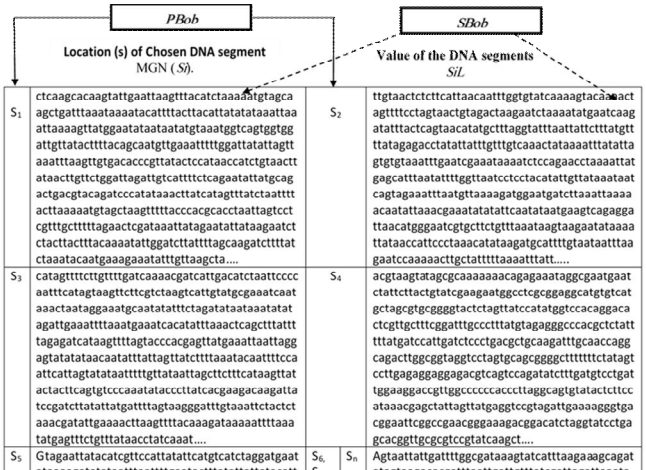

Figure 3: PBob and SBob using the DNA segments in GenBank

The steps of the proposed method for the both sides (Alice and Bob) can be summarized in the figure 4, which illustrates the encryption and decryption processes in a form of algorithm.

Alice	Bob
1. get the plaintext (m)	1. get the ciphertext (c)
2. encrypted m using PBob MGN(S_i) MGN = The location(s) of the DNA segments within GenBank	2. decrypted c using a. the PBob MGN(S_iL), b. SBob (S_iL) S_iL = The value of the DNA segment and its length.
3. send the ciphertext (c) to Bob	3. get the Plaintext (m)

Figure 4: The steps of encryption and decryption methods on both sides (Alice and Bob)

The principle of the proposed method compared to the original Public key scheme is summarized in figure 5.

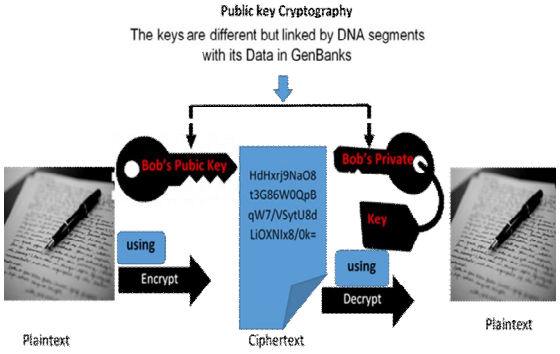

Figure 5: The proposed Public key Cryptography Principle

The proposed encryption and decryption processes utilize the public key concept, without the need to use module or elliptic curve or other calculated methods. The huge amount of DNA data which are available on GenBank makes it possible to be applied. For the sender (Alice), the location of the segments is only known as a key and nothing else, and the value of the DNA segment and its length are known by the receiver (Bob) only. The process of evaluating DNA value to use it as private key is also is related to the Bob alone.

3. DISCUSSION

The proposed method is more secure since it can be used to encrypt and decrypt data without the need to transfer the secret key on any public channel. The public key will be sent to many parties depending on a specific agreement. The sender will use this key to encrypt the massage m. This key can be just a number or a code without any meaning, and this number or code is used to specify a location within a GenBank which contains a huge amount of locations with each holding a billion number of DNA segments. This is one of the advantages of GenBank as a factor for our proposed encryption method because the attacker does not have the ability to analyze this key even if he knows it. In addition, it is impossible for the attacker to check all locations in GenBank in terms of computing capability and time. Furthermore, the sender will know nothing about the secret key of the receiver, so the secret or the private key will be known by the receiver only.

The receiver uses his public and private keys to decrypt the ciphertext. He uses the public key to locate the chosen segment within the GenBank, and uses the value of the chosen DNA segment to obtain his private key to use in the decryption process. This private key can be in any form depending on the techniques or approaches the receiver using to obtain the private key. Regardless to approaches used, it will give a robustness security since this key represented by the DNA segments and techniques are known by the receiver only.

According to [3], on June 2019 the GenBank can have more than 213383758 DNA sequences, and 329835282370 DNA bases. So, the probability of obtaining the location of one segment from these segments is very difficult in terms of calculation, where this segment represents the chosen DNA segment used as a public key. The search process will be so consuming since it has to try all possible locations. Moreover, even if the attacker was able find the exact location, i.e. get the public key which could be any one of these locations, he will get nothing than a fuzzy meaningless number. Also, the private key which could be any group of the DNA bases is also difficult or impossible if we know that each base could be represented by two bits and the 4 bases form one byte. As 00-A, 01-C, 10-G, and 11-T. So, the private key will be a group of DNA bases and a chosen technique done on them by the receiver. The number and the locations of DNA bases in which be used as a private key is known only by the receiver himself.

