T.C.
DÜZCE ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

BİYOTEKNİK DONANIMLI SİMETRİK BİR DNA ŞİFRELEME METODU

OĞUZHAN KENDİRLİ

DOKTORA TEZİ
ELEKTRİK – ELEKTRONİK ve BİLGİSAYAR MÜHENDİSLİĞİ
ANABİLİM DALI

DANIŞMAN
DR. ÖĞR. ÜYESİ ESRA ŞATIR

DÜZCE, 2022

T.C.
DÜZCE ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

BİYOTEKNİK DONANIMLI SİMETRİK BİR DNA ŞİFRELEME METODU

Oğuzhan KENDİRLİ tarafından hazırlanan tez çalışması aşağıdaki jüri tarafından Düzce Üniversitesi Lisansüstü Eğitim Enstitüsü Elektrik - Elektronik ve Bilgisayar Mühendisliği Anabilim Dalı'nda **DOKTORA TEZİ** olarak kabul edilmiştir.

Tez Danışmanı

Dr. Öğr. Üyesi Esra ŞATIR

Düzce Üniversitesi

Jüri Üyeleri

Dr. Öğr. Üyesi Esra ŞATIR
Düzce Üniversitesi _____

Prof. Dr. İnan GÜLER
Gazi Üniversitesi _____

Prof. Dr. Resul KARA
Düzce Üniversitesi _____

Dr. Öğr. Üyesi İrem DÜZDAR ARGUN
Düzce Üniversitesi _____

Dr. Öğr. Üyesi Engin EŞME
Konya Teknik Üniversitesi _____

Tez Savunma Tarihi: 16/06/2022

BEYAN

Bu tez çalışmasının kendi çalışmam olduğunu, tezin planlanmasından yazımına kadar bütün aşamalarda etik dışı davranışımın olmadığını, bu tezdeki bütün bilgileri akademik ve etik kurallar içinde elde ettiğimi, bu tez çalışmasıyla elde edilmeyen bütün bilgi ve yorumlara kaynak gösterdiğimi ve bu kaynakları da kaynaklar listesine aldığımı, yine bu tezin çalışılması ve yazımı sırasında patent ve telif haklarını ihlal edici bir davranışımın olmadığını beyan ederim.

16 Haziran 2022

Oğuzhan KENDİRLİ

TEŞEKKÜR

Doktora öğrenimimde ve bu tezin hazırlanmasında gösterdiği her türlü destek ve yardımdan dolayı danışmanım Dr. Öğr. Üyesi Esra ŞATIR'a en içten dileklerimle teşekkür ederim.

Tez çalışmam boyunca değerli katkılarını esirgemeyen Prof. Dr. İnan GÜLER'e ve Prof. Dr. Resul KARA'ya da şükranlarımı sunarım. Ayrıca desteklerinden ötürü Dr. Öğr. Üyesi İrem DÜZDAR ARGUN'a ve Dr. Öğr. Üyesi Engin EŞME'ye de teşekkür ederim.

Bu çalışma boyunca yardımlarını ve desteklerini esirgemeyen sevgili aileme ve çalışma arkadaşlarıma sonsuz teşekkürlerimi sunarım.

Bu tez çalışması, Düzce Üniversitesi BAP- 2016.06.01.502 numaralı Bilimsel Araştırma Projesiyle desteklenmiştir

16 Haziran 2022 **Oğuzhan KENDİRLİ**

İÇİNDEKİLER

Sayfa No

ŞEKİL LİSTESİ

Sayfa No

ÇİZELGE LİSTESİ

Sayfa No

KISALTMALAR

A	Adenin
AES	Advanced Encryption Standard
ASCII	American Standard Code for Information Interchange
bp	Base Pair
CDMB	Central Dogma of Molecular Biology
C	Sitozin
CPU	Central Processing Unit
DES	Data Encryption Standart
D-GET	DNA-Genetic Encryption Technique
DNA	Deoksiribonükleik Asit
EB	Eksabayt
F	Function
G	Guanin
GA	Genetic Algorithm
GHz	Gigahertz
mRNA	Messenger Ribonucleic Acid
MSE	Mean Square Error
NP	Non-Deterministic Polynomial
NPCR	Number of Changing Pixel Rate
PSNR	Peak Signal to Noise Ratio
RAM	Random Access Memory
RC4	Rivest Cipher 4
RSA	Rivest–Shamir–Adleman
S_Box	Substitution Box
SSL	Secure Sockets Layer
T	Timin
U	Urasil
UACI	Unified Averaged Changed Intensity
XOR	Exclusive OR
ZB	Zetabayt
3DES	Triple Data Encryption Standart

ÖZET

BİYOTEKNİK DONANIMLI SİMETRİK BİR DNA ŞİFRELEME METODU

Oğuzhan KENDİRLİ
Düzce Üniversitesi
Lisansüstü Eğitim Enstitüsü,
Elektrik – Elektronik ve Bilgisayar Mühendisliği Anabilim Dalı
Doktora Tezi
Danışman: Dr. Öğr. Üyesi Esra ŞATIR
Haziran 2022, 50 sayfa

İnternet/ağ teknolojilerinin her geçen gün artan hızı, dünyadaki veri oluşumunu önemli ölçüde artırmaktadır. Ağ üzerinde bilgi akışı arttıkça, kullanıcılar için zaman güvenliği tehditleri de artmaktadır. Verileri korumak için geçmişten günümüze kriptografi ve steganografi teknikleri kullanılmıştır. Kriptografinin amacı, mesajı gönderen ile alıcı arasında gözlemcinin anlayamayacağı bir şekilde iletmektir. Günümüzde DNA kriptografisi ise kriptografi alanında parlayan bir daldır. Buradaki birincil amaç, DNA'yı taşıyıcı ve modern biyolojik teknikleri uygulama araçları olarak kullanmaktır. Bu çalışmada, DNA şifreleme ve DNA operatörleri Feistel ağ yapısına entegre edilerek bir DNA kriptografi tekniği önerilmiştir. Burada taşıyıcı olarak görüntü, metin veya video gibi geleneksel dijital medya yerine DNA'nın kendisi kullanılmış, biyolojik teknikleri ise uygulama araçları olarak kullanılmıştır. Ayrıca geliştirilen simülasyon yazılımı ve sentezlenen DNA dizisi, özel olarak oluşturulmuş biyoteknik donanıma dijital ve biyolojik olarak entegre edilmiştir. Deneysel sonuçlar, önerilen çalışmanın kriptografik gereksinimler için verimli çıktılara sahip olduğunu %100'e yakın kapasite, tek blok için yaklaşık 12×10^6 yıl kaba kuvvet saldırı, tek blok için 2^{80} anahtar uzayı ve 2'ye yakın entropi analizi sonuçları ile göstermiştir. Ayrıca önerilen yöntemin uygulanması, laboratuvar deneyleri ile de doğrulanmıştır.

Anahtar sözcükler: DNA hesaplama, DNA kodlama, DNA şifreleme, Metin şifreleme, Güvenlik analizi.

ABSTRACT

A SYMMETRIC DNA ENCRYPTION PROCESS WITH A BIOTECHNICAL HARDWARE

Oğuzhan KENDİRLİ
Düzce University
Institute of Graduate Studies,
Department of Electrical – Electronics and Computer Engineering
Doctoral Thesis
Supervisor: Assoc. Prof. Dr. Esra ŞATIR
June 2022, 50 pages

The growing rate of internet/network technologies day by day dramatically increases the formation of data in the world. As the flow of information increases on the network, time security threats are also increasing for users. In order to protect data, cryptography and steganography have been used from the past to the present. The goal of cryptography is to transfer the message between sender and receiver in a way that is incomprehensible to the observer. Nowadays, DNA cryptography is a shining branch in the field of cryptography. The primary purpose here is to employ DNA as a carrier and to employ modern biological techniques as application tools. In this study, a DNA cryptography technique was proposed by integrating DNA encoding and DNA operators into the Feistel network structure. Here, DNA itself was used as a carrier instead of traditional digital media such as image, text or video, while its biological tools were being used as implementation tools. Besides, the developed simulation software and the synthesized DNA sequence were digitally and biologically integrated into specifically created biotechnical hardware. Experimental results demonstrated that the proposed study has efficient outcomes for cryptographic requirements; capacity of nearly 100%, brute force attack nearly 12×10^6 years for only one block, key space that is 2^{80} for only one block, and entropy analyses close to 2. Besides, the implementation of the proposed method has been verified by vitro experiments.

Keywords: DNA computing, DNA encoding, DNA encryption, Text encryption, Security analysis.

EXTENDED ABSTRACT

A SYMMETRIC DNA ENCRYPTION PROCESS WITH A BIOTECHNICAL HARDWARE

Oğuzhan KENDİRLİ
Düzce University
Institute of Graduate Studies,
Department of Electrical – Electronics and Computer Engineering
Doctoral Thesis
Supervisor: Assoc. Prof. Dr. Esra ŞATIR
June 2022, 50 pages

1. INTRODUCTION

The growing rate of internet/network technologies day by day dramatically increases the formation of data in the world. As the flow of information increases on the network, time security threats are also increasing for users. In order to protect data, cryptography and steganography have been used from the past to the present. The goal of cryptography is to transfer the message between sender and receiver in a way that is incomprehensible to the observer. Nowadays, Deoksiribonükleik Asit (DNA) cryptography is a shining branch in the field of cryptography. The primary purpose here is to employ DNA as a carrier and to employ modern biological techniques as application tools. In this study, a DNA cryptography technique was proposed by integrating DNA encoding and DNA operators into the Feistel network structure. Here, DNA itself was used as a carrier instead of traditional digital media such as image, text or video, while its biological tools were being used as implementation tools.

2. MATERIAL AND METHODS

In this part of the thesis, a DNA cryptography technique is proposed by integrating DNA coding and DNA operators into the Feistel network structure. Here, DNA itself was used as the carrier instead of traditional digital media such as images, text or video, while modern biological tools were used as tools of application. In addition, the simulation software developed and the synthesized DNA sequence were integrated into the specially created biotechnical hardware both digitally and biologically. The amount of data (bits) that can be embedded in a DNA base (A, C, G and T) becomes a major issue here, as data is also intended to be stored in DNA media. When the studies in the literature are examined, it has been seen that the number of bits that can be encoded per nucleotide

does not exceed two. In the proposed study, this rate is increased with the original compression algorithm performed before the encryption process. When the biologically synthesized DNA sequence was compared with the simulated DNA sequence obtained by the simulation decoding process, 100% match success was achieved. The simulation software is also integrated into the hardware that can work as plug and play.

3. RESULTS AND DISCUSSIONS

In this section, biological experiments, performance and safety analyzes and results are presented by giving concrete examples as much as possible. In the proposed study, the coded DNA sequence was synthesized from 9588 base pairs (bp) obtained through simulation software with the support of Düzce University Scientific Research Projects Coordinatorship. After sequencing of 9588 bp, it was observed that the synthesized and sequenced DNA chains matched without loss of information. The synthesized 9588 bp DNA strand was encapsulated in a plasmid.

In the proposed study, it is clear that most of the prescribed requirements are met. Since the proposed work converts plain text to binary, it is possible to encode the entire character set over DNA. However, dynamic coding is not possible. Because DNA coding and Feistel Network-integrated DNA operators are used and different round keys are generated in only one session per cycle, a unique sequence is in place to encode each character of the plaintext into the DNA sequence. These are the main features that support the robustness and dynamism of encryption.

The proposed method has a more efficient throughput ratio than existing studies in the literature. Capacity values are also mentioned in the graph. In this study, the key length is at least 80 bits, since the input has at least one block with a length of 12 bits. The key is obtained from 40 cycles, each containing a 2-bit round key. It is almost impossible to crack it with a brute force attack using today's technology. Since there are 40 cycles in the proposed scheme, the total length of the key is $40\times2=80$ bits. Here we have 2^{80} possible combinations to extract the secret key in the proposed scheme. Also, this number of combinations depends on the length of the plain text. The longer the plaintext, the longer the key. It is seen from that the estimated entropy values are close to 2. Considering that there are 4 base possibilities (A, C, G, T), this is a very efficient ratio.

4. CONCLUSION AND OUTLOOK

In this thesis, a new encryption approach is derived by combining DNA carrier medium, DNA coding and DNA Exclusive OR (DNA XOR) process with Feistel network structure, depending on the principle of central molecular biology for encoding and decoding systems. In addition, the proposed DNA coding process is both digitally and biologically integrated into the designed biotechnical hardware.

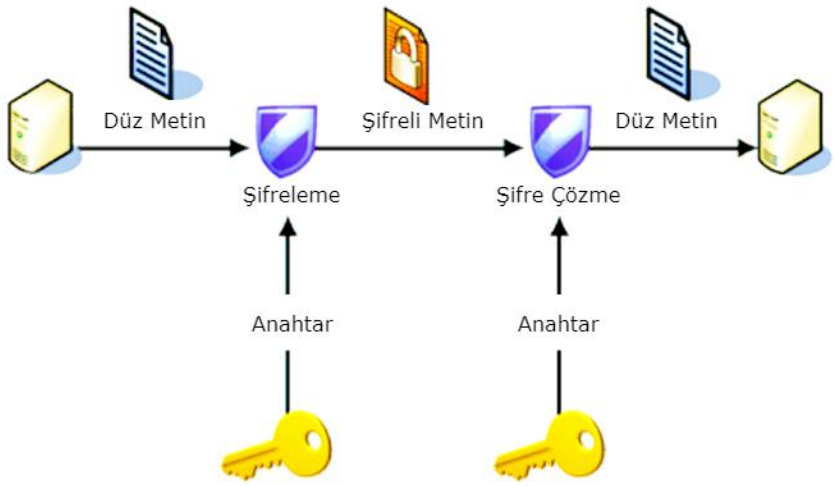
Experiments have shown that DNA cryptography has a high potential to be a new method in data security. Experimental results showed that the proposed work has productive results in terms of capacity, brute force attack, key space and entropy analysis. In addition, the application of the proposed method has been confirmed by laboratory experiments.

1. GİRİŞ

İnternet ve ağ teknolojilerinin günden güne büyümesi ve ağ üzerindeki bilgi akışından ötürü, kullanıcılar için güvenlik tehditleri artmaktadır. Kritik bilgilerini elde etmek veya veri bütünlüğünü bozmak için sisteme sızmaya çalışan birçok tehdit mevcuttur. Bu nedenle, bilgi güvenliği, modern hesaplama sistemlerinin vazgeçilmez gereksinimi olmuştur. Gizli bilgilerini sızdırma gibi bir lüksü olmayan kamu birimleri, bankacılık vb. bazı sektörler vardır. Geçmişten günümüze, veriyi korumak için kriptoloji ve steganografi sıklıkla kullanılan yöntemlerdir. Kriptoloji, veriyi kodlarken steganografi veriyi saklamaktadır. Kriptografide, şifreleme ve şifre çözme işlemleri anahtar yardımıyla gerçekleştirilmektedir [1].

Kriptografi, güvenlik sorunları için en dikkat çekici çözümlerden biridir. Kriptografide veri, gönderici ve alıcı arasında güvenilmeyen bir ortam üzerinden aktarılır. Kriptografik algoritmalar, şifreleme ve şifre çözme sırasında anahtarların kullanımına bağlı olarak sınıflandırılır. Bu kategoriler simetrik kriptografi ve asimetrik kriptografidir. İlkinde, aynı anahtar kullanılarak şifreleme ve şifre çözme işlemleri gerçekleştirilmektedir. İkincisinde ise farklı anahtarlar kullanılarak şifreleme ve şifre çözme işlemleri gerçekleştirilmektedir. Burada her bir taraf, özel anahtar ve genel anahtar olarak adlandırılan bir çift anahtara sahiptir. Özel anahtar her zaman gizli tutulurken genel anahtar paylaşılabilir [2].

Kriptografide amaç, mesajın, gözlemcinin anlayamayacağı bir şekilde alıcı ve gönderici arasında transfer edilmesidir. Temel terminolojiye kısaca değinilecek olursa, düz metin, doğal bir dile uygun, sonlu bir alfabeden alınan karakter sırasıdır. Şifreleme, bilinen bir algoritma ve gizli bir anahtar ile düz metni karmaşıklaştırma işlemidir. Çıktı ise, şifreli metin olarak bilinen karakter dizisidir. Şifre çözme, şifreli mesajı orijinal haline dönüştüren tersine bir işlemdir. Şifrelemenin amacı, gizli anahtarı bilmeyen bir gözlemcinin, şifreyi çözmesini engellemektir [3]. Şekil 1.1'de bu süreç gösterilmektedir [4].



Şekil 1.1. Şifreleme ve şifre çözme süreci.

