**International Journal of Science Academic Research**
Vol. 04, Issue 06, pp.5756-5760, June, 2023
Available online at http://www.scienceijsar.com

**ISSN: 2582-6425**

**Research Article**

# AN ANALYSIS OF THE USE OF DNA COMPUTING IN CRYPTOGRAPHIC ENCRYPTION METHODS

***Karan Chawla**

Ashoka University, 131029, India

**Received** 28th April 2023; **Accepted** 24th May 2023; **Published online** 26th June 2023

**Abstract**

In the newly growing discipline of DNA computing, DNA is employed as a data store medium and biological computation is used to answer problems. The study of DNA transcription and replication is where the idea originated. DNA computing is one type of biological computing. In the field of computer science, the use of DNA molecules as a data storage and processing medium is a relatively new concept. The interests and security of humanity have recently been significantly challenged by regular network assaults. Multiple detection techniques have been researched to counter this hazard, some of which have had promising outcomes but due to the advancing nature of network connectivity technology, enormous volumes of network data and a great deal of redundant data have been generated. The sample sizes for each attack type in the dataset are severely out of balance due to the difficulties in gathering samples due to the regularly changing forms of assaults. The robustness of current detection techniques is significantly decreased by these two issues, and current research techniques do not offer a satisfactory answer. Similarly, Electronic health records (EHRs) are being exchanged across the open network in historic numbers in this developing era where encryption is a trusted method to safeguard data in EHRs but the issue is that they are vulnerable to a variety of security risks. This paper addresses the application of dna computing in cryptographic encryption methods for NSGA3, Medical Imaging and Cloud Computing.

**Keywords:** Hand Strength, Grasp weight cuff, Hand grip, Hand fatigue rehabilitation, Stroke hand function.

## INTRODUCTION

DNA is used as a data storage medium and biological computation is used to address issues in the emerging field of DNA computing. The concept comes from the investigation of DNA transcription and replication. However, it has only been used for straightforward computational tasks like locating the best answers to combinatorial puzzles. DNA-based storage computing is one of the most disruptive data technologies because it employs biological and molecular hardware instead of conventional electronic computing. A form of biological computing is DNA computing. The use of DNA molecules as a data store and processing medium is a relatively recent development in the world of computer science. Molecular computing and biological parallel computation are other names for DNA computing. The goal of DNA computing is to use nature's capacity to carry out complicated operations at cheap cost, with minimum energy use, and with no environmental impact, rather than to replace conventional computers. DNA computing can assist in the resolution of issues that are challenging for electrical computers and humans to handle, such as: forecasting future climatic changes on Earth by inputting all relevant data into the computer system; launching an artificial satellite into orbit to use its telescope to look at various things there in order to determine if there is any life there; evaluating if drinking water from lakes or rivers that may have been contaminated by pollutants is safe for people to consume; making fresh medications from substances already present in plants or animals DNA computing has a highly promising future. In fact, it may be computer science's next big thing. Compared to electrical computing, DNA computing uses a lot less energy to do computations. We now have new ways to think about information management, which may result in more effective methods of storing data on hard drives and improved methods of interfacing with gadgets like smartphones and laptops.

## NSGA3

DDoS attacks, advanced persistent threats (APT) attacks, the use of ransomware, data theft, and other intrusions have mushroomed in recent years, paralysing various networks and interconnected devices, leading to frequent network security incidents worldwide, and becoming a nightmare for people, businesses, and even the national security. Traditional attack detection techniques, however, are failing due to the aberrant network traffic's rising complexity, concealment, and variety as a result of the network's exponential expansion. To counter this trend, academics have created a variety of attack detection techniques based on deep learning, machine learning, and rough set theory. Traditional detection techniques, for instance, have the benefit of being able to identify network assaults even in the presence of sparse data since they are based on rough set theory. Compared to conventional detection techniques, machine learning-based approaches can improve classifier performance through continuous learning and the accumulation of experience. When building a nonlinear network structure using hidden layers, deep learning may first understand the underlying principles and expression levels of training data. This is not restricted to a fixed problem and offers the advantages of increased efficiency and autonomous model development. However, these most recent detection techniques ignore feature subset optimisation and detection imbalance, focusing exclusively on the overall detection accuracy and feature dimension reduction of network assaults. For instance, the amount of feature information it includes grows when network traffic grows quickly. However, this data is muddled with additional noise and irrelevant data. The close relationship between this noise and redundant information and the useful information makes attack detection more computationally complex and time-consuming, and it also