329835282370 DNA bases, each of which can be A, C, G, T. So, the probability to get the value of the DNA chosen DNA segment will be 4329835282370 if the segment length consists only of 4 DNA bases. This probability will increase if we use the binary coding for the DNA bases as every base can be represented by two bits with deferent coding as in table 1: The publicly availability of these NIH genetic sequence database makes it available on line which give an advance to use it to obtain the keys anytime and anywhere.

For the attacker, it is very difficult to estimate the public and the private keys since it has to check all the GenBank over the whole. He also has to get the value of the DNA segment among billions which can be considered as impossible issue. Finally, he has to know the approach used by the attacker which is only known by the receiver himself. If the private key was 4 DNA bases then the attacker will be need to try (4329835282370)16! possible bases, in this case, the DNA bases were treated as one unit that consists of 4 bases, of course the possibility will increase dramatically if these 4 bases were not one unit, but instead were collected depending on a special sequence as a key. No mathematics or other calculations are required for achieving the proposed encryption method. However, it could provide a strong security using some biological issues as the DNA data within

GenBank as well as the properties of the DNA sequences that stored within these Banks

3.1 Work Scenario

Information connections can be attacked in different manners, the man in middle attack is one of the most popular attacks, this attack defined as the person in the midst breaks off the data hand on by the sender pretended to be the real sender and send

The attacker of MITM as well as the black hole attempts to interfere with the data, If the attacker attempts to access the information, the data must be decrypted. However, the attacker has no chance of knowing the data provided by a sender in transmission, In addition an increase in the DNA data volume contributes to data storage and data privacy problems, but the attacker has no idea of the private key of the receiver or any information about the transmitted data. Besides, the attacker receives the encrypted information cipher text. But the DNA code is altered Therefore, the attacker cannot interrupt the encrypted information from the previous plaintext and cipher text.

4. CONCLUSION

This article proposed a new cryptography method through using two keys (public and private) for encryption and decryption. The public key is used to encrypt the message by the sender, while the receiver uses both public and private keys to decrypt the message. The sender does not know anything about the receiver and the private key is not exchanged using any channel. The method utilizes the Banks of DNA data and the segments stored in these banks. Based on the analysis, this method provides a robust security with less mathematical calculations. The simple requirements of the proposed method make it. The simple requirements of the proposed method make it possible to be employed in the field of IoT security since the connected smart devices has a limited resource in term of power and storage. “affect” (usually a verb) and “effect” (usually a noun), “complement” and “compliment,” “discreet” and “discrete,” “principal” (e.g., “principal investigator”) and “principle” (e.g., “principle of measurement”). Do not confuse “imply” and “infer.”

REFERENCES

1. M. Bishop and B. Matt, **Introduction to computer security:** Pearson Education India, (2006).
2. S. William and W. Stallings**, Cryptography and Network Security, 4/E**: Pearson Education India, (2006).
3. K. Sailaja, R. Srinivasa, and P. Ramesh, **A New Circle based Symmetric key Encryption Technique for Text Data**, International Journal of Advanced Trends in Computer Science and Engineering, vol. 8, pp. 2573-2576, (2019). https://doi.org/10.30534/ijatcse/2019/106852019
4. D. A. Benson, M. Cavanaugh, K. Clark, I. Karsch-Mizrachi, J. Ostell, K. D. Pruitt, et al., **GenBank," Nucleic acids research**, vol. 46, pp. D41-D47, (2018).
5. I. Karsch-Mizrachi, T. Takagi, G. Cochrane, and I. N. S. D. Collaboration, **The international nucleotide sequence database collaboration**, Nucleic acids research, vol. 46, pp. D48-D51,(2017). https://doi.org/10.1093/nar/gkx1097
6. N. Silvester, B. Alako, C. Amid, A. Cerdeño-Tarrága, L. Clarke, I. Cleland, et al., **The European nucleotide archive in 2017**, Nucleic acids research, vol. 46, pp. D36-D40, (2017).
7. Y. Kodama, J. Mashima, T. Kosuge, E. Kaminuma, O. Ogasawara, K. Okubo, et al., **DNA data bank of japan: 30th anniversar**y, Nucleic acids research, vol. 46, pp. D30-D35, (2017).
8. N. R. Coordinators, Database resources of the National Center for Biotechnology Information, Nucleic Acids Research, vol. 46, pp. D8-D13, (2017).
9. S. Byun, **Gateway-based Resource Control for Reliable IoT Environments**, International Journal of Advanced Trends in Computer Science and Engineering, vol. 8, pp. 1881-1885, (2019). https://doi.org/10.30534/ijatcse/2019/11852019
10. W. Hussein, L. Kamarudin, H. Hussain, M. Hamzah, and K. Jadaa, **Technology Elements that Influence the Implementation Success for Big Data Analytics and IoT-Oriented Transportation System**, International Journal of Advanced Trends in Computer Science and Engineering, vol. 8, pp. 2347-2352, (2019). https://doi.org/10.30534/ijatcse/2019/74852019