Günümüzde kriptografi alanında yeni gelişmekte olan ve DNA kriptografisi denilen bir dal mevcuttur. Bu metodun temel amacı, düz metni şifreleyerek DNA içinde gizlemektir. Ayrıca, DNA kriptografisinde diğer modern yöntemlerin aksine çok büyük miktarlarda veri için anahtar üretilebilmektedir [1]. DNA, çok küçük bir hacimde oldukça büyük miktarlarda veri depolayabilmesinden ötürü de oldukça dikkat çekici bir ortamdır. Klasik elektriksel ve optik ortamların kapasitesi, DNA ile aşılmıştır. 1 gram DNA, 10^{21} baz ya da 10^{8} terabayt civarında bilgi içerebilmektedir. Bu nedenle, bir kaç gram DNA, tüm dünyadaki bilgiyi depolama potansiyeline sahip olabilir [3]. Çizelge 1.1'de görüldüğü üzere veri depolamada hard disk, USB bellek ve bakteriyel DNA için okuma yazma, veri koruma, harcanan güç ve veri yoğunluğu karşılaştırılmasında DNA açık ara farkla klasik depolama tekniklerinden üstündür. Bu çizelge göstermektedir ki; dünyada ki mevcut tüm veriyi depolamak için gereken yaklaşık 1 kg DNA'dır [5]. Bilgiyi kodlamak ve kodunu çözmek için matematiksel birçok teknik ve sistem geliştirilmiştir. Ancak bu sistemler, DNA kriptografi teknikleri ile daha ileri götürülebilir [6].

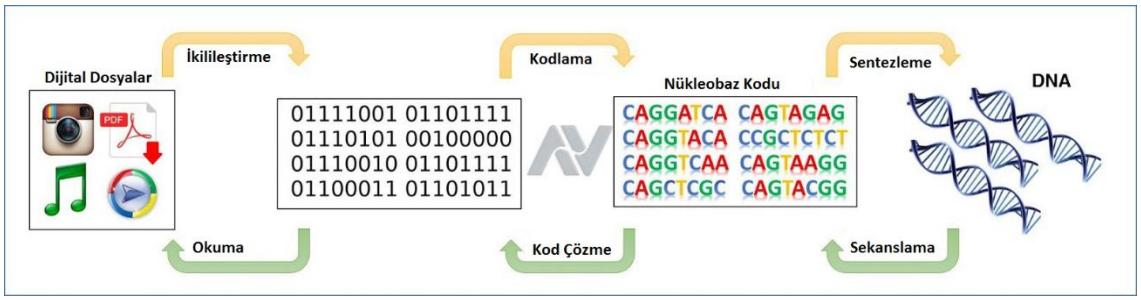
Çizelge 1.1. Hard disk, USB bellek ve bakteriyel DNA için okuma/yazma hızı, veri saklama, harcanan güç ve veri yoğunluğu karşılaştırılması.

Depolama Limitleri			
	Hard disk	USB bellek	Bakteriyel DNA
Okuma/yazma hızı (bit başına μs)	~ 3.000-5.000	~100	<100
Veri saklama (yıl)	>10	>10	>100
Harcanan güç (Gigabayt başına Watt)	~ 0.04	~0.01-0.04	$<10^{-10}$
Veri yoğunluğu (cm^3 başına bit)	~ 10^{13}	~10^{16}	~10^{19}

DNA kriptografisi, DNA'nın taşıyıcı olarak kullanıldığı, modern biyolojik tekniklerin ise uygulama aracı olarak kullanıldığı yeni bir daldır [7]. DNA kriptografisinde DNA, metin, resim, video vb. herhangi bir ortam yerine bir bilgi taşıyıcısıdır. Bu nedenle, bu yaklaşım şifreleme elde etmek için biyolojik teknolojiden yararlanır. Başka bir deyişle, taşıyıcı ortam olarak DNA kullanılırken, uygulama araçları olarak modern biyolojik teknikler kullanılmaktadır. Ancak pahalı deneysel ekipman, karmaşık işlemler ve karmaşık biyoteknoloji gibi dezavantajları vardır. Bu nedenle, kriptografi alanında hala yaygın olarak uygulanamamaktadır. Bu sorunların üstesinden gelmek amacıyla, bilgiyi karıştırmak için bazı DNA hesaplama işlemleri kullanılır [8].

Günümüzde DNA kavramı karmaşık bir yapıya sahip olduğu için bilgi güvenliği alanında popülerdir. Burada 0 ve 1 gibi geleneksel bilgisayar hesaplama kavramlarından farklı olarak, veriler DNA bazları, Adenin (A), Sitozin (C), Guanin (G) ve Timin (T) kullanılarak şifrelenir ve saklanır. Böylece gönderici, verilerin güvenliğini artırarak şifreleme işlemi sırasında bu bazların herhangi bir kombinasyonunu seçebilir [2].

Verileri kodlamak ya da kodunu çözmek için birçok matematiksel teknik ve sistem geliştirilmiştir. DNA kriptografi teknikleri sayesinde bu sistemler daha ileri götürülebilir [6]. Aşağıda verilen Şekil 1.2'de genel DNA şifreleme ve şifre çözme işlemini özetleyebiliriz [9].



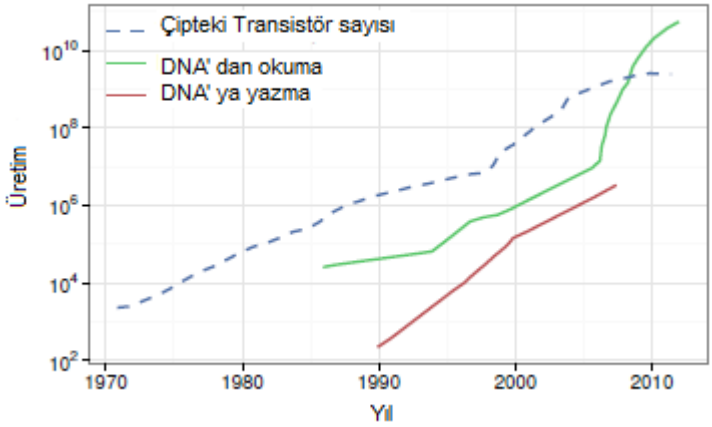
Şekil 1.2. DNA şifreleme ve şifre çözme işlemi süreci.

Bu çalışmanın amacı, kriptografi alanında verimli güvenlik alternatifleri üretmek için doğal biyolojik modellerden faydalı metaforlar oluşturmaktır. DNA'yı taşıyıcı olarak kullanırken modern biyoloji tekniklerini uygulama araçları olarak kullanmak hedeflenmiştir. Bu çalışmanın farkı ise DNA şifreleme ve DNA operatörlerinin Feistel ağ yapısına entegre edilerek yeni bir DNA kriptografi tekniği önerilmesidir. Burada taşıyıcı olarak görüntü, metin veya video gibi geleneksel dijital medya yerine DNA'nın kendisi kullanılırken, biyolojik araçlar ise uygulama araçları olarak kullanılmıştır. Ayrıca geliştirilen simülasyon yazılımı ve sentezlenen DNA dizisi, özel olarak oluşturulmuş biyoteknik donanıma dijital ve biyolojik olarak entegre edilmiştir. Böylece kriptografide geleneksel çoklu ortam taşıyıcıları (resim, metin, video vb.) yerine taşıyıcı ortam olarak DNA'yı kullanarak gizli bit sayısı (kapasitesini) artırılmıştır.

1.1. DNA HESAPLAMA VE ŞİFRELEME ÇALIŞMALARI ÜZERİNE

Çağımızın vazgeçilmez unsuru ve aynı zamanda önemli bir problemi olan büyük veri depolama için DNA'nın oldukça uygun bir zemin teşkil ettiği, hatta hayat devam ettiği müddetçe veri depolama amacıyla DNA moleküllerinin kullanılmasının son derece mantıklı olduğu aşikârdır. Ayrıca DNA'nın oldukça verimli bir kapasite yoğunluk dengesi, uzun ömürlülüğü, zor koşullarda bile dayanıklı olması yanında çevre dostu ve yaşamın temeli olması tercih edilme sebebini güçlendirici unsurlardır. Ancak günümüz koşullarında DNA tabanlı veri depolama teknolojisi halen mükemmel değildir. Sentezleme ve dizileme maliyeti, nükleotid başında düşen hata payı ve veriye DNA üzerindeki erişim zamanı gibi zorluklar mevcuttur. Ancak Moore yasası seyri ile DNA sentez ve dizileme teknolojilerindeki gelişim karşılaştırıldığında, durumun oldukça ümit verici olduğu da Şekil 1.3' te görülmektedir. Tek bir çip üzerine yerleştirilebilen transistörlerin sayısıyla ölçülen bilgisayarların hızının her yıl iki katına çıkacağını belirten Moore yasasının sarsılmaz doğruluğu 40 yıldır kabul edilirken artık yeni teknolojilerin

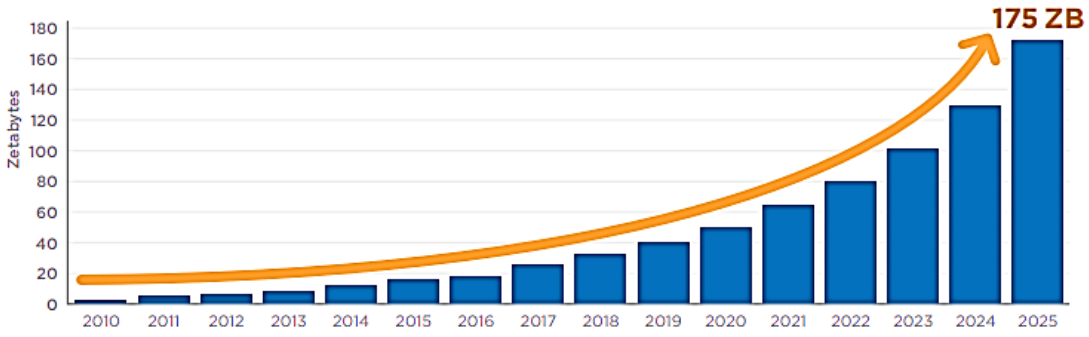
çıkması gerekliliği de ortaya konulmuştur [10]. Günümüzde DNA sentezleme ve dizileme, veri depolama bakımından uygulanabilir olmasa da, tarihsel gelişimi üssel bir şekilde gerçekleşmiştir [11]. Moore yasasına karşı, maliyet azalması ve verimlilik bakımından bu gelişim verilen grafikte de gösterilmektedir. Dizileme teknolojisindeki son gelişmelerin Moore yasasını gölgede bıraktığı, dizileme üretiminin Moore yasasından daha hızlı büyüdüğü Şekil 1.3'te görülmektedir [11].



Şekil 1.3. Carlson eğrileri: DNA sentez ve dizileme teknolojilerinin Moore yasası ile karşılaştırılması.

Günümüzde depolama aygıtları olan hard diskler, optik sürücüler, DVD ve hatta Blu-ray bile hem çok kolay tahribata uğrayıp içerisindeki veriler kaybolabilmekte hem de ömürleri de kısa olmaktadır. Özellikle akıllı telefonların yaygınlaşması ve bulut sisteminde giderilmesi ümit edilen güvenilirlik ve gizlilik gibi konuların yanı sıra hızlı veri üretimi ve yoğunluğu düşünüldüğünde bu teknolojiye olan ihtiyaç daha fazla ortaya çıkmaktadır.

Dijital dünya yani internetteki bütün verinin 2025 yılında 175 zetabaytı (ZB) aşacağı Şekil 1.4' te öngörülmektedir [12]. 2022 itibarı ile 80 zetabayt veri üretmiş olabileceğimizi düşünürsek dijital bilgi birikmeye devam ederken, yüksek yoğunluklu ve uzun süreli depolama çözüm olarak DNA tabanlı veri saklamak birçok avantajı ile ileri teknolojiyi oluşturmaktadır.



Şekil 1.4. Global veri kümesinin yıllık boyutu.

Veri saklamanın en temel gereği, saklanan verinin bozulmadan alabildiğince uzun yıllar saklanabilmesi ve bu verinin istenildiğinde erişilebilir olmasıdır. USB diskte saklanan bir veri yaklaşık 5 yıl, manyetik bantta saklanan bir veri yaklaşık 15-30 yıl, DNA'da saklanan bir veri ise yüzyıllar boyu korunabilir [13]. Bu açıdan bakıldığında da DNA tabanlı veri saklama, veri saklamanın en temel özelliğini fazlasıyla karşılamaktadır.

Sosyal medyanın ve nesnelerin internetinin yaygınlaşmasıyla birlikte internette paylaşılan dijital veri 600 eksabayta (EB) ulaşırken dijital veri depolama teknolojilerinden sabit disk sürücüleri veya USB belleklerle ihtiyacımızı karşılamak imkansız hale gelmektedir. Ayrıca sabit disk sürücüleri ve USB bellekler fazlasıyla elektrik tüketmektedir. Eğer internetteki bütün dijital veriyi depolamayı sağlayacak depolama teknolojileri olsa bile dünya olarak global elektrik üretimi bunları çalıştırmaya yetmeyecektir. Ayrıca, uzun süreli depolamaya elverişli olmayan DVD'ler ve manyetik kasetler birkaç yıl içinde bozulmaktadırlar. Bu yüzden DNA tabanlı veri depolama teknolojileri geleneksel teknolojilere göre çevre dostu depolanabilir teknoloji olma özelliği taşımaktadır.

Carlson Eğrisi, Moore Yasası'nın biyoteknolojide ki eşdeğeri olarak gösterilmektedir. Carlson, DNA dizileme hızının en az Moore Yasası'ndaki kadar hızlı olacağını, yani her 18 ayda bir DNA dizilemenin iki katına çıkacağını söylemektedir. Bu da DNA tabanlı veri saklama teknolojilerinin gelecekte çok daha yaygın olarak kullanılacağının bir kanıtıdır.

DNA geleneksel dijital bilgisayar ile karşılaştırıldığında bit başına depolama alanı ciddi bir azalma sağlar. Teorik olarak DNA nükleotid başına 2 bit veya DNA gramında 455 eksabayta kodlanarak veriyi A, C, G ve T bazları kullanılarak volümetrik bir şekilde saklar. DNA yoğunluk, sağlamlık, stabilite ve enerji verimliliği avantaj özellikleri ile depolama amacıyla kullanmak kaçınılmazdır. Ayrıca birçok araştırma sonunda DNA

tabanlı veri saklamanın bir şifreleme mekanizması olmadan bile çok daha özel ve güvenli dijital depolama sunabileceği öngörülmüştür [14].

Bilgi taşıyıcıları olarak mikroskobik polimerleri (DNA nükleotidlerini) kullanma fikri yeni değildir. Bu konuda ilk çalışmalar 1964 ve 65 yıllarında Neiman tarafından gerçekleştirilmiştir. Bu çalışmalar ışığında kriptolojik amaçlar ile DNA hesaplama yapan algoritmalar 90'lı yıllarda önerilmeye başlanmıştır [15].

DNA hesaplama kavramı ise ilk olarak 1994 yılında Adleman tarafından Hamilton yol probleminin çözülmesi amacıyla ele alınmıştır. 1995 yılında Lipton, başka bir Non-Deterministic Polynomial - Deterministik Olmayan Polinom (NP) karmaşıklık sınıfına giren, 2 bitlik sayılar için çizelgeleme problemini, bir test tüpünün içinde DNA moleküllerini kullanarak çözmüştür. 1996 yılında ise Dan Boneh ve arkadaşları Adleman ve Lipton tarafından kullanılan DNA hesaplama yaklaşımlarını simetrik şifreleme gerçekleştiren Data Encryption Standard – Veri Şifreleme Standardı (DES) algoritmasını kırmak için uygulamışlardır. DNA dizileri üzerinde, ikilik tabandaki dizgileri kodlama etkisi olan özütleme, polimerizasyon, amplifikasyon gibi temel biyolojik işlemleri de uygulamışlardır [1]. 2003 yılında Chen tarafından moleküler hesaplama ve tek kullanımlık şifrelemeye dayalı DNA şifreleme algoritmaları tasarlamışlardır. Şifreleme/şifre çözme adımları 2 boyutlu görüntüler üzerinde gerçekleştirilmiştir [16].

2005 yılında Tanaka ve arkadaşları asimetrik şifrelemeye dayalı bir DNA şifreleme algoritması önermiştir. Çalışmada, özdeş olmayan karışımlara yardımcı olarak ortak anahtarlar sentezlenmiştir. Anahtar üretildikten sonra, mesaj ilk açık anahtarla kodlanmış ve ikinci ortak anahtarla bir DNA dizisine bağlanmıştır. Asimetrik şifreleme simetrik şifrelemeden daha güvenli ancak hız açısından simetrik şifreleme ile karşılaştırıldığında simetrik şifrelemeden daha yavaş kalmıştır [17].

2006 yılında, Amin ve arkadaşları Tarafından [18] simetrik şifrelemeye dayalı bir DNA kriptografisi önerilmiştir. Bu metotta, tek bir American Standard Code for Information Interchange (ASCII) karakter 4 nükleotid ile temsil edilerek, ilgili DNA zinciri içinde aranmakta ve çıktı olarak ilgili karakterin zincir içindeki konumunu veren bir işaretçi üretilmiştir.

2008 yılında, Verma ve arkadaşları, Mobil Ad hoc ağlarında güvenli aktarım amacıyla sözde DNA kriptografi metodunu kullanmışlardır. Simetrik şifrelemenin kullanıldığı bu yöntem, kablosuz mobil ağlarla sınırlı kalmıştır. Ayrıca, burada maliyet önemli bir

sorundur çünkü uygulamaya geçildiği takdirde, protein sentezi, DNA sentezinden çok daha maliyetli olduğu görülmüştür [19].