***Corresponding Author:** ***Karan Chawla***
Ashoka University, 131029, India

causes some of the useful information to be deleted while some of the noise and redundant information is retained during feature dimension reduction. This poses significant difficulties for assault detection technologies. The actual number of incidents is low, and data imbalance is pervasive, as several attack types in the available detection dataset are impossible to sample. The problem of detection imbalance is made worse by the fact that earlier feature selection techniques tended to ignore associations between tags with few instances and features while favoring interactions between tags with numerous instances and features. A dataset with a class imbalance is made up of samples from the majority and minority classes. The robustness of the detection classification system is severely impacted since detection systems trained on such datasets frequently produce detection results with significant errors for a small number of attack categories. A type of bionic optimisation technique called DNA computing is based on biological DNA coding and evolutionary principles. For the purpose of tackling challenging combinatorial optimisation issues, this method works incredibly well. The fact that it takes use of both the high parallelism of biological reactions and the enormous store capacity of DNA molecules is its biggest benefit.

Studies revealed that DNA computing offered benefits including rich population diversity and a quick convergence rate. Additionally, they simulated different processes using DNA molecules to find and process information while concurrently acquiring and updating information in the evolution process and employed DNA codes to represent the information held by the system. In practical domains, multi-objective optimization is a frequent issue. For instance, the several sorts of cyberattacks that might be encountered include regular, DDoS, worm, R2L, and generic assaults. Therefore, there are at least 5 objectives in the optimization of assault detection. However, there are only 2 or fewer goals in the current feature selection schemes. For instance, only one or a maximum of two targets may be dealt with by the genetic algorithm (GA), NSGA, and NSGA2. Complex computations and inefficient reorganization processes are inevitable when dealing with optimization issues involving three or more goals. However, NSGA3 can be effective for three to fifteen objectives and can ensure benefits like strong searchability while offering a variety of alternatives. The fact that it doesn't require any additional settings is one of its primary features. Without adding any additional parameters, the method may be expanded to increase processing limitations, and the computational efficiency is significantly increased. By using this method, the reference point set's association state, which is based on the reference point, is updated adaptively in real time over several generations.

DNA coding comes first in DNA computing, then biological processes. All DNA sequences are based on the four different types of DNA base coding, and the feature set is first initialized at random as the parent population of the DNA sequence. Then, through DNA biochemical processes including crossover, mutation, and inversion, new individual and progeny populations are created. Calculating the categorization performance index is the second step. The relevant base set is chosen as the feature subset once the DNA sequence has been decoded. After that, a hyperparameter optimization technique is used to optimize the detection model. Finally, values for the imbalance index and the ideal feature subset index are calculated after calculating the accuracy of various classes. The fitness level of the DNA sequence is calculated using these data. Population selection is done as the third phase. The populations of the parents and offspring are first merged. Then, using a nondominated sorting algorithm based on reference points, all DNA sequences are arranged in order of fitness. The parent population of the following generation is then chosen in accordance with the findings of the sorting. In a biochemical environment, DNA biochemical reactions take place between DNA sequences or at the molecular level of humans. These reactions in this article include base exchanges across DNA sequences, base or sequence exchanges inside a single DNA sequence, inversions of specific sequences, and base variations. The population is initially converted into a DNA chain population by adding DNA molecules, and the pertinent data is simultaneously stored in each DNA chain. Before achieving a stable population state, the DNA population goes through orderly selection, crossover, mutation, and reverse ordering processes in a particular biochemical environment. The required problem's answer is then discovered by decoding the DNA chains. A DNA sequence's fitness is a gauge of whether it has a chance of surviving in the DNA population. A DNA sequence's fitness is evaluated in this article using the classification performance index, which primarily considers the imbalance index, the ideal feature index, and the accuracy of the DNA sequence. Because it bases its DNA evolutionary search only on fitness rather than upon outside knowledge, the categorization performance index is crucial to the development of the DNA population. The K-nearest neighbor (KNN) method was utilized as the primary fitness calculation strategy and was augmented by hyperparameter optimization to rapidly and efficiently acquire the most accurate fitness values.