2011 yılında Kumar ve Singh, DNA dizilerinde gizli verilerin yazılmasına dayanan yeni bir yöntem önermişlerdir. Önerilen algoritmada, "HELLO" kelimesini düz metin olarak 350 bitlik tek kullanımlık bir anahtarla şifrelemişlerdir. İlk olarak "HELLO" kelimesi ASCII kod eşdeğerine dönüştürülmüştür. Daha sonra 35 bit ve 35x10=350 bit Tek Zamanlı Pad dizisinde ikili örüntü eşdeğeri üretilmiştir. Üretilen bu nükleotid dizisini ve single-stranded DNA (ssDNA) anahtarını kullanarak, şifreleme/şifre çözme yoluyla verileri mükemmel bir şekilde gönderip almışlardır [13].

Yine 2012 yılında Zhang ve arkadaşları Tarafından [20] önerilen başka bir çalışmada da düz metin, uzun bir DNA zincirine dönüştürülmüş ve bu zincir rastgele parçalanarak anahtar üretmek suretiyle önerilen algoritma ile kodlanarak şifreli metin üretilmiştir.

2013 yılında, Tornea ve Monica [21] tarafından önerilen DNA şifreleme algoritmasında, hali hazırdaki gen bankaları şifreleme amacıyla kullanılmıştır. Öncelikle düz metin, klasik yolla (A-00, G-01, C-10, T-11) DNA bazlarına çevrilmiştir. Anahtar üretimi ise, düz metnin her bir baytının baz karşılığının, bu gen bankalarındaki eşleniklerine indekslenmesi yoluyla gerçekleştirilmiştir. Bu çalışmada, anahtar uzayı, gen bankalarının kullanımı sayesinde epey genişletilmiştir. Ancak yine DNA zincirinin verimli kullanım sorunsalı göz önünde bulundurulmuştur.

2015 yılında Monika ve Upadhyaya tarafından DNA dizileri ile şifrelenen RSA (Rivest–Shamir–Adleman) şifreleme algoritmasına dayalı bir yöntem önerilmiştir. İlgili yöntem, iletişim sırasında daha yüksek düzeyde güvenlik sağlamak için Secure Socket Layer - Güvenli Giriş Katmanı (SSL) teknolojisine uyarlanmıştır [22].

2016 yılında Mousa, D-GET adlı bir DNA-Genetik Şifreleme Tekniği (D-GET) önermiştir. Burada ki birincil amaç, tekniği daha güvenli ve daha az tahmin edilebilir hale getirmekti. Bu teknikte, herhangi bir dijital verinin ikili biçimi DNA dizilimine dönüştürülür, yeniden şekillendirilir, şifrelenir, çaprazlanır, mutasyona uğratılır ve ardından yeniden şekillendirilir. D-GET'in bu ana adımları üç veya daha fazla kez tekrarlanır. Deneysel sonuçlar, önerdiği tekniğin farklı saldırılara karşı çok katmanlı koruma aşamalarına ve çok aşamalı ve genetik operasyonlara dayalı daha yüksek bir güvenlik düzeyine sahip olduğunu göstermiştir [23]. Aynı yıl Zhang ve arkadaşları tarafından [24] önerilen çalışmada, görüntüler üzerine uygulanan kaotik kodlama, tek

kullanımlık şifreye dayalı DNA hesaplama prensibine ek olarak uygulanmıştır.

2016'da Goyat ve Jain, uygulama ve tasarım için bulut sunucularının tüm iletişimini ve depolanmasını güvence altına almak için DNA tabanlı bir şifreleme algoritması önermişlerdir. DNA tabanlı kriptografik teknikleri, temel olarak ikame ve diğer temel operatör uygulamaları kullanılarak geliştirmişlerdir. Çalışmalarının zaman ve mekan karmaşıklığı açısından deneysel performansı, bulut tabanlı sistemlerde veri güvenliği için etkili ve düşük kaynak tüketen teknikler sağlanmıştır. [25].

2017'de Ahmed ve Muhammed, DNA ve Rivest Cipher 4 (RC4) kullanarak yeni bir hibrit güvenlik algoritması geliştirmişlerdir. Bu algoritma, steganografi çerçevesi kapsamında yüksek karmaşıklık ile güvenli veri gizleme sağlamak için hem simetrik akış şifresi RC4 hem de DNA indeksleme algoritmalarını kullanmışlardır. Önerilen şemanın performans değerlendirmesi şu üç parametre dikkate alınarak ölçülmüştür; koşullu entropi, rastgelesellik testleri ve şifreleme süresi. Sonuçları, yerel RC4 ile karşılaştırıldığında güvenlikte daha iyi performans göstermiş ve hibrit şifrede oldukça karmaşık çıkmıştır [26].

2018'de Thangavel ve Varalakshmi, DNA içindeki orijinal verileri güvence altına almak için bir DNA şifreleme sistemi önermişlerdir. Yöntemlerine Gelişmiş ElGamal şifreleme sistemi adını vermişlerdir. Yöntemleri, bulutta anahtar yönetim sorunlarını çözme amacına sahip asimetrik bir şifreleme sistemidir. Anahtar dosyasını veri sahibi ile kullanıcısı arasında güvenli bir şekilde aktararak bu sorunu aşmayı amaçlamaktadır. Gelişmiş ElGamal şifreleme sisteminin daha iyi kullanıcı kimlik doğrulama sonuçları sağladığından bahsetmişlerdir. Ayrıca güvenlik saldırılarına karşı daha iyi performans gösterdiğini eklemişlerdir [27].

2018 yılında Narendren ve arkadaşları RSA algoritmasına ek bir katman görevi gören önerilen DNA tabanlı algoritma önermişlerdir. Yöntemlerinde, DNA'dan protein sentezi sürecine dayanan bir döngü fonksiyonu kullanılmıştır. Önerdikleri planın güvenlik analizinin iyi sonuçlar verdiğini kanıtladıklarını belirtmişlerdir [28]. Yine 2018'de Sharma ve Sohal, bulut sistemleri için yeni bir şifreleme tekniği sunmuşlardır. Yaklaşımları, verileri buluta yüklemeden önce şifrelemek için istemci tarafı veri şifrelemesini kullanmaktadırlar. Ayrıca, DNA kriptografisine dayanan çok katlı simetrik anahtarlı bir kriptografi tekniği olarak da düşünülebilir. Çalışmalarını mevcut simetrik anahtar algoritmalarıyla karşılaştırdılar; DNA, Advanced Encryption Standard (AES),

Data Encryption Standart (DES) ve Blowfish. Deneysel sonuçlarının, şifreli metin boyutu, şifreleme süresi ve çıktı açısından geleneksel algoritmalarla karşılaştırıldığında daha iyi sonuçlar gösterdiğini belirtmişlerdir [29].

2019 yılında Basu ve arkadaşları verilerin güvenliğini sağlamak amacıyla biyo-ilham ve makine öğrenimi tekniklerinin kullanıldığı modern bir kriptografi biçimi olan biyo-ilham verici kriptosistemlerden yararlanarak genetik kodlama (ikiliden DNA bazlarına dönüşüm), transkripsiyon (DNA'dan mRNA'ya dönüşüm) ve translasyon (mRNA'dan Protein'e dönüştürme) gibi doğal süreçleri simüle ederek şifreleme ve şifre çözme algoritmaları için Central Dogma of Molecular Biology - Moleküler Biyolojinin Santral (Merkez) Dogma'sına (CDMB) dayalı bir sistem önermişlerdir [30].

2020 yılında Tahir ve arkadaşları bulut verilerindeki veri bütünlüğü ve gizlilik sorunlarıyla başa çıkmak için CryptoGA adlı bir Genetik Algoritma (GA)'ya dayalı yeni bir model geliştirmişlerdir. Yöntemlerinde, bulut verilerinin gizliliğini ve bütünlüğünü sağlamak için bir şifreleme algoritması ile entegre edilen şifreleme ve şifre çözme anahtarlarını oluşturmak için GA kullanılmıştır. Deneylerde, değerlendirme ve karşılaştırma için yürütme süresi, aktarım hızı, anahtar boyutu ve çığ etkisi göz önünde bulundurulmuştur. Deneysel sonuç analizinin bütünlüğü sağladığını ve yetkisiz taraflara karşı kullanıcı verilerinin gizliliğini koruduğunu ifade etmekmişlerdir. Ayrıca, CryptoGA'nın DES, Triple Data Encryption Standard (3DES), RSA, Blowfish ve AES gibi diğer son teknoloji şifreleme algoritmaları ile karşılaştırıldığında sağlam olduğunu ve seçilen parametreler üzerinde daha iyi bir performans sağladığını belirtmişlerdir [31].

2020 yılında Indrasena ve arkadaşları biyo-ilhamlı bir kriptografik DNA sistemi önermişlerdir. Yöntemleri üç aşamadan oluşmaktadır: şifreleme, anahtar oluşturma ve şifre çözme. Ayrıca, CMDB kullanılmıştır. Yöntemlerini geleneksel kriptografik tekniklerle karşılaştırdılar ve sırasıyla şifreleme işlemi ve şifre çözme işlemi için işlem süresinin %67 arttığını bildirmişlerdir [32].

2021 yılında Thabit ve arkadaşları bulut bilişim güvenliğini artırmak için iki katmana sahip bir şifreleme algoritması tasarlamışlardır. Çalışmalarında, ilk katman Shannon'ın difüzyon ve karışıklık teorisinden esinlenilmiştir. İkinci katman, kriptografik amaçlı genetik kodlama yapılarından esinlenilmiştir. Genetik kriptografinin, transkripsiyonun ve translasyonun doğal süreçlerinin simülasyonunu kullanmışlardır. Boyut ve yürütme süresi açısından bulut bilişimdeki uygulamaları güvenceye almak için kullanılabilecek

olağanüstü deneysel sonuçlar elde etmişlerdir [33].

Aljazaery ve arkadaşları 2022'de yaptıkları çalışmada 2 ve 3 boyutlu renkli görüntülerin kodlanması için yeni bir yöntem ortaya koymuşlardır. Bu yöntemde DNA zinciri yapısı, yöntemin yapılandırılması için temel olarak kullanılmıştır. DNA bazları, DNA kodlama tablosu ve XOR yardımı ile oluşturulan yeni bir yöntem oluşturmuşlardır. Bu yöntem ile yaptıkları çalışmadan çıkan görüntüleri, orijinal görüntülerle karşılaştırmak için Mean square error (MSE) ve Peak signal to noise ratio (PSNR) değerlerini hesaplamış ve yüksek değerli sonuçlar elde etmişlerdir [34].

2022'de Singh ve arkadaşları, Industrial Internet of Things - Endüstriyel Nesnelerin İnterneti (IIoT) dijital görüntülerini korumak için şifreleme tekniği olarak, kaotik haritalar ve DNA kriptografisi kullanan bir görüntü güvenlik frameworklerini ele almışlardır. Önerdikleri algoritmada, üç anahtar oluşturmak için çok düzeyli bir biçimde bir çadır, daire, Chebyshev iterasyonu ve 3B lojistik harita kullanmışlardır. Bu anahtarları, alt blokların satır-sütun dönüşü için kullanmışlardır. Alt bloklar üzerinde DNA kodlama-çözme gerçekleştirecek kuralı belirler ve şifrelenmiş görüntüyü elde etmek için DNA XOR işleminin gerçekleştirildiği bir anahtar görüntü oluşturur. Önerilen şemanın sonuç analizleri olan ortalama Number of Changing Pixel Rate – Piksel Değişim Oranı (NPCR) (%99.6566), Unified Averaged Changed Intensity – Birleşik Ortalama Değişim (UACI) (%33.4588), Entropi (7.9971) ve 10195'lik daha büyük anahtar alanı değerlerinin mevcut şemalardan daha iyi olduğunu ve farklı saldırılara karşı dirençli olduğunu ortaya koymuşlardır [35].

Bu çalışma literatürden farklı olarak DNA şifreleme ve DNA operatörleri Feistel ağ yapısına entegre edilmiiş bir DNA kriptografi tekniği önerilmiştir. Burada taşıyıcı olarak görüntü, metin veya video gibi geleneksel dijital medya yerine DNA'nın kendisi kullanılırken, modern biyolojik araçlar uygulama araçları olarak kullanılmıştır. Ayrıca geliştirilen simülasyon yazılımı ve sentezlenen DNA dizisi, özel olarak oluşturulmuş biyoteknik donanıma hem dijital hem de biyolojik olarak entegre edilmiştir. Verilerin DNA ortamında saklanması da amaçlandığından, bir DNA bazına (A, C, G ve T) gömülebilecek veri (bit) miktarı burada önemli bir sorundur. Literatürdeki çalışmalar incelendiğinde nükleotit başına kodlanabilecek bit sayısının ikiyi geçmediği görülmüştür. Önerilen çalışmada, şifreleme işleminden önce gerçekleştirilen orijinal sıkıştırma algoritması ile bu oran artırılmıştır. Biyolojik olarak sentezlenen DNA dizisi, simülasyonda kod çözme işlemi ile elde edilen simüle edilmiş DNA dizisi ile

karşılaştırıldığında %100 eşleşme başarısı elde edilmiştir. Simülasyon yazılımı da tak - çalıştır şeklinde çalışabilen donanıma entegre edilmiştir. Bunlar, önerilen çalışmayı DNA tabanlı veri depolama açısından çok yönlü yapan ana özelliklerdir. Önerilen çalışmanın kriptografik gereksinimler için verimli sonuçlara sahip olduğu %100'e yakın kapasite, tek blok için yaklaşık 12×10^6 yıl kaba kuvvet saldırısı, tek blok için 2^{80} anahtar uzayı ve 2'ye yakın entropi analizi sonuçları ile gösterilmiştir. Ayrıca önerilen yöntemin uygulanması laboratuvar doğrulama testleri ile doğrulanmıştır [36].

1.2. TEZİN ORGANİZASYONU

Bu tez çalışması dört bölüm olarak düzenlenmiştir.

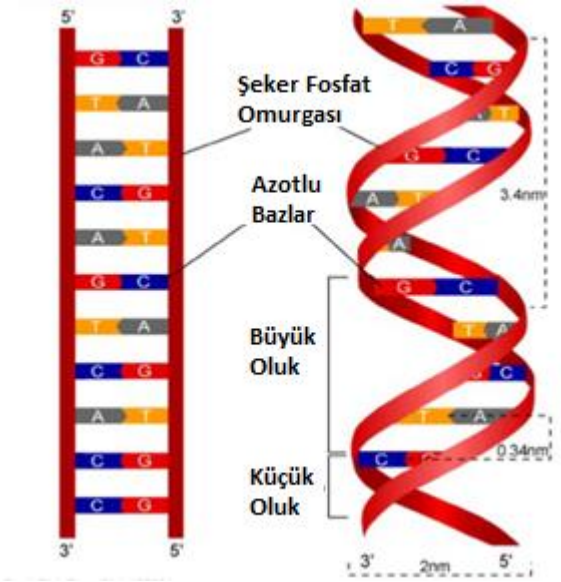
1. Bölümde literatürdeki geçmiş ve benzer çalışmalardan bahsedilmiştir. Çalışmanın genel bir anlatımı yapılmış ve tezin organizasyonu sunulmuştur.
2. Bölümde kullanılan metodolojilerden bahsedilmiş ve önerilen yaklaşım ayrıntılı örnekler verilerek adım adım açıklanmıştır. DNA'nın biyolojik yapısı, DNA sentezleme, dizileme ve kriptografisinden bahsedilmiş daha sonra uygulanan metodun süreçleri olan sıkıştırma, şifreleme ve şifre çözmeden ayrıntılı olarak bahsedilmiştir. Ayrıca yazılımın simülasyonu ve biyoteknik donanımın tasarımı görsellerle ayrıntılı olarak anlatılmıştır.
3. Bölümde gerçekleştirilen deneylerin sonuçları verilmiş ve tartışılmıştır. Burada kapasite, kaba kuvvet saldırı ve anahtar uzay analizleri sonuçlarına değinilmiş ayrıca entropi hesabına yer verilmiştir.
4. Bölümde yapılan çalışmanın sonucu ortaya konmuş ve çalışmanın olumlu / olumsuz yönlerine değinilmiştir. Çalışmanın gelecek çalışmalar için faydalı olabilecek önerileri de bu bölümde sunulmuştur.

2. MATERYAL VE METOT

Bu bölümde öncelikle DNA'nın biyolojik yapısı kısaca ifade edilerek DNA kriptografisinin temel prensiplerinden bahsedilmiştir. Daha sonra önerilen algoritma, somut örnekler verilerek her bir alt adım detaylı olarak açıklanmıştır.

2.1. DNA YAPISI

Deoksirübo Nükleik Asit (DNA), tüm canlılar için kalıtsal bir madde olup, genetik bilgiyi taşımaktadır. DNA'nın çift sarmal yapısı paralel olmayan birbiri etrafına sarılı iki biyopolimer (canlı organizmalarda bulunan birleşik moleküller; protein, karbonhidrat, nükleik asit gibi) dizi içermektedir. Her bir DNA dizisi ise, 4 çeşit nükleotid içermektedir: A,G,C ve T. DNA yapısında, Adenin- Timin ile Guanin-Sitozin ile eşleşmektedir ve buna Watson-Crick tümleyeni denmektedir. Bu yapıdaki her iki dizi Şekil 2.1'te görüldüğü gibi anti paraleldir; yani bir dizi 3' ile başlayıp 5' ile bitiyorsa, diğer dizi, 5' ile başlayıp 3' ile bitmektedir (Bu mekanizma, karbon atomları ile belirlenen bir sıradır) [1] Şekil 2.1'te DNA'nın çift sarmal yapısı görülmektedir [22].



Şekil 2.1. DNA'nın çift sarmal yapısı.

2.1.1. DNA Sentezleme

Sentezin amacına dayalı olarak birden çok tanımı vardır: “(a) Bir bütün oluşturacak şekilde parçaların veya öğelerin bileşimi veya birleşimi. (b) kimyasal elementlerin, grupların veya daha basit bileşiklerin birleşmesi veya karmaşık bir bileşiğin bozunması yoluyla bir maddenin üretilmesi. (c) Sıklıkla farklı kavramların tutarlı bir bütün halinde birleştirilmesi: bu şekilde oluşan kompleks”. Diğer bir tanım ise “(a) tümdengelimli akıl yürütme. (b) Tez ve antitezin gerçeğin daha yüksek bir aşamasına diyalektik birleşimi” Sonuç olarak, sentezi, tutarlı bir bütün oluşturmak için ayrı öğeleri veya bileşenleri birleştirmeye atıfta bulunduğu analiz prosedürünün karşısındaki prosedür olarak tanımlayabiliriz [37]. DNA sentezlenmesi ise PCR teknolojisi modifiye edilerek yapılmıştır. Veri depolama için DNA sentezi uzun zamandır katı destek üzerinde hareketsiz hale getirilmiş büyüyen bir oligonükleotit zincirine istenen nükleotidin eklenmesini içeren dört aşamalı bir döngüsel reaksiyon olan fosforamidit (Oligonükleotit sentezinde en yaygın yöntem) yöntemine dayanmaktadır. Ancak, bu yöntemin çeşitli nedenlerle çözülmemiş sınırlamaları olduğunu unutmamak gerekir [38].

2.1.2. DNA Dizileme

Nükleik asitler (DNA ve RNA) canlılardaki hücrelerde meydana gelen metabolik olayların gerçekleştirilmesinde, kontrolünde rol alan ve kalıtımı sağlayan temel yapı taşı moleküllerdir. Nükleik asitler, nükleotit (pürin/pirimidin bazı + 5 karbonlu bir şeker + bir fosfat grubu) olarak adlandırılan basit birimlerden oluşurlar. Nükleik asit moleküllerindeki nükleotit bazları (A, G, C, T) sırasının belirlenmesi DNA dizileme olarak adlandırılmaktadır.

Watson ve Crick, 1953'te Rosalind Franklin ve Maurice Wilkins tarafından üretilen ve hem DNA replikasyonu hem de nükleik asitlerdeki proteinleri kodlamak için kavramsal bir çerçeveye katkıda bulunan kristalografik verilerden çalışarak DNA'nın üç boyutlu yapısını çözmüşlerdir. Watson ve Crick’in 1953 yılında DNA’nın çift sarmal yapısını keşfi ve diğer benzer araştırmalar nükleik asit dizileme sistemlerinin kökenini oluşturmuştur [39], [40].

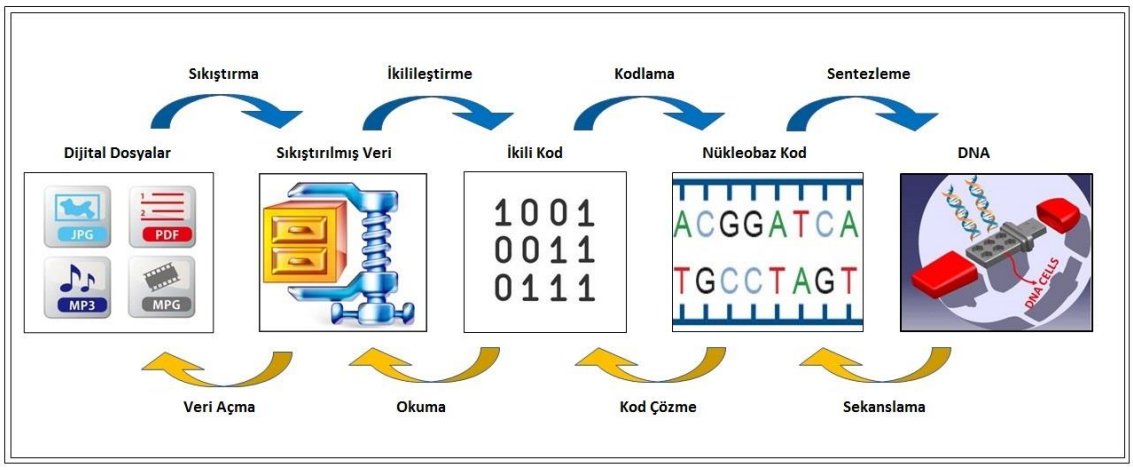
Moleküler biyoloji biliminin hızla gelişmesiyle DNA çalışmaları ön plana çıkmış ve 1965 yılından bu yana farklı dizileme yollarına, değişik örnek hazırlama stratejilerine, immobilizasyona ve dizileme kimyasına sahip çok sayıda DNA dizileme yöntemleri geliştirilmiştir. Bu yöntemler biyoenformatik yazılımlarla birleştirilerek biyolojik

araştırmalarda genom dizileme, teşhis-tanı, transkriptomik, ekolojik ve epidemiyolojik amaçlı çalışmalarda başarılı bir şekilde kullanılmaktadır [40].

2.2. DNA KRİPTOGRAFİSİ

DNA kriptografisinde, bilgi taşıyıcıları olarak DNA çiftleri kullanılır. Diğer yöntemlerle karşılaştırıldığında, DNA moleküllerinin büyük işlem gücü, DNA kriptografisini daha gelişmiş kılmaktadır. Sonuç olarak, gelecekte DNA çipi (donanım) teknolojisinin mevcut silikon çiplerin yerini alması beklenmektedir. Böylece bilgisayarlarda veri işleme büyük ölçüde artacaktır. Ek olarak, DES, RSA gibi kriptografik algoritmaların geleneksel kavramları kırılabileceğinden daha güvenli bir algoritmaya ihtiyaç vardır. DNA kriptografisinin avantajları, DNA'nın olağanüstü depolama kapasitesi, düşük güç tüketimi ve dikkate değer işleme performansıdır [22].

Bu tez çalışmasında önerilen sistemin adımları şu şekilde ifade kısaca ifade edilebilir. İlk olarak DNA kodlamasını gerçekleştirmek için her türlü dijital veri ikiliye dönüştürülür. Daha sonra DNA kodlaması ile verilerin ikili formu DNA bazlarına dönüştürülür. Burada DNA XOR, DNA kaydırma vb. DNA operatörlerinden faydalanılarak şifreleme gerçekleştirilir. Daha sonra simüle edilmiş DNA dizisi sentezlenir. Sentezlenen DNA daha sonra başlangıçta orijinal verileri elde etmek için dizilenir. Burada DNA kodlama aşaması sonucunda DNA bazlarını elde etmek için şifre çözme işlemi yapılır. Daha sonra verilerin ikili formunu oluşturmak için DNA kod çözme işlemi gerçekleştirilir. Son olarak, bu ikili form başlangıçta okunur ve dijital dosyaya dönüştürülür. Bu önerilen çalışmanın grafiksel özeti Şekil 2.2'de yer almaktadır.



Şekil 2.2. Çalışmanın grafiksel özeti.

2.3. ÖN HAZIRLIKLAR ve KULLANILAN DEĞİŞKENLER

Bu alt bölümde, denklemlerde kullanılan sembollerin kısa bir açıklaması verilmiştir.

$I_{1,n}$: Girdi dosyası dizisi.

$B_{1,n\times 8}$: Baz 2'de giriş dosyası vektörü

$G_{1,m}$: Kelimeleri üç bitten oluşan dize dizisi.

w: Uzunluğu 3 bit olan *G'*nin kelimesi

$D_{1,l}$: Sıkıştırılmış DNA dizisi

$B'_{1,l\times 2}$: DNA Kodlaması ile *D'*nin Bit Dizisi

$G'_{1,j}$: Elemanları 12 bit uzunluğunda olan B'

w': Uzunluğu 12 bit olan G' kelimesi

L_i: w' nin sol yarısı

R_i: w' nin sağ yarısı

L_{i+1}: w'nin sonraki sol yarısı

R_{i+1}: w'nin sonraki sağ yarısı

s: S-kutusu çıktısı

$E_{1,j}$: Şifreli DNA dizisi

$B''_{1,j\times 2}$: DNA kodlaması yoluyla *E*'nin ikili formu

$T_{1,j\times 2}$: Elemanları 12 bit uzunluğunda olan B''

Aşağıda verilen Çizelge 2.1'de DNA XOR işlemini göstermektedir [41].

Çizelge 2.1. DNA XOR işlemi.

XOR	A	G	T	C
A	A	G	T	C
G	G	A	C	T
T	T	C	A	G
C	C	T	G	A

Çizelge 2.2'de DNA baz bit dönüşümüne yer verilmiştir [42].

Çizelge 2.2. DNA baz bit dönüşümü.

Nükleotit	İkili Sayı
A	00
C	01
G	10
T	11

Sıkıştırma işlemi: Önerilen çalışmada sıkıştırma yöntemi, iki boyutlu düzlem ve iyi bilinen DNA baz-bit dönüşümü (Patent Pending Identifier: 2017/00459) kullanılarak oluşturulmuştur. Böylece, sıkıştırma kapasitesi yaklaşık %180 oranında artırılırken girdi ve çıktı arasındaki karmaşık ilişkiye katkıda bulunulmuştur.

2.4. SIKIŞTIRMA SÜRECİ

Adım 1: Bu adımda, girdi verisi olarak bir metin dosyası alınır ve Denklem 2.1'de belirtildiği gibi ikili verilere dönüştürülür:

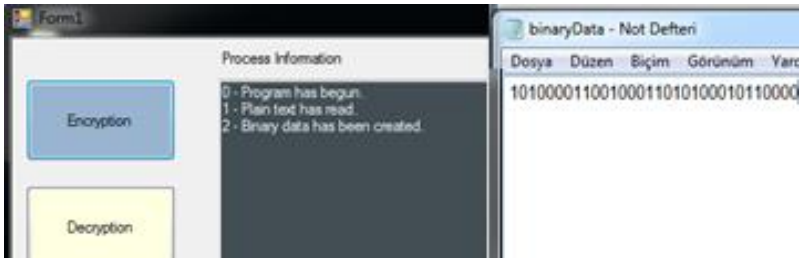
$$B_{1,n\times 8} = Baz\ (I_{1,n}) \tag{2.1}$$

Örneğin:

Metin: *ESRA SATIR DÜZCE ÜNİVERSİTESİ.* [30 karakter]

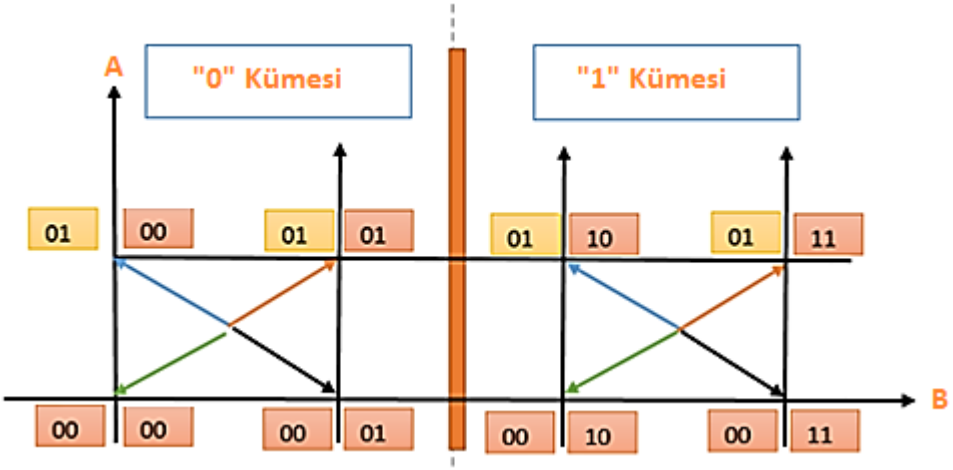
İkili sistemde oluşan veri: (*01000101, 01010011, 01010010, 01000001, 00100000, 01010011, 01000001, 01010100, 01001001, 01010010, 00100000, 01000100, 01010101, 01011010, 01000011, 01000101, 00100000, 01010101, 01001110, 01001001, 01010110, 01000101, 01010010, 01010011, 01001001, 01010100, 01000101, 01010011, 01001001, 00101110*) (30×8=240 bit)

Şekil 2.3'te girdi olarak sunulan metnin ikili veriye dönüşümünün simülasyonda gösterimi aşağıdaki gibidir.



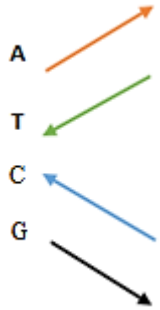
Şekil 2.3. Girdi olarak sunulan metnin ikili veriye dönüşümünün gösterimi.

Adım 2: Bu adımda, elde edilen veriler, her biri üç bit içeren gruplara bölünür (oluşan bit dizisi üçün katı değilse gerekli sayıda 0 eklenir). Burada bahsedilen sıkıştırma işlemine uyularak DNA bazları vektörler olarak ifade edilir. Kümeler, iki boyutlu düzlemin yardımıyla ayarlanır. DNA'nın baz vektörleri, *B* eksenindeki her küme içinde yer alırken, *A* eksenindeki koordinatlar iki bit olarak ifade edilir. İlk olarak, her bir üçlü bitin koordinatları bulunur ve ardından karşılık gelen baz atanır. Daha sonra, bazın karşılık gelen kümesi belirlenir ve her ikili küme, bit-baz dönüşümüne tabi tutulur. Bu işlemler için tasarlanan iki boyutlu düzlem ve küme eşdeğerleri Şekil 2.4'te gösterilmiştir.



Şekil 2.4. Tasarlanan iki boyutlu düzlem ve küme eşdeğerleri.

İki boyutlu düzlemi iki bölgeye ayrılmaktadır; *"0" Kümesi* ve *"1" Kümesi*. Her kümede ki dört vektörde aynıdır. Şekil 2.5'te, bu dört vektör, anlaşılmasını kolaylaştırmak için ayrıntılı olarak verilmiştir. Yani DNA bazlarının vektör temsilleri belirtilmiştir. Bunları baz vektörleri olarak adlandıralım:



Şekil 2.5. DNA bazlarının vektör gösterimleri; baz vektörleri.

Bu adımın matematiksel süreci Denklem 2.2'de gösterildiği gibidir:

$$G_{1,m} = \{w: |w| = 3 \ \& \ w \in B\} \tag{2.2}$$

G = (010, 001, 010 ,101, 001, 101, 010, 010, 010, 000, 010, 010, 000, 001, 010, 011, 010, 000, 010, 101, 010, 001, 001, 001 , 010, 100, 100, 010, 000, 001, 000, 100, 010, 101, 010, 101, 101, 001, 000, 011, 010, 001, 010, 010, 000, 001, 010, 101, 010 , 011, 100, 100, 100, 101, 010, 110, 010, 001, 010, 101, 001, 001, 010, 011, 010, 010, 010, 101, 010, 001, 000, 101, 010, 100 , 110, 100, 100, 100, 101, 110)

G dizisinde toplam 240 bitimiz var. İlgili DNA bazları, Şekil 2.4 ve Şekil 2.5'te görüldüğü üzere *Küme Eşdeğerleri* ve *Primer bazları*, bahsedilen sıkıştırma prosedürü dikkate

alınarak oluşturulmaktadır.

Örneğin; w = 010 olduğunu varsayalım. Bu durumda aslında w' = 0010'dır. Tasarlanan düzlemde tüm üç bitlik dizlerin dört bitlik tasarımında ilk dizi karakteri 0 olduğu görülmektedir. İki boyutlu düzlemde w' = 0010'u ikişer bit olarak ayırıp aradığımızda *"1" Kümesi* bölgesinde sol altta oluştuğu ve bu w' için yeşil baz vektörünün eşleştiğini görebiliriz, bu baz vektörü ise *T* bazına karşılık gelmektedir. Bu durumda w = 010 bit dizisi için *T* bazı elde edilir.

Tüm DNA bazları bu şekilde elde edilir. Bu adım sıkıştırmanın gerçekleştiği adımdır. Burada her üç bit için bir *DNA bazı* ve bir *Küme Eşdeğeri* elde edilmiş olur. Yukarıda verilen bit dizisi için oluşan *DNA Bazları* ve *Küme Eşdeğerleri* aşağıdaki gibi oluşacaktır.