## Medical Imaging

The pace of technological progress during the last few decades has skyrocketed. Information may now be transmitted across the internet more easily because of developments in communications. Digital photographs, movies, and records are just a few examples of the data that is transported instantly around the globe. Telemedicine is the most extraordinary answer to the current healthcare dilemma in the medical industry. Digital medical pictures of internal human organs are utilized for quick diagnosis and effective therapy. Healthcare professionals must handle sensitive patient information due to cybercrimes. In this case, encryption methods are required to guarantee the secure transfer of private medical data. Information is encoded using a secret key during the encryption process so that only permitted individuals may access it. Due to the large pixel capacity and high redundancy of medical pictures, traditional security procedures like Advanced Encryption Standard (AES) and Data Encryption Standard (DES) are ineffective for encrypting them. For the encryption process to be completed, these techniques require a lot of processing time. Due to its great randomization performance, a chaotic system has recently captured the attention of many academics for usage in various cryptosystems. In this case, encryption methods are required to guarantee the secure transfer of private medical data. Information is encoded using a secret key during the encryption process so that only permitted individuals may access it. Due to the large pixel capacity and high redundancy of medical pictures, traditional security procedures like Advanced Encryption Standard (AES) and Data Encryption Standard (DES) are ineffective for encrypting them. For the encryption process to be completed, these techniques require a lot of processing time. Due to its great randomization performance, a chaotic system has recently captured the attention of many academics for usage in various cryptosystems. This may be altered in the spatial and transform domains to secure a digital medical picture. Numerous researchers have experimented with encryption methods in this field by integrating chaotic systems with DNA approaches. a combined 2D linked logistic map that nearly reached 8 entropy and improved the security level of DICOM pictures. The Unified Average Change in Intensity (UACI), in contrast, has fallen short of the predicted value. S Sun proposed a unique DNA-based encryption method that incorporates DNA encoding, bit- and pixel-level scrambling, and is resistant to known plain text assaults.

The use of DNA sequences and IWT in a dual-domain strategy has been proposed as a novel method. A combined chaotic 3D Lorenz attractor and logistic map with a sufficient entropy of 7.998 were used to achieve the goal. The special qualities of image encryption in the transform domain make it stronger than image encryption in the spatial domain. Guan et al. used a combined technique of 4D hyperchaotic maps and DNA encoding to accomplish picture encryption in the frequency domain. Modern encryption approaches have been studied, therefore the cryptosystem must be resistant to statistical and differential assaults. It must be extremely careful with its secret keys and use a bigger keyspace in order to prevent data retrieval by an unauthorized user. The major characteristics of the suggested method are investigated in terms of statistical attack, selected plain text assault, keyspace, key sensitivity, differential attack, encryption quality, and visual quality analysis due to the rapid growth of DNA cryptography. By using the IWT process, a DNA-fused chaos algorithm with two stages of shuffling and diffusion has been developed. To evaluate the complexity and security of the created approach, many indicators are assessed. Additionally, the known and selected plain text attack analysis is reviewed for a complete black and white image and produces an entropy of 15.1692 to verify the algorithm's resilience. Because of its high level of security, the suggested algorithm will thus be effectively adopted or viewed as a better alternative than other algorithms. Future research may be carried out on a Field Programmable Gate Array to achieve a high level of parallelism with tamper-proof picture encryption architecture.