DNA Bazları:

TGTAGATTTTTTTGTGTTTATGGGTCCTTGTCTATAAGTGTGTTTTTATGCCCATCTGTAGGT GTTTATGTATCCCCCAC (80 baz)

Küme Eşdeğerleri:

1,0,1,0,0,0,1,1,1,0,1,1,0,0,1,1,1,0,1,0,1,0,0,0,1,0,0,1,0,0,0,1,0,0,0,1,0,1,0,0,0,1,1,0,1,1,1,0,0,1,0,1,1, 0,0,0,0,1,1,1,1,1,0,0,0,1,1,1,1,1,0,1,0,0,0,1,0,1, 0,0,0,0,1 (80 bit)

Elde edilen *Küme Eşdeğerleri* ise kendi içerisinde her iki bit sırasıyla Çizelge 2.3'te Primer tablosu yardımı ile DNA baz dizisi elde edilir.

Çizelge 2.3. Primer tablosu.

A	00
T	01
G	10
C	11

Primerler:

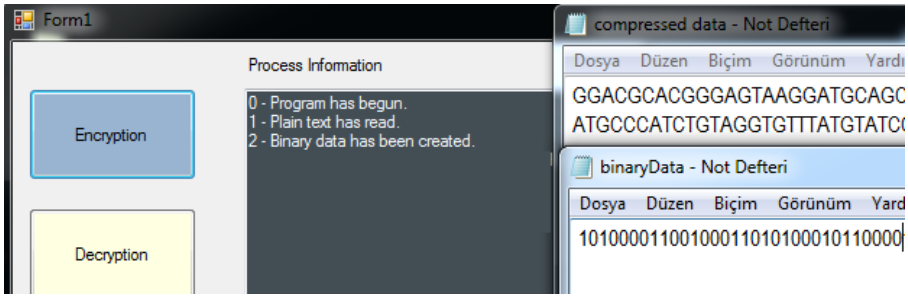
GGACGCACGGGAGTAAGGATGCAGCAACCGACCGGAGGAT (40 baz)

Adım 3: Bu adımda, Denklem 2.3'te verilen DNA bazları ve Primer Bazları bir araya getirilerek sıkıştırılmış DNA dizisi elde edilir:

$D_{1,l}$ = Primer Bazlar // DNA Bazları (2.3)

D=(*GGACGCACGGGAGTAAGGATGCAGCAACCGACCGGAGGAT // TGTAGATTTTTTTGTGTTTATGGGTCCTTGTCTATAAGTGTGTTTTTATGCCCATCTGTAGGT GTTTATGTATCCCCCAC)* (40+80 = 120 baz)

Şekil 2.6'da yukarıda oluşturulan sıkıştırılmış DNA dizisine ait simülasyon ekran görüntüsü aşağıda gösterilmiştir.

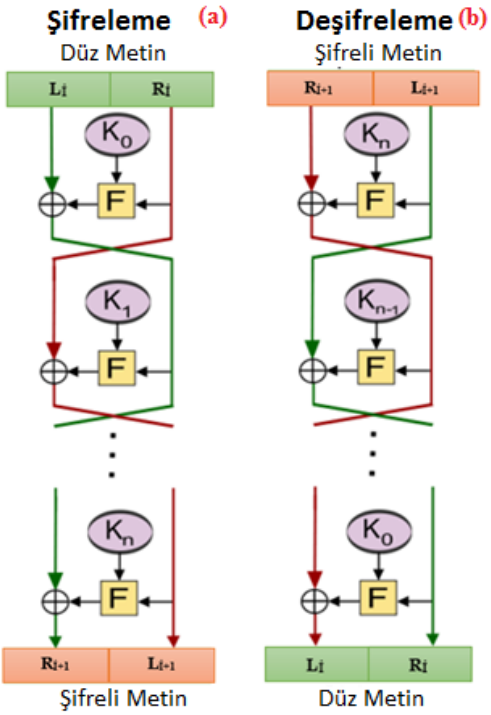


Şekil 2.6. Sıkıştırılmış DNA dizisi.

2.5. ŞİFRELEME SÜRECİ

Önerilen çalışmada DNA kodlaması [42] ve DNA XOR [43], Feistel ağına entegre edilmiştir. Feistel ağı, geleneksel kriptografide şifreleri bloke etmek için uygulanan simetrik bir yapıdadır. Feistel'in fiziksel tekrarı, donanım üzerinde uygulamayı kolaylaştırır ancak geleneksel yapısı 16 tur işlem gerektirmektedir [44]. Bu nedenle, Feistel'in yapısının gereği değiştirilebilir olması birçok şifreleme çalışmasında Feistel'in geliştirilerek kullanılmasını sağlamıştır [45], [46]. Feistel ağı Şekil 2.7'de gösterilmiştir [47]. Bu entegrasyon, geleneksel DES S-kutuları kullanılarak gerçekleştirilmiştir. Feistel ağlarında, yaklaşımın özgünlüğünü ve karmaşıklığını ölçtüğü için *F* fonksiyonunun tasarımı çok önemlidir. Burada, Şekil 2.8(b)'de gösterildiği gibi DNA XOR ve DNA kaydırmasından yararlanılarak yeni bir *F* fonksiyonu geliştirilmiştir. Şekil 2.8(a), DES S-kutularının kullanımını göstermektedir. Sonuç olarak, önerilen şema simetrik şifrelemenin bir örneğidir.

Matematiksel olarak, Feistel şifrelemesinin yapısı Denklem 2.4 ve 2.5'te görülmektedir.



Şekil 2.7. Feistel Şifrelemenin blok şeması (a)Şifreleme süreci (b)Şifre çözme süreci.

$$L_{i+1}=R_i \tag{2.4}$$

$$R_{i+1}=L_i \oplus F(R_i,K_i) \tag{2.5}$$

Adım 1: Bu adımda, sıkıştırmadan sonra elde edilen baz dizisi DNA Baz-Bit transformasyonu yoluyla ikiliye dönüştürülür, Denklem 2.6 ile aşağıdaki gibi ifade edilir:

$$B'_{1,l\times 2} = DNA\ Encode(D_{1,l}) \tag{2.6}$$

Adım 2: Bu adımda, elde edilen bit dizisi Denklem 2.7 aracılığıyla 12 bitlik blok gruplarına bölünür.

$$G'_{1,j=}\{w' : |w'| = 12\ \&\ w' \in B'\} \tag{2.7}$$

Burada her grup yine kendi içinde iki eşit parçaya bölünür. Bu 6 bitlik gruplara sol, L_i ve sağ R_i adı verilir. Böylece, L_i ve R_i, Çizelge 2.4'te açıklandığı gibi Feistel ağ yapısına girmeye hazırdır.

Çizelge 2.4. Sıkıştırılmış verileri Feistel Ağ yapısına uyarlama süreci.

D	TGTAGATTTTTTGTGTTTATGGGTCCTTGT…							
B’	1110110010111111111111111011101111110011101010110101111111011…							
G’	111011001011		111111111111		101110111111		…	
6 bitlik gruplar	111011	001011	111111	11111	101110	111111	…	…
Feistel Parçaları	L_i	R_i	L_i	R_i	L_i	R_i	L_i	R_i

Adım 3: Feistel ağlarında *F* fonksiyonu, tasarımcıya göre yapıyı özgün kılar. Önerilen çalışmada, DES S-kutuları, DNA XOR ve DNA kaydırma operatörleri kullanılarak yeni bir *F* fonksiyonu oluşturulmuştur. Bu işlemin sözde kodları, anlaşılmasını kolaylaştırmak için aşağıda verilmiştir:

a) w' iki eşit parçaya bölün (R_i, L_i)
b) Sol ve sağ parçaları değiştirin
(L_i) ↔ (R_i)
For x = 1 to 8
{
s= S_kutusu_Sx (R_i)
For y = 1 to 5
e’ = DNA_F_Function(s) (bkz Şekil 2.8.*b)*
}

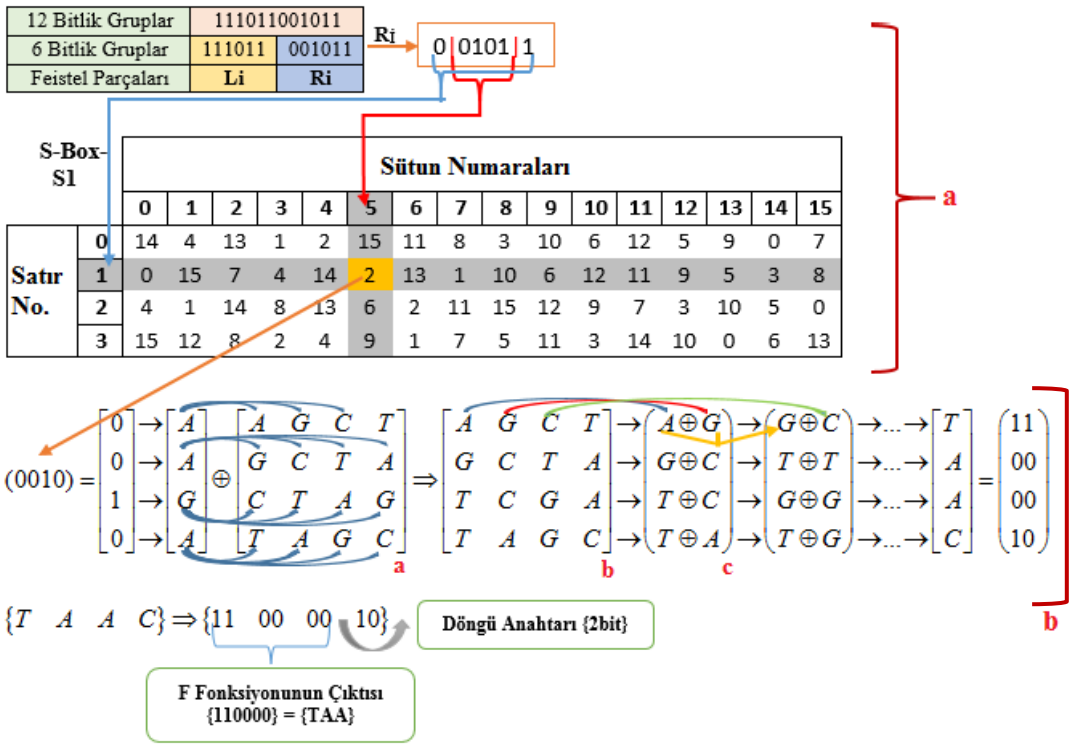
Şekil 2.8(a)'da, Çizelge 2.4’te ki R_i, *S-kutusu_S1*'e girer [48]. Dolayısıyla çıktı 4 bit uzunluğunda olur. Elde edilen 4 bite karşılık gelen bazlar Şekil 2.8(b)' de *Matris_1* ile DNA XOR'a tabi tutulur. Sonuç olarak, *Matris_2* elde edilir. *Matris_2* tekrar Şekil 2.8(b)'de yer alan *3*.üncü işlemde gösterildiği gibi DNA XOR işlemine tabi tutulur. Sonuç hala 4 baz ve 8 bitten oluşur. Ayrıca bu, *F* fonksiyonumuzun çıktısıdır. Burada, her 8 bitin 2 LSB biti, Şekil 2.8(b)'de gösterildiği gibi döngü anahtarlarıdır. Kalan 6 bit yani 3 baz Feistel ağında *Li* ile XOR işlemine girmeye hazırdır. Toplamda 8 S-kutusu *(S1,...,S8)* için 8 döngü yapılır ve her *S-kutusu* döngüsü için *Matris_1,* 5 kez kaydırılır . Böylece 40 döngüden sonra L_{i+1} ve R_{i+1} oluşur. Sonunda, kodlanmış DNA zinciri elde edilir. Şekil 2.8(a) ve (b)'de, geliştirilen *F* fonksiyonunun sadece bir turu sunulmuştur. Bu işlemin sözde kodları, anlaşılmasını kolaylaştırmak için aşağıda verilmiştir:

$L_{i+1} = R_i$

$R_{i+1} = e'\ XOR\ L_i$

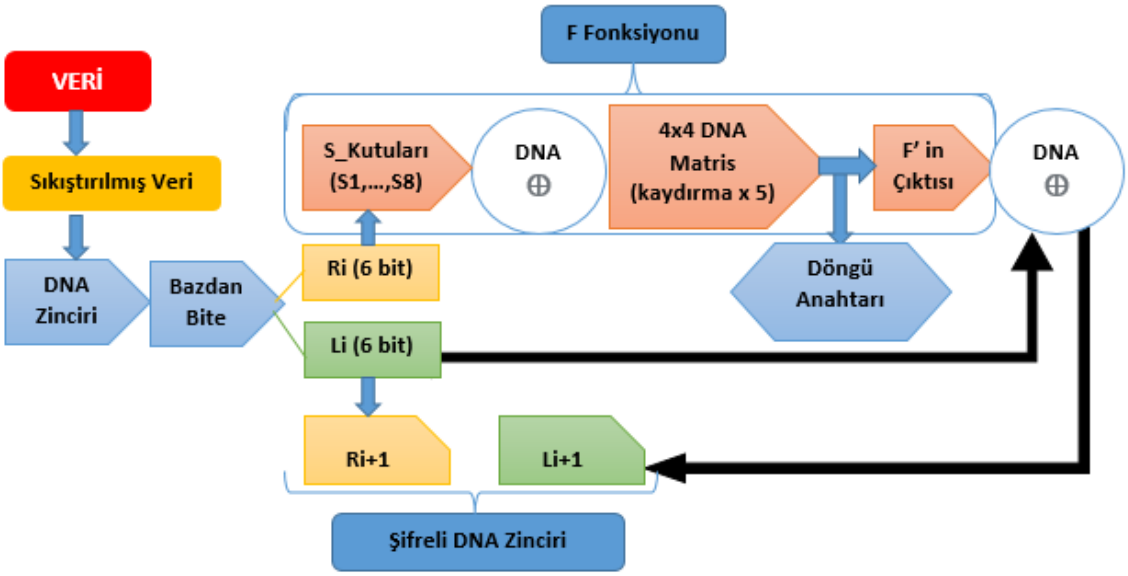
$e = L_{i+1} // R_{i+1}$

$E = \{e: e = L_{i+1} // R_{i+1}\}$



Şekil 2.8. (a) DES S_Kutusu_S1'in kullanımı (b) Tasarlanan *F* fonksiyonu.

Şekil 2.9'da her bir adım ayrılarak ve işaretlenerek şifreleme işleminin bir akış şeması verilmiştir.



Şekil 2.9. Şifreleme işleminin akış şeması.

2.6. ŞİFRE ÇÖZME SÜRECİ

Şifreleme işleminin simetrik yapısı sayesinde, şifreleme işlemi tersine çevrilerek kod çözme işlemi kolaylıkla elde edilebilir. Şifreli DNA dizisi Şekil 2.7(b)'de görüldüğü gibi Feistel ağ yapısına uyarlanarak düz veri elde edilebilir.

Denklem 2.8 ve 2.9 aracılığıyla matematiksel olarak aşağıdaki gibi gösterilebilir:

$$R_i = L_{i+1} \tag{2.8}$$

$$L_i = R_{i+1} \oplus F\,(L_{i+1}, K_i) \tag{2.9}$$

Adım 1: Bu adımda, şifrelenmiş DNA zinciri, Şekil 2.7(b)'de gösterildiği gibi Feistel ağını sunmak için 6 bazlık gruplara bölünmüştür. Elde edilen bu 6-bazlık gruplar, L_{i+1} ve R_{i+1} olarak ifade edilir. Yine Feistel yapısına uyum sağlamak için iki yarım şeklindedirler ama bu sefer tersine akan bir süreç vardır. Bu işlem Denklem 2.10 ile gösterilir:

$$B''_{1,j\times 2} = Baz_2(E_{1,j}) \tag{2.10}$$

Adım 2: Bu adımda B'' her biri 12 bit uzunluğunda kelimelere ayrılır. Ortaya çıkan diziye T diyoruz. Şifre çözme işleminde, L_{i+1}, F fonksiyonunun girişidir. L_i, doğrudan ve burada F alır işlev, şifreleme işleminin tersi sırayla çalışır. R_{i+1}'de oluşturulan 6 bitlik gruplar ve F fonksiyonunun sonucu DNA XOR işlemeye girer. Dolayısıyla, elde edilen sonuç R_i'dir. Bu süreç Denklem 2.11 ile gösterilmiştir:

$$T_{1,j\times 2} = \{w'' : |w''| = 12 \ \& \ w'' \in (B''_{1,j\times 2})\} \quad (2.11)$$

Bu işlemin sözde kodları, anlaşılmasını kolaylaştırmak için aşağıda verilmiştir:

a) w′ iki eşit parçaya bölün ($R_{i'}$, $L_{i'}$)

Şuna dikkat edin: $L_{i'} = L_{i+1}$ *ve* $R_{i'} = R_{i+1}$ çünkü şifre çözme işlemi şifrelemenin ters fonksiyonudur.

```
For y = 1 to 5
 {
        e'= tersine çevir (DNA_F_fonksiyonu (Li')
        For x = 1 to 8
        s= tersine çevir (S_kutusu_Sx (e'))
 }
```

Adım 3: Feistel'in yapısı nedeniyle, L_{i+1} yeni L_i olur. Elde edilen L_i ve R_i grupları Bit-Baz'a tabi tutulur ve dönüştürülerek düz DNA dizisi elde edilir.

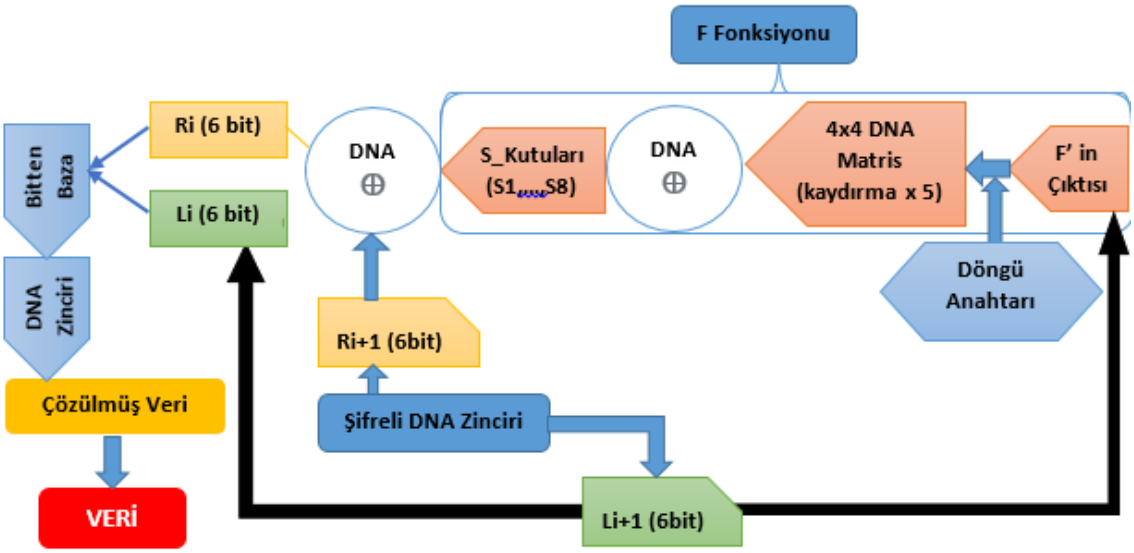
$R_{i+1} = s \ XOR \ L_{i'}$

$L_{i+1} = R_{i'}$

$b' \in [B'] = L_{i+1} // R_{i+1}$

D'nin *B*'nin DNA kodlamasına karşılık geldiğine dikkat edin. Burada sıkıştırma işleminin tersi olan veri açma işleminden sonra *B* elde edilir. *B*, giriş verilerinin *I* ikili biçimi olduğundan, ilk giriş dosyası elde edilmiştir.

Şekil 2.10'da, her bir adım ayrılarak ve işaretlenerek şifre çözme işleminin akış şeması verilmiştir.



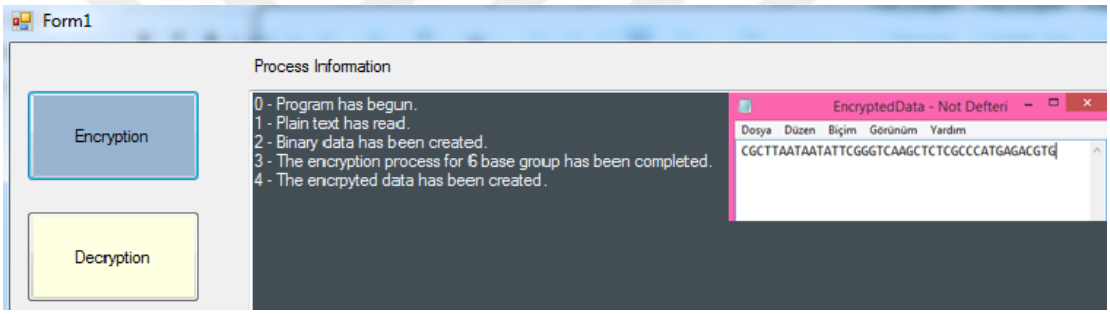
Şekil 2.10. Şifre çözme işleminin akış şeması.

2.7. SİMÜLASYON VE TASARIM

Bu bölümde yazılımın simülasyonu adım adım ekran çıktıları ile aktarılmış ve biyoteknik donanımın detayları görsellerle desteklenerek paylaşılmıştır.

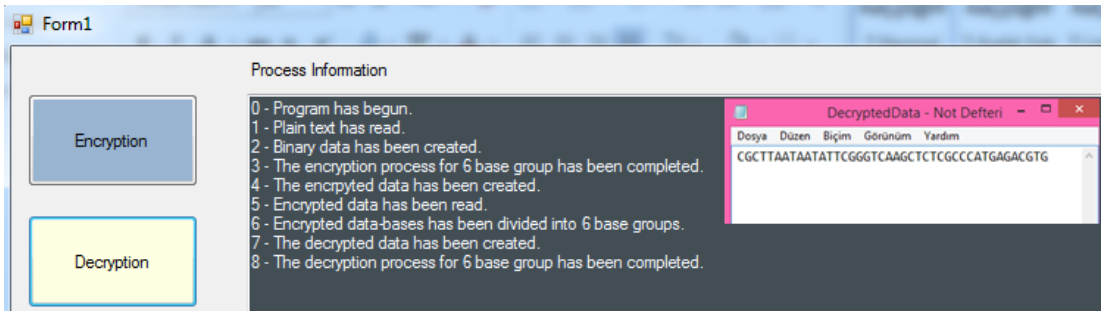
2.8. METODUN SİMÜLASYONU

Önerilen yöntemin simülasyonu C# dili kullanılarak gerçekleştirilmiştir. Yazılım testi Intel 7, 2.2 gigahertz (GHz) işlemci, 6 gigabayt (GB) Random-access memory – Rasgele erişimli bellek (RAM) ve Windows 7 64bit işletim sistemine sahip bir bilgisayar tarafından yapılmıştır. Şifreleme adımını içeren simülasyon çıktısı Şekil 2.11'de sunulmuştur.



Şekil 2.11. Simülasyonun şifreleme aşaması.

Şekil 2.12'de de şifre çözme aşaması simülasyon ekran çıktısı aşağıda görülmektedir.



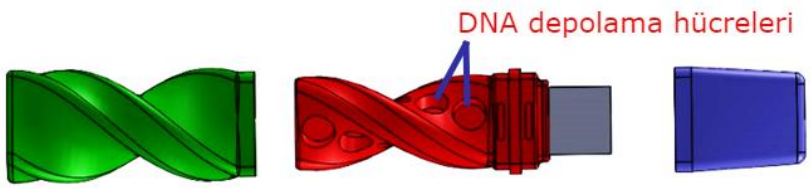
Şekil 2.12. Simülasyonun şifre çözme aşaması.

Şekil 2.12'de "Süreç Bilgisi" bölümünde simülasyon yazılımının akışı izlenebilmektedir. Şifreleme ve şifre çözme işlemlerinden sonra ilgili çıktı, arayüzün sağ tarafında verilen not defteri dosyası aracılığıyla alınabilmektedir.

2.9. BİYOTEKNİK DONANIMIN TASARIMI

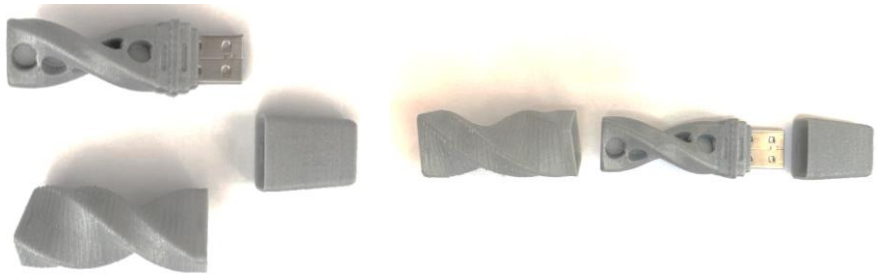
Biyoteknik donanım, CATIA *V5* programında çizilen benzersiz bir tasarıma sahiptir. Oluşturulan tasarım 3 boyutlu yazıcı yardımıyla üretilmiştir. Ayrıntılı bir görünüm için Şekil 2.13'ün dikkate alınması gerekir. Şekil 2.13'te ki biyoteknik donanımdaki DNA hücreleri, sentezlenen DNA'nın saklanmasına olanak sağlamaktadır. Öte yandan, şifreleme veya şifre çözme sonucu ortaya çıkan ilgili sayısal DNA dizisini üreten simülasyon yazılımı, Windows tabanlı herhangi bir sistemde çalıştırılmaya hazırdır.

Çalışmaların genel özellikleri incelendiğinde DNA hesaplama işlemlerinin giderek arttığı görülmüştür. Ancak, bu sürecin her zaman doğrusal olarak biyolojik sürecin hızına bağlı olduğu bilinmektedir. Ayrıca, yapılan çalışmaların çoğu simülasyonla sınırlı kalmıştır yani sadece yazılımın simülasyonu yapılmış ancak laboratuvar deneyleri yapılmamıştır. Önerilen çalışmada bu eksiklikler giderilmiştir. Öncelikle DNA'nın doğasına uygun yeni bir DNA kriptografi algoritması geliştirilmiştir. Daha sonra simülasyon yazılımı ile üretilen DNA zinciri biyolojik olarak üretilmiştir. Ayrıca yazılım ve üretilen DNA zinciri tak-çalıştır biyoteknik donanıma entegre edilerek simülasyonda kalan sınırlamalar çözülmüştür. Geliştirilen biyoteknik donanımın içyapısı Şekil 2.14'te gösterilmiştir. Burada entegre DNA hücreleri, sentezlenen DNA zincirini (dizisini) DNA damlacıkları şeklinde depolamak üzere tasarlanmıştır. Bu nedenle, her DNA hücresini 1 mg DNA depolayacak şekilde tasarlanmıştır. 1 gram DNA'nın 10^8 terabayt veri depolayabildiği yaygın olarak bilinmektedir. Tasarlanan biyoteknik donanımda 6 DNA hücresi yer aldığı için 6 x 10^5 terabayt veri depolayabilmektedir.



Şekil 2.13. Geliştirilen biyoteknik donanımın simülasyonu.

Geliştirilen biyoteknik donanımın somut olarak üretilmiş hali (çıktısı) aşağıda Şekil 2.14'te görülmektedir.



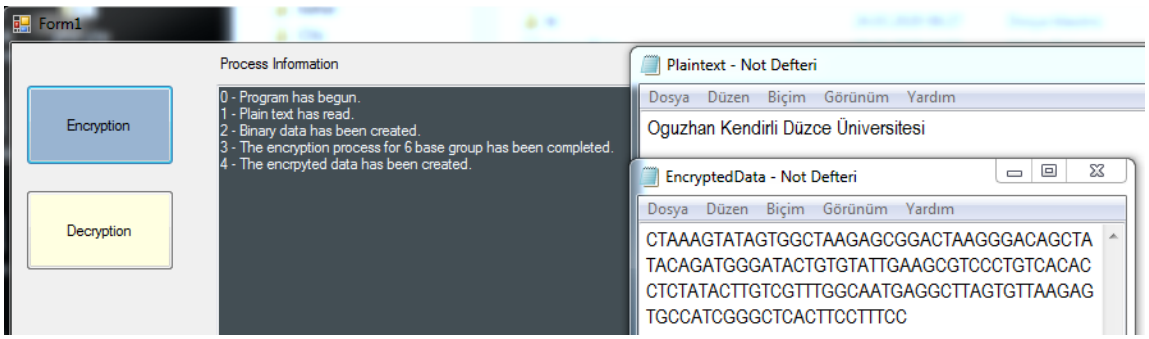
Şekil 2.14. Geliştirilen biyoteknik donanım.

3. DENEYSEL SONUÇLAR VE TARTIŞMA

Bu bölümde mümkün olduğunca somut örnekler verilerek yapılan biyolojik deneyler, performans ve güvenlik analizleri ve sonuçları sunulmuştur. Bu bölümde sunulan bulgular ve karşılaştırmalar yapılırken kapasite, kaba kuvvet saldırı, anahtar uzay analizi ve entropi alt başlıkları dikkate alınmıştır. Bu alt başlıklar ışığında karşılaştırmalarda sonuçları verilen çalışmalar tarafımızca ulaşılabilir olanlardır. Ayrıca verilen alt başlıklar deneysel parametreler olarak bir araya getiren başka bir çalışmaya da rastlanmamıştır. Bu nedenle tüm deneysel parametrelerin bu tez çalışmasında yapılan atıflar ile karşılaştırılarak desteklenmesi mümkün olmamıştır.

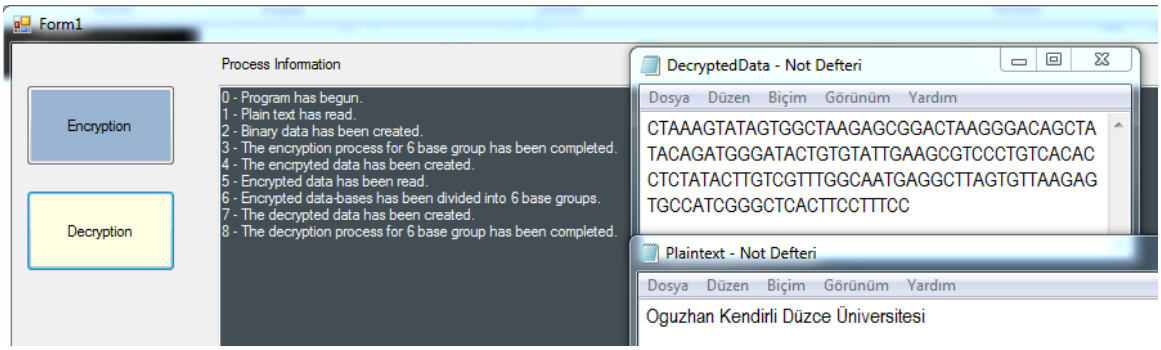
3.1. LABORATUVAR DENEYLERİ

Önerilen çalışmada düz metin örnek olarak "Oğuzhan Kendirli Düzce Üniversitesi" alınmıştır. Daha sonra şifreleme sonucunda elde edilen içerik Şekil 3.1'de gösterilmiştir.



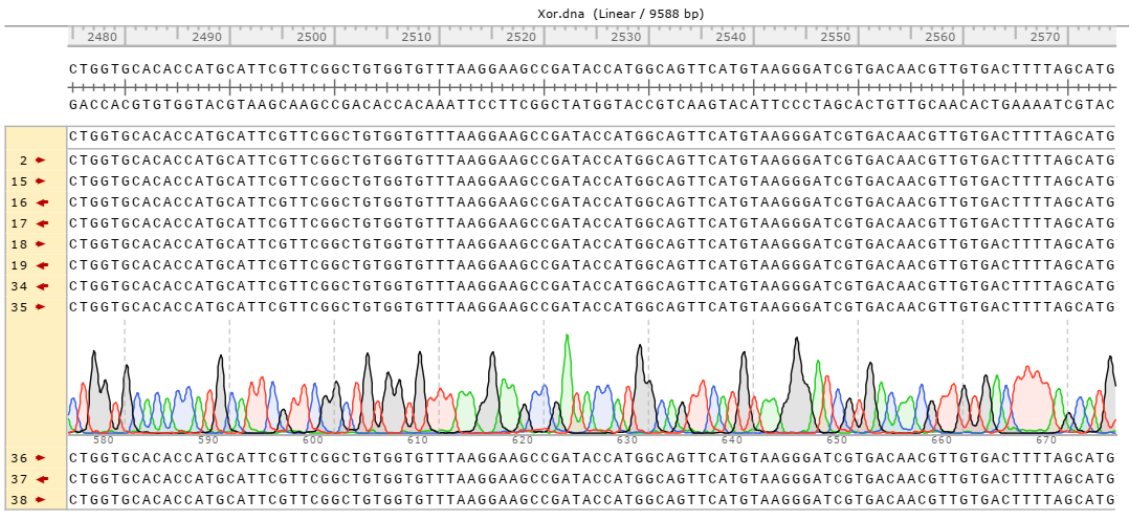
Şekil 3.1. Kodlama işlemi için örnek simülasyon çıktısı.

Elde edilen kodlanmış veriler, kod çözme işlemine tabi tutulursa, sonuç Şekil 3.2'de gösterilen düz veri olarak kendini gösterir.