## Cloud Computing

In the era of large-scale computation, when a great number of distributed and parallel computers are networked, cloud computing is one of the benchmarks. In this case, a number of technologies, including utility computing, virtualization, network systems, and distributed processing, are coupled to provide a number of services, including a server, space, network hardware, pay-per-use, and many more. In a cloud environment, there are three primary types of entities: (i) the Data Owner (DO), (iii) the Cloud Service Provider (CSP), and (iii) the User. The DOs and users use the cloud services offered by the CSP, and numerous users keep their private data on the cloud server here. Today's cloud environment is anticipated to be able to manage massive data, which is a key competency. Big data is now regarded as one of the most advanced study fields. Big data is a broad category of data that includes both organized and unstructured information. It must be gathered, handled, examined, and visualized. Big data uses sophisticated methodologies and strategies for data processing and information extraction. However, due to the abundance of attackers and bad users, access control and data security are two of the most pressing issues with big data in a cloud setting. A method through which a user or customer can access any data or any type of service from the cloud server is known as access control. The CSP detects all attempts to access any cloud server service using an effective Access Control Model (ACM) and only permits authorized users to access large data or cloud services.

Since the CSP serves as the central authority or administrator for all cloud server operations, it should offer excellent security for the user's sensitive or secret huge data. One of the most popular options for large data security is cryptography. Cryptography's main objective is to transport data or messages between sender and recipient across an unreliable channel in a way that prevents an attacker or malicious user from reading the original data content. Based on how encryption and decryption keys are used, there are two types of cryptosystems: symmetric key encryption and asymmetric key encryption. The encryption and decryption steps are carried out using the same key in symmetric key encryption. Different keys or values are used to encrypt and decode data or messages in asymmetric key encryption. In this case, a private key and public key are given to each person separately. While the public key can be distributed to anybody, the private key must always be kept a secret. Many academics have recently put up various plans to enhance data or information security and boost the functionality of the cloud environment. One of the most cutting-edge disciplines is DNA computing, which uses hardware, DNA, molecular biology, and biochemistry to encode genetic information in computers. Adleman was the first to employ DNA in the realm of computing. This sophisticated method was initially used to resolve NP-hard issues. However, it was quickly realized that DNA computing would not be the greatest approach to handling this kind of issue. Following that, several methods for employing DNA computing to solve numerous issues, including the SAT problem, the 0-1 planning problem, the integer planning problem, the optimum problem, graph theory, and various Turing machines, have been devised. The secret key creation is done by the data owner. The DNASK is only generated by the DO for legitimate or authorized users. After confirming the user's legitimacy using the CSP, the DO retrieves crucial user characteristics.

The DO divides the plaintext into 1024-bit blocks before starting the large data encryption process. In this case, each 1024-bit block is divided into four equal blocks with 256 bits each. After that, an EXOR operation is carried out between the 256-bit blocks of the plaintext and the key. As a result, each block of plaintext that is 1024 bits long undergoes four EXOR procedures. Numerous tests are conducted to compare the performance of the proposed system to that of the current methods, including the client-side data encryption scheme, the zigzag Morse code-based ACM, and the reversible data concealment scheme. These schemes are used as the benchmark schemes for comparisons since they were only recently introduced and are closely equivalent to the proposed scheme. When requesting any huge data in DNABDS, all users must make requests to the appropriate DOs for the secret key and user access rights. Therefore, it takes some time for the suggested technique to create the DNASK. Once the users of DNABDS have the secret key, they can use it going forward.