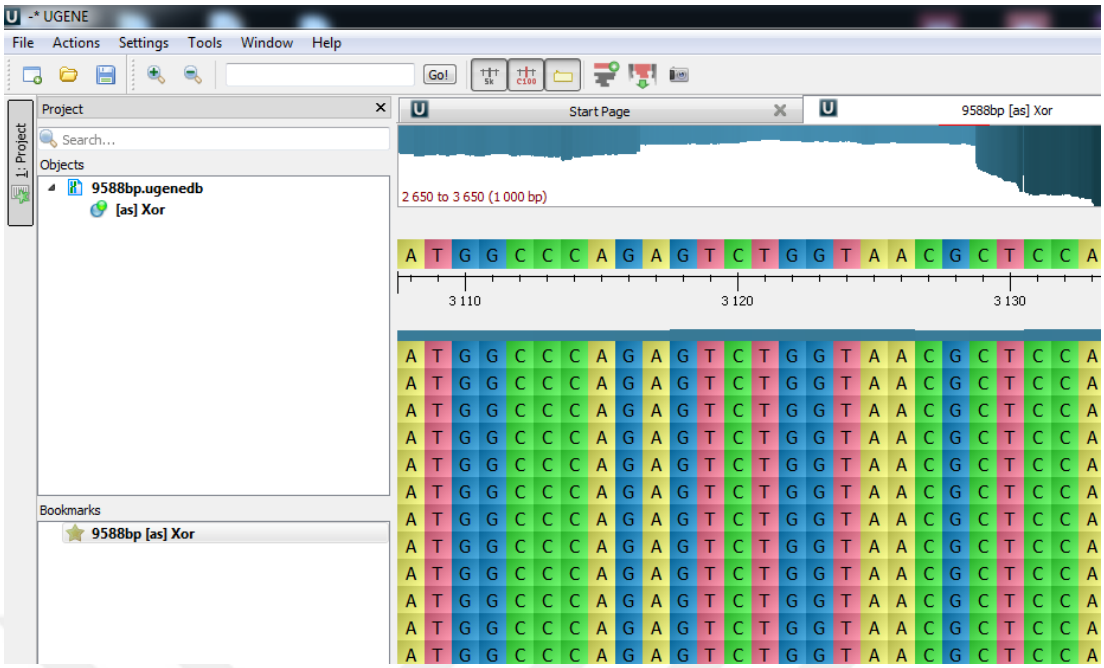
Şekil 3.2. Kod çözme işlemi için örnek simülasyon çıktısı.

Önerilen çalışmada Düzce Üniversitesi Bilimsel Araştırma Projeleri Koordinatörlüğünün desteğiyle elde edilen 9588 base pair – baz çiftinin (bp) şifreli DNA dizisi sentezlenmiştir. DNA zinciri sentezlendikten sonra, orijinal DNA dizisini yeniden okuyarak elde etmek temel bir gerekliliktir. Bu işleme DNA dizileme (sekanslama) denir. Şekil 3.3, sentezlenen DNA zincirinin dizileme sonucunu göstermektedir. Şekil 3.4'de ise 9588 bp DNA zincirinin dizileme sonucu yer almaktadır ve burada DNA zincirlerinin bilgi kaybı olmadan eşleştiği görülmektedir. Sentezlenen 9588 bp DNA zinciri ise aşağıda Şekil 3.5'te görüldüğü üzere bir plazmit içinde kapsüllenmiştir.



Şekil 3.3. DNA zincirinin sentez sonucu.

Şekil 3.4'te sentezlenmiş olan DNA zincirinin sekans sonucu yer almaktadır.



Şekil 3.4. DNA zinciri dizileme (sekanslama) sonucu.

Şekil 3.5'te, sentezlenen 9588 bp DNA zincirinin tüplerde depo edilmiş hali görülmektedir.

Şekil 3.5. Sentezlenen 9588 bp DNA zincirinin tüplerdeki görseli.

3.2. KARŞILAŞTIRMA KRİTERLERİ

Her DNA şifreleme algoritması tarafından yerine getirilmesi gereken bir dizi gereksinim vardır [6]. Önerilen çalışmada, karşılaştırma için verilen bu gereksinimler dikkate alınmıştır. Çizelge 3.2'de mevcut DNA şifreleme algoritmalarının yerine getirmeleri gereken gereksinimler açısından karşılaştırma sonuçları sunulmuştur.

Her DNA şifreleme algoritması tarafından bir dizi gereksinim yerine getirilmelidir. Bu

gereksinimler belirlenmiş ve mevcut şifreleme algoritmalarında gözlemlenen sınırlamalar Çizelge 3.1'de listelenmiştir. Önerilen çalışmada bu gereksinimler temel alınmıştır. Çizelge 3.2'de DNA şifreleme algoritması çalışmaları yapan mevcut çalışmaların gereksinimleri yerine getirime durumlarını göstermektedir [6]. Çizelge 3.2 incelendiğinde UbaidurRahman [6] ve Thabit [33]'in arkadaşlarının yapmış olduğu çalışmalar ile kriterleri karşılama bakımından başa baş kaldıkları görülmüştür. UbaidurRahman [6] ve Thabit [33]'in arkadaşlarının yapmış olduğu çalışmalarda laboratuvar doğruluğu testleri parametresi dışında diğer parametreleri sağlamış iken, önerilen bu tez çalışmasında ise dinamik kodlama tablosu oluşturma kriteri dışında ki diğer tüm kriterler sağlanmıştır.

Çizelge 3.1. DNA şifreleme algoritmasının yerine getirmesi gereken gereksinimler ve açıklamalar.

No.	Gereksinimler	Açıklamalar
1	Tam karakter setinin DNA kodlamasının gerçekleştirilmesi [6]	DNA kodlama işleminde tüm karakter setini (alfabeler, sayılar ve özel karakterler) işlemelidir. Düz metnin tüm karakter setini DNA dizisine kodlamak için kullanılabilir olmalıdır.
2	Dinamik kodlama tablosu oluşturma [6]	Daha yüksek bir güvenlik düzeyi sağlamak için, kodlama tablosu periyodik olarak yeniden oluşturulmalıdır. Karakter setinin her bir elemanı için rastgele farklı DNA dizileri oluşturmak da önemlidir.
3	Düz metnin her karakterinin DNA dizisine kodlanması için benzersiz dizi [6]	Alıcı ve gönderici arasındaki her oturumda, düz metnin DNA dizisine kodlanması, kodlama tablosunun her neslinde karakter kümesinin her bir öğesi için benzersiz olmalıdır.
4	Kodlamanın sağlamlığı [6]	Saldırılara karşı direnci artırmak için, düz metnin DNA kodlaması, deşifre edilmesi zor olan sağlam bir kodlama şemasına sahip olmalıdır.
5	Biyolojik süreç simülasyonu [6]	DNA şifreleme ve şifre çözme algoritmaları, dijital verilerle uyumlu bir şekilde çalışan simüle edilmiş biyolojik süreçlere dayanmalıdır.
6	Şifreleme sürecinin dinamikliği [6]	Aynı düz metnin farklı şifreli metinler üretebilmesi için her oturum için benzersiz bir DNA kodlama tablosu oluşturmak gerekir ve bu dinamik yapı oldukça önemlidir.
7	Laboratuvar Doğrulama Testleri	Modelin DNA sentezleme ve sekanslama aşamalarının gerçekleştirildiği somut laboratuvar deneyleriyle desteklenmesi önemlidir.

Çizelge 3.2. Mevcut DNA şifreleme algoritmalarının gereksinimler açısından karşılaştırma sonuçları.

Yazarlar	Tam karakter setinin DNA kodlamasının gerçekleştirilmesi	Dinamik kodlama tablosu oluşturma	Düz metnin her karakterinin DNA dizisine kodlanması için benzersiz dizi oluşturma	Kodlamanın sağlamlığı	Biyolojik süreç simülasyonu	Şifreleme sürecinin dinamikliği	Laboratuvar Doğrulama Testleri
G. Cui ve arkadaşları [49]	X	X	X	X	*	*	X
Q. Zhang ve arkadaşları [50]	X	X	X	X	*	X	X
S. Sadeg ve arkadaşları [51]	X	X	X	X	√	X	X
S.T. Amin ve arkadaşları [18]	X	X	X	X	√	X	X
O. Tornea & M.E. Borda [52]	X	X	X	X	*	X	X
M. Sabry ve arkadaşları [53]	X	X	X	X	*	*	X
X. Wang & Q. Zhang [54]	X	X	X	X	X	X	X
A. Akanksha ve arkadaşları [55]	√	X	X	X	X	X	X
K. Ning [56]	X	X	X	X	√	X	X
S. Goyat ve S. Jain [25]	*	X	X	*	√	*	X
H.M. Mousa [23]	*	*	*	√	√	*	X

Çizelge 3.3. (devam) Mevcut DNA şifreleme algoritmalarının gereksinimler açısından karşılaştırma sonuçları.

R. Ahmed ve I. Muhammed [26]	*	√	X	*	√	X	X
M. Thangavel ve P. Varalakshmi [27]	*	X	√	√	*	*	X
Narendren ve arkadaşları [28]	*	X	X	*	√	X	X
S. Basu ve arkadaşları [30]	*	√	*	*	√	√	X
M. Tahir ve arkadaşları [31]	*	√	*	√	√	√	X
F. Thabit ve arkadaşları [33]	√	√	√	√	√	√	X
N. H. UbaidurRahman ve arkadaşları [6]	√	√	√	√	√	√	X
E. Şatır & O. Kendirli [36]	√	X	√	√	√	√	√
X- Asgari destek seviyesi göstergesi. √ - Kabul Edilebilir Destek Düzeyinin Göstergesi. * - Kısmen yerine getirme							

Önerilen çalışmada, öngörülen gereksinimlerin çoğunun karşılandığı açıktır. Önerilen çalışma düz metni ikiliye çevirdiğinden, tüm karakter setini DNA üzerinden kodlamak mümkündür. Ancak, dinamik kodlama mümkün değildir. DNA kodlaması ve Feistel Ağ entegreli DNA operatörleri kullanıldığından ve her döngüde sadece bir oturumda farklı döngü anahtarlar üretildiğinden, düz metnin her karakterini DNA dizisine kodlamak için benzersiz bir dizi yerine getirilmiştir. Bu aynı zamanda şifreleme sürecini dinamik hale getirir. Çünkü sadece 12 bitlik bloğu kodlamak için en az 40 döngü gerçekleştirilir. Her döngüde, uzunluğu 2 bit olan bir döngü anahtarımız var. Toplamda 12 bitlik blokların şifrelenmesi için anahtar uzunluğu en az 80 bittir ve her blok için anahtarlar da değiştirilir. Bunlar, şifrelemenin sağlamlığını ve dinamikliğini destekleyen ana özelliklerdir. Önerilen algoritma simüle edilmiş ve biyolojik süreç simülasyonu konusu yukarıda yer alan alt bölüm 3.1'de açıklandığı gibi laboratuvar deneyleri de gerçekleştirilmiştir.

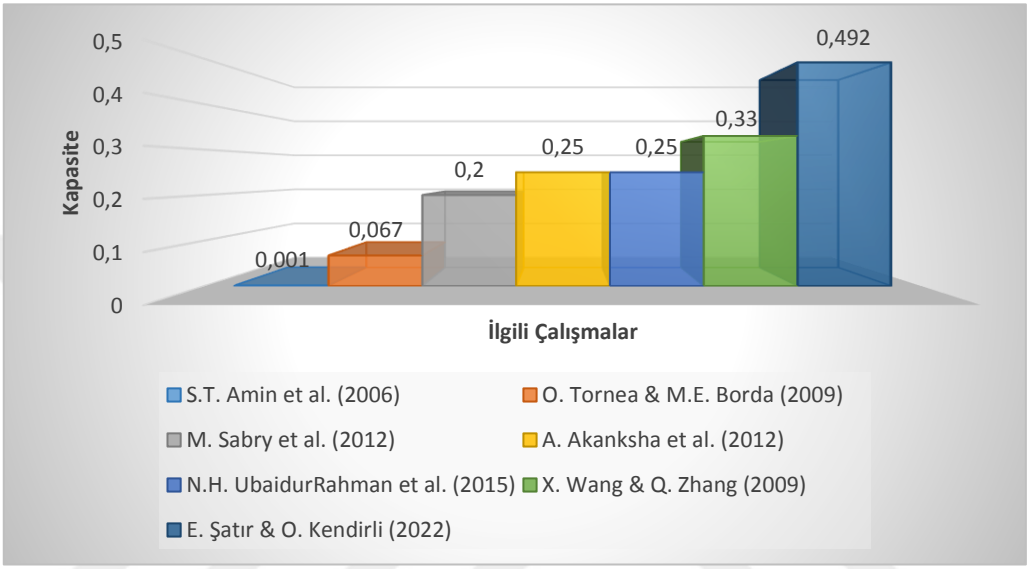
3.3. KAPASİTE ANALİZİ

Bu bölümde önerilen çalışmanın kapasite oranları ile literatürde ulaşılan mevcut çalışmalar karşılaştırılmaktadır. Kapasite, şifreli metnin sahip olduğu bilgi aktarımının yoğunluğudur. DNA şifrelemesinde, nükleotit başına gömülebilen veri oranı olarak ifade edilir. Çizelge 3.4, mevcut çalışmaların ve önerilen çalışmanın kapasite hesaplarının karşılaştırmalarını göstermektedir.

Çizelge 3.4. Kapasite hesabı karşılaştırması.

Yazarlar	Giriş Düz Metin Uzunluğu (bit)	Çıkış Şifreli Metin uzunluğu (bit)	Kapasite (Giriş/Çıkış)
S.T. Amin ve arkadaşları [18]	90300	88410189	0,001
O. Tornea ve M.E. Borda [52]	10	148	0,067
M. Sabry ve arkadaşları [53]	4	20	0,2
A. Akanksha ve arkadaşları [55]	1	4	0,25
N.H. UbaidurRahman ve arkadaşları [6]	4	16	0,25
X. Wang ve Q. Zhang [54]	3	9	0,33
M. Sohal ve S. Sharma [29]	40000	46640	0,85
E. Şatır ve O. Kendirli [36]	40000	40320	0,99

Çizelge 3.4’te açık metinlerin, şifreli metinlerin uzunlukları ve oranları yani kapasite değerleri verilmiştir. Burada düz metinlerin ve şifreli metinlerin uzunlukları bit olarak verilmiştir. Çizelge 3.4'te gösterildiği gibi, önerilen yöntem literatürdeki mevcut çalışmalardan daha verimli bir kapasite oranına sahiptir. Şekil 3.6'da ki grafikte de kapasite hesaplarının karşılaştırılması grafiksel olarak gösterilmiştir.



Şekil 3.6. Kapasite değerleri grafiği.

3.4. KABA KUVVET SALDIRI ANALİZİ

Kaba kuvvet algoritmasının değerlendirilmesi basittir, ancak parolanın şifresini çözmek için işlemek için çok sayıda adımı vardır. Bu amaçla saldırgan, parola adayları için her olası kombinasyonu denemeye başlar ve bunun şifresi çözülmüş bir dosyayla sonuçlanıp sonuçlanmadığını görmektedir [57].

Bu çalışmada, giriş 12 bit uzunluğunda en az bir bloğa sahip olduğundan anahtar uzunluğu en az 80 bittir. Anahtar, her biri 2 bitlik bir döngü anahtar içeren 40 döngüden elde edilir. Günümüz teknolojisi kullanılarak gerçekleştirilen bir kaba kuvvet saldırısı ile bunu kırmak neredeyse imkansızdır. Önerilen çalışma için kaba kuvvet saldırısı gerçekleştirildiğinde;

Günümüzde ortalama bir bilgisayar için ortalama Merkezi İşlem Birimi - Central Processing Unit (CPU) hızı 3×10^9 Hertz (Hz)'dir.

Bir yıl ise yaklaşık 3×10^7 (31557600) saniyedir,

Yıl cinsinden işlem zamanı Denklem 3.12'de ki formül ile hesaplanır ise

$$T = \frac{2^{anahtar\ uzunluğu}}{CPU\ hızı \times saniye} \tag{3.12}$$

Kaba kuvvet saldırısı ile sadece bir blok sonunda (40 döngü x 2 döngü anahtarı) üretilen 80 bit döngü anahtarını çözmek için gerekli işlemi süresi Denklem 3.13'te 12×10^6 yıl olarak bulunmaktadır:

$$T = \frac{2^{80}}{3\times10^9\times3\times10^7} \approx 12 \times 10^6 \text{ Yıl.} \tag{3.13}$$

Ayrıca önerilen yöntemin DNA'nın doğası gereği biyolojik bir süreci de vardır. Bu durum biyolojik olarak gerçekleştirilen ekstra bir güvenlik katmanı da oluşturmuş oluruz. Bu nedenle, klasik saldırı teknikleri önerilen şemayı kırmak için anlamsız kalmaktadır [36].

3.5. ANAHTAR UZAY ANALİZİ

Anahtar uzayı, şifreleme algoritması için kullanılan olası anahtar sayısını temsil eder. Daha yüksek bir anahtar uzayı, belirli bir zaman aralığında anahtar elde etme hesaplamalarını hesaplama açısından olanaksız hale getirdiğinden, algoritmayı kaba kuvvet saldırılarına karşı daha güvenli hale getirir. Yani, bir şifreleme şemasına ilişkin anahtar uzayı, olası anahtarların toplam sayısını belirtir [58]. Önerilen şemada şifreleme aşamasında yalnızca bir giriş verisi bloğu için en az 40 döngümüz var. Her döngüde, iki bit uzunluğunda bir döngü anahtar üretilir. Bu nedenle, her döngüde üretilen döngü anahtarı ($2^2 = 4$) bulmak için dört kombinasyonumuz var. Önerilen şemada 40 döngü olduğu için anahtarın toplam uzunluğu $40\times2=80$ bittir. Burada, önerilen şemadaki gizli anahtarı çıkarmak için 2^{80} olası kombinasyonumuz var. Ayrıca, bu kombinasyon sayısı düz metnin uzunluğuna bağlıdır. Düz metnin uzunluğu arttıkça anahtarın uzunluğu da artar [36].