Thus, the time required for key creation is decreased for current or past users who already possess a secret key based on DNA. As a result, DNABDS generates a zigzag curve. Secret key generation in ZMCACM involves a number of steps and takes a lot of time. To create the secret key based on a DNA sequence, both RDHS and CSDES employ several EXOR processes. In addition, all of these existing techniques are far more complicated than the one that is being presented, and all users or customers must obtain the secret key each time they access huge data in order to continue using it for other purposes. As a result, all of these current methods encounter linear growing curves. Big data security concerns in the context of cloud computing are growing daily. This research suggests a lengthy 1024-bit DNA-based secret key or password creation approach based on the user's hidden traits. Big data based on DNA computing is encrypted using the same key. The confidential credentials are only shared with authorized users, and the secret key is safeguarded by labeling the randomness during DNA base assignment. The suggested technique can withstand a variety of assaults in a cloud setting. In this case, the data owners can be online while supplying the DNA-based secret key and access certificate, and they can become offline after the credentials have been given.

As a result, the system overhead is reduced.

## Conclusion

Various strategies to identify cyberattacks on cyber networks and systems have been put forth by numerous researchers. Due to an imbalanced number of classification cases, the techniques have often performed poorly overall and are unable to successfully provide appropriate feature subsets to decrease system training, runtimes, and overhead. In order to prevent local optimum, it can first enrich the population by taking into account DNA-encoded traffic characteristics. Second, it may use nondominated sorting and decoding to optimize feature subsets and take into account imbalanced cases in detection. Finally, the system performance may be enhanced by the implemented hyperparameter optimisation model. Experiments demonstrate that ADDC can efficiently select the best subset from all datasets, enhance attack detection's overall performance, and optimize the imbalance problem in both binary and multiclass classification. When compared to previous optimum techniques, the total detection rate for some datasets increased by more than 10%, and the balance increased by more than 50%. Even if there are two forms of assaults that cannot be detected in the UNSW-NB15 dataset, ADDC can still resolve the imbalance problem in scenarios when there are few incidences, proving that it is not a perfect solution. As a result, there are still significant risks to the Internet of Vehicles' security. Thus, future effort will concentrate on creating a comprehensive plan for attack detection systems.

## REFERENCES

1. Namasudra S., An improved attribute-based encryption technique towards the data security in cloud computing, in: Concurrency and Computation: Practice and Exercise, 2017, http://dx.doi.org/10.1002/cpe.4364.
2. Namasudra S., P. Roy, B. Balamurugan, Cloud computing: fundamentals and research issues, in: Proceedings of the 2nd International Conference on Recent Trends and Challenges in Computational Models, IEEE, Tindivanam, India, 2017.
3. Namasudra S., Taxonomy of DNA-based security models, in: S. Namasudra, G.C. Deka (Eds.), Advances of DNA Computing in Cryptography, Taylor & Francis, 2018, pp. 53–68.
4. Devi D., S.K. Biswas, B. Purakayastha, Redundancy-driven modified Tomek-link based undersampling: a solution to class imbalance, Pattern Recognit. Lett., 93, 3-122017.
5. Deka G.C., M. Kathing, D.P. Kumar, Library automation in cloud, in: Proceedings of the International Conference on Computational Intelligence and Communication Networks, Mathura, 2013.
6. Namasudra S., G.C. Deka, Introduction of DNA computing in cryptography, in: S. Namasudra, G.C. Deka (Eds.), Advances of DNA Computing in Cryptography, Taylor & Francis, 2018, pp. 27–34.
7. Deka G.C., M.D. Borah, Cost benefit analysis of cloud computing in education, in: Proceedings of the International Conference on Computing, Communication and Applications, 2012, pp. 1–6.
8. Ferraiolo D.F., D.R. Kuhn, Role-based access controls, in: Proceedings of the 15th National Computer Security Conference, Baltimore, USA, 1992, pp. 554–563.
9. Bethencourt J., A. Sahai, B. Waters, Ciphertext-policy attribute based encryption, in: Proceedings of the IEEE Symposium on Security and Privacy, Berkeley, CA, 2007, pp. 321–334.
10. Tiri K., Side-channel attack pitfalls, in: Proceedings of the 44th ACM/IEEE Design Automation Conference, IEEE, San Diego, USA, 2007.
11. Aashiq BS, Amirtharajan R (2020) A robust medical image encryption in dual domain: chaos-DNA-IWT combined approach. Med Biol Eng Comput 58:1445–1458. https://doi.org/10.1007/s11517-020-02178-w
12. Liu J, Ma Y, Li S, Lian J, Zhang X (2018) A new simple chaotic system and its application in medical image encryption. Multimed Tools Appl 77:22787–22808. https://doi.org/10.1007/s11042-017-5534-8
13. Dzwonkowski M, Rykaczewski R (2019) Secure quaternion feistel cipher for DICOM images. IEEE Trans Image Process 28:371–380. https://doi.org/10.1109/TIP.2018.2868388
14. Priya S, Santhi B (2019) A novel visual medical image encryption for secure transmission of authenticated watermarked medical images. Mobile Netw Appl. https://doi.org/10.1007/s11036-019-01213-x
15. Aqeel-ur-Rehman LX, Hahsmi MA, Haider R (2018) An efficient mixed inter-intra pixels substitution at 2bits-level for image encryption technique using DNA and chaos. Optik 153:117–134. https://doi.org/10.1016/j.ijleo.2017.09.099
16. Kumar S, Panna B, Jha RK (2019) Medical image encryption using fractional discrete cosine transform with chaotic function. Med Biol Eng Comput 57:2517–2533. https://doi.org/10.1007/s11517-019-02037-3
17. Ravichandran D, Praveenkumar P, Rayappan JBB, Amirtharajan R (2017) DNA chaos blend to secure medical privacy. IEEE Trans Nanobiosci 16:850–858. https://doi.org/10.1109/TNB.2017.2780881
18. Ravichandran D, Rajagopalan S, Upadhyay HN, Rayappan JBB, Amirtharajan R (2019) Encrypted biography of biomedical image - a pentalayer cryptosystem on FPGA. J Signal Process Syst 91:475–501. https://doi.org/10.1007/s11265-018-1337-z

19. Chai X, Gan Z, Yuan K, Chen Y, Liu X (2019) A novel image encryption scheme based on DNA sequence operations and chaotic systems. Neural Comput & Applic 31:219–237. https://doi.org/10.1007/s00521-017-2993-9
20. Belazi A, Talha M, Kharbech S, Xiang W (2019) Novel medical image encryption scheme based on chaos and DNA encoding. IEEE Access 7:36667–36681. https://doi.org/10.1109/ACCESS.2019.2906292
21. Almaiah Mohammed Amin. 2021. Classification of Cyber Security Threats on Mobile Devices and Applications. Artificial Intelligence and Blockchain for Future Cybersecurity Applications. Springer, Cham, 107–123.
22. Jianping X., Chun L., Jing Z., et al. 2021. A survey on network intrusion detection based on deep learning. Frontiers of Data and Computing 3, 3 (2021), 59–74
23. Soon Hui Fern, Amiza Amir, and Saidatul Norlyana Azemi. 2022. Multi-class imbalanced classification problems in network attack detections. In Proceedings of the 6th International Conference on Electrical, Control and Computer Engineering. Springer, Singapore, 1057--1069.
24. Almseidin Mohammad, Al-Sawwa Jamil, and Alkasassbeh Mouhammd. 2022. Generating a benchmark cyber multi-step attacks dataset for intrusion detection. Journal of Intelligent & Fuzzy Systems. Preprint, 1–15.
25. Azayeri N. and Sajedi H.. 2020. DNAVS: An algorithm based on DNA-computing and vortex search algorithm for task scheduling problem. Evolutionary Intelligence 14, 4 (2020), 1763–1773.

*******