3.6. ENTROPİ

Bilgi entropisi, bir kaynağın rastgeleliğini ölçmek için en önemli özelliktir. Bir mesaj kaynağının m bilgi entropisi *H(m)* , Denklem 3.14'te aşağıdaki formülle tahmin edilebilir:

$$H(m) = \sum_{i=0}^{L-1} p\ (m_i) log_2(p(m_i)) \quad (3.14)$$

Burada L, m'deki toplam sembol sayısıdır ve $p(m_i)$, m_i sembolünün ortaya çıkma olasılığıdır. Aynı olasılık entropi değerine sahip 256 sembol olduğunu varsayarsak, $H(m)=8$ [59]. Önerilen DNA kriptografi şemasında, kaynak mesaj dört sembolden oluşmaktadır; A, G, C, T [36]. Çizelge 3.5'te tahmin edilen entropi değerlerinin 2'ye yakın olduğu görülmektedir. 4 tane baz olasılığı (A, G, C, T) olduğunu düşündüğümüzde bu oldukça verimli bir orandır.

Çizelge 3.5. Önerilen şemanın tahmini entropi değerleri.

	Düz Metin			Şifreli Metin		
Uzunluk	*Frekans*	*Olasılık*	*Entropi H(x)*	*Frekans*	*Olasılık*	*Entropi H(x)*
100 baz	A = 17	0,17	1,933	A = 21	0,21	1,9385
	G = 19	0,19		G = 29	0,29	
	C = 36	0,36		C = 16	0,16	
	T = 28	0,28		T = 34	0,34	
1.000 baz	A = 215	0,215	1,9694	A = 120	0,12	1,906
	G = 186	0,186		G = 235	0,235	
	C = 290	0,290		C = 300	0,3	
	T = 309	0,309		T = 345	0,345	
10.000 baz	A = 3125	0,3125	1,9648	A = 2504	0,2504	1,9937
	G = 2964	0,2964		G = 2487	0,2487	
	C = 1982	0,1982		C = 2756	0,2756	
	T = 1929	0,1929		T = 2253	0,2253	
100.000 baz	A = 40009	0,40009	1,8805	A = 32155	0,32155	1,9238
	G = 15694	0,15694		G = 17569	0,17569	
	C = 28952	0,28952		C = 16541	0,16541	
	T = 15345	0,15345		T = 33735	0,33735	
Ortalama:			1,9369			1,9405

4. SONUÇLAR VE ÖNERİLER

Bu çalışmada, sistemlerin şifrelenmesi ve şifresinin çözülmesi için moleküler biyolojinin santral (merkezi) dogmasına (ilkesine) bağlı olarak, DNA taşıyıcı ortamı, DNA kodlaması ve DNA XOR işlemini Feistel ağ yapısı ile birleştirerek yeni bir şifreleme yaklaşımı türetilmiştir. Ayrıca, önerilen DNA şifreleme işlemi, tasarlanan biyoteknik donanıma hem dijital hem de biyolojik olarak entegre edilmiştir.

Yapılan deneyler DNA kriptolojisinin veri güvenliğinde yeni bir yöntem olma potansiyelinin yüksek olduğunu göstermiştir. Deneysel sonuçlar, önerilen çalışmanın kapasite, kaba kuvvet saldırısı, anahtar uzay ve entropi analizleri açısından verimli sonuçlara sahip olduğunu göstermiştir. Ayrıca önerilen yöntemin uygulanması laboratuvar deneyleri ile doğrulanmıştır.

Önerilen tez çalışması ile sıkıştırılıp şifrelenmiş olan DNA zinciri sentezlenerek bir plazmid içinde saklanmaktadır. Bu bakımdan çalışmanın veri gizleme sistemi olarak ele alınarak değerlendirilmesi de gerekmektedir.

Bir veri gizleme sisteminin en önemli gereksinimleri algılanamazlık, sağlamlık ve kapasite olarak bilinmektedir. Bu gereksinimlerin her biri, bir veri gizleme sistemindeki sihirli üçgenin köşelerini temsil ettiği düşünülmektedir ve bu çakışan gereksinimler arasında her zaman bir ödünleşim mevcuttur [60].

Kapasite, örten ortama gömülebilen verinin bit miktarını ifade etmektedir. Güvenlik, bir gözlemcinin saklı bilgiyi çıkarma becerisiyle ilgilidir. Sağlamlık ise saklı bilgiyi modifiye etme veya yok etmeye karşı direnme imkânı ile alakalıdır [61].

Bu tez çalışmasında elde edilen kapasite değerleri nükleotid başına saklanabilen bit miktarı oldukça verimlidir (bkz. 3.3. Kapasite analizi). Sağlamlık ise anahtar uzayı ve kaba kuvvet saldırı analizleri ile ölçülmüştür (bkz. 3.4. kaba kuvvet saldırı analizi, bkz. 3.5. Anahtar uzayı analizi) ilgili ölçümlerden alınan sonuçlar ele alınan her iki analizinde dikkate değer çıktılar ile desteklendiğini göstermektedir.

Her canlıda ki DNA dizilimi o türe özgüdür ve değişiklik göstermektedir. DNA zincirinin çözümlenebilmesi günümüzde halen ele alınan bir problemdir. Ayrıca canlılarda bulunan DNA zincirinde milyarlarca nükleotid bulunmaktadır. Bu nükleotidlerden anlamlı ve

anlamsız olanlarının ayrıştırılması da halen süregelen bir sorundur. Bu çalışmada sıkıştırılıp şifrelenmiş olan DNA zinciri sentezlenerek bir plazmid içinde saklanmaktadır. Plazmid içinde yer alan milyarlarca nükleotidden oluşan DNA zincirlerinin hangisinde saklanan bölümün yer aldığını bulmak bile gözlemci tarafından laboratuvar koşulları olmadan mümkün olamamaktadır. Bu koşulların olduğunu varsayarsak gözlemcinin milyarlarca DNA bazını ele alarak sayısız kombinasyon kurması gerekir. Tüm bu işlemler sadece saklanan kısmın bulunması içindir. DNA bazlarının diziliminde önceden belirli belli başlı kurallar yerine yalnızca minimal kısıtlar (aynı bazından çok sık tekrarlamaması vb.) yer aldığından, algılanamazlık bakımından bu tez çalışması en uygun ortamlardan birini sunmaktadır.

Ancak, bu çalışma kapsamında hala bazı engeller bulunmaktadır. Öncelikle bu çalışma simetrik şifreleme işlemine dayanmaktadır. Daha güvenli bir şema için asimetrik şifreleme tabanlı algoritmaların incelenerek DNA yapısına ve ortamına uygun bir modelin oluşturulması gerekmektedir.

DNA yapısında, bu teknoloji yepyeni bir alan olduğu için verilere rastgele erişim başka bir sorundur. Burada silikon ortamı yerine sentezlenen DNA'nın işlem süresini ve hızını düşündüğümüzde tüm süreci hızlandırmak için ekstra tekniklere ihtiyacımız var. Yukarıda bahsedildiği gibi, her karakterin dinamik olarak kodlanması, bir algoritmanın sağlamlığını artıran bir diğer gerekliliktir. Ancak tasarımın bu süreçte hafızayı işgal etmeden ve herhangi bir hız kaybına neden olmadan gerçekleştirilmesi gerekmektedir. Bu tez çalışmasındaki deneyler gerçekleştirilirken kapasite, kaba kuvvet saldırı, anahtar uzay analizi ve entropi alt başlıkları dikkate alınmıştır. Verilen tüm bu alt başlıkları deneysel parametreler olarak bir araya getiren başka bir çalışmaya da rastlanmamıştır. Bu nedenle tüm deneysel parametrelerin bu tez çalışmasında yapılan atıflar ile karşılaştırılarak desteklenmesi mümkün olmamıştır. Tüm bu konuların zaman faktörü ile birlikte gözlemlenip işlenerek ilerleyen çalışmalarda ele alınması planlanmaktadır.

5. KAYNAKLAR

[1] A. Kaundal ve A. Verma, "DNA based cryptography: a review", *International Journal of Information and Computation Technology*, c. 4, sayı 7, ss. 693–698, 2014.

[2] S. Namasudra, D. Devi, S. Kadry, R. Sundarasekar, ve A. Shanthini, "Towards DNA based data security in the cloud computing environment", *Computer Communications*, c. 151, ss. 539–547, 2020.

[3] A. Gehani, T. LaBean, ve J. Reif, "DNA-based cryptography", *Aspects of Molecular Computing*, ss. 167–188, 2003.

[4] V. Kollati ve S. Krishnan, "A symmetric multiple random keys (SMRK) Model Cryptographic Algorithm", *The International Journal of Innovative Research in Computer and Communication Engineering*, c. 3, ss. 10896–10903, 2020.

[5] A. Extance, "How DNA could store all the world's data", *Nature*, c. 537, sayı 7618, 2016.

[6] N. H. UbaidurRahman, C. Balamurugan, ve R. Mariappan, "A novel DNA computing based encryption and decryption algorithm", *Procedia Computer Science*, c. 46, ss. 463–475, 2015.

[7] G. Cui, L. Qin, Y. Wang, ve X. Zhang, "An encryption scheme using DNA technology", *2008 3rd International Conference on Bio-Inspired Computing: Theories and Applications, BICTA 2008*, sayı November, ss. 37–41, 2008.

[8] Y. Wang, P. Lei, H. Yang, ve H. Cao, "Security analysis on a color image encryption based on DNA encoding and chaos map", *Computers & Electrical Engineering*, c. 46, ss. 433–446, 2015.

[9] Ashutush Viramgama, "DNA data storage – synthetic DNA– future of storage". Erişim 20 Temmuz, 2022. https://ashutoshviramgama.com/dna-data-storage-synthetic-dna-future-of-storage.

[10] E. Mollick, "Establishing Moore's law", *IEEE Annals of the History of Computing.*, c. 28, sayı 3, ss. 62–75, 2006.

[11] J. Bornholt, R. Lopez, D. M. Carmean, L. Ceze, G. Seelig, ve K. Strauss, "A DNA-based archival storage system", *ACM SIGARCH Computer Architecture News*, c. 44, sayı 2, ss. 637–649, 2016.

[12] IDC, "The digitization of the world from edge to core", Erişim: 20 Temmuz, 2022. https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf.

[13] DELL, "Solid state drive (SSD) FAQ". Erişim: 6 Haziran, 2022. http://www.dell.com/downloads/global/products/pvaul/en/Solid-State-Drive-FAQ-us.pdf.

[14] S. Shrivastava ve R. Badlani, "Data storage in DNA", *International Journal of Electrical Energy*, c. 2, sayı 2, ss. 119–124, 2014.

[15] M. S. Neiman, "On the molecular memory systems and the directed mutations", *Radiotekhnika*, c. 6, ss. 1–8, 1965.

[16] J. Chen, "A DNA-based, biomolecular cryptography design". *Proceedings - IEEE International Symposium on Circuits and Systems*, c. 3, ss. 822–825, 2003.

[17] K. Tanaka, A. Okamoto, ve I. Saito, "Public-key system using DNA as a one-way function for key distribution", *BioSystems*, c. 81, sayı 1, ss. 25–29, 2005.

[18] S. T. Amin, M. Saeb, ve S. El-Gindi, "A DNA-based implementation of yaea encryption algorithm", *Proceedings of the 2nd IASTED International Conference on Computational Intelligence*, ss. 116–120, 2006.

[19] A. K. Verma, M. Dave, ve R. C. Joshi, "DNA cryptography: A novel paradigm for secure routing in mobile ad hoc networks (MANETs)", *Journal of Discrete Mathematical Sciences and Cryptography*, c. 11, sayı 4, ss. 393–404, 2008.

[20] Y. Zhang, B. Fu, ve X. Zhang, "DNA cryptography based on DNA fragment assembly", içinde *Proceedings - ICIDT 2012, 8th International Conference on Information Science and Digital Content Technology*, 2012, c. 1, ss. 179–182.

[21] O. Tornea ve M. E. Borda, "Security and complexity of a DNA-based cipher", *IEEE Roedunet. International Conference, 2013*. ss. 1-5, Sinaia, Romanya

[22] Monika ve S. Upadhyaya, "Secure communication using DNA cryptography with secure socket layer (SSL) protocol in wireless sensor networks", *Procedia Computer Science*, c. 70, ss. 808–813, 2015.

[23] H. M. Mousa, "DNA-genetic encryption technique", *International Journal of Computer Network and Information Security*, c. 8, sayı 7, ss. 1–9, 2016.

[24] Y. Q. Zhang, X. Y. Wang, J. Liu, ve Z. L. Chi, "An image encryption scheme based on the MLNCML system using DNA sequences", *Optics and Lasers in Engineering*, c. 82, ss. 95–103, 2016.

[25] S. Goyat ve S. Jain, "A secure cryptographic cloud communication using DNA cryptographic technique", *International Conference on Inventive Computation Technologies (ICICT)*, c. 3, ss. 1–8, 2016. Hindistan.

[26] R. Ahmed ve I. Mohammed, "Developing a new hybrid cipher algorithm using DNA and RC4", *International Journal of Advanced Computer Science and Applications*, c. 8, s. 71, 2017.

[27] M. Thangavel ve P. Varalakshmi, "Enhanced DNA and ElGamal cryptosystem for secure data storage and retrieval in cloud", *Cluster Computing*, c. 21, sayı 2, ss. 1411–1437, 2018.

[28] Narendren, Y. B. Yathish, ve C. Mohan, "A cryptosystem using two layers of security-DNA and RSA cryptography", *International Journal of Pure and Applied Mathematics*, c. 119-7, ss. 217-224, 2018.

[29] M. Sohal ve S. Sharma, "BDNA-A DNA inspired symmetric key cryptographic technique to secure cloud computing", *Journal of King Saud University - Computer and Information Sciences*, c. 34, ss. 1417-1425, 2022.

[30] S. Basu, M. Karuppiah, M. Nasipuri, A. K. Halder, ve N. Radhakrishnan, "Bio-inspired cryptosystem with DNA cryptography and neural networks", *Journal of Systems Architecture*, c. 94, ss. 24–31, 2019.

[31] M. Tahir, M. Sardaraz, Z. Mehmood, ve S. Muhammad, "CryptoGA: a cryptosystem based on genetic algorithm for cloud data security", *Cluster Computing*, c. 24, sayı 2, ss. 739–752, 2021.

[32] M. Indrasena Reddy, A. P. Siva Kumar, ve K. Subba Reddy, "A secured cryptographic system based on DNA and a hybrid key generation approach", *Biosystems*, c. 197, s. 104207, 2020.

[33] F. Thabit, S. Alhomdy, ve S. Jagtap, "A new data security algorithm for the cloud computing based on genetics techniques and logical-mathematical functions", *International Journal of Intelligent Networks*, c. 2, ss. 18–33, 2021.

[34] I. A. Aljazaery, H. T. S. ALRikabi, ve A. H. M. Alaidi, "Encryption of color image based on DNA strand and exponential factor", *International Journal of Online and Biomedical Engineering*, c. 18, sayı 3, ss. 101–113, 2022.

[35] A. K. Singh, K. Chatterjee, ve A. Singh, "An image security model based on chaos and DNA cryptography for IIoT images", *IEEE Transactions on Industrial Informatics*, s. 1-1, 2022.

[36] E. Şatir ve O. Kendirli, "A symmetric DNA encryption process with a biotechnical hardware", *Journal of King Saud University - Science*. c. 34, sayı 3, s. 101838, 2022.

[37] S. Al-Janabi ve A. Alkaim, "A novel optimization algorithm (Lion-AYAD) to find optimal DNA protein synthesis", *Egyptian Informatics Journal*, c. 23, sayı 2, ss. 271–290, 2022.

[38] E. Yoo, D. Choe, J. Shin, S. Cho, ve B.-K. Cho, "Mini review: Enzyme-based DNA synthesis and selective retrieval for data storage", *Computational and Structural Biotechnology Journal*, c. 19, ss. 2468–2476, 2021.

[39] J. M. Heather ve B. Chain, "The sequence of sequencers: The history of sequencing DNA", *Genomics*, c. 107, sayı 1, ss. 1–8, 2016.

[40] M. Zeki Kizmaz, İ. C. Paylan, ve S. Erkan, "DNA dizilemenin tarihsel gelişimi", *Gaziosmanpaşa Bilimsel Araştırma Dergisi*, ss. 47–53, 2017.

[41] D. Huo, D. fu Zhou, S. Yuan, S. Yi, L. Zhang, ve X. Zhou, "Image encryption using Exclusive-Or with DNA complementary rules and double random phase encoding", *Physics Letters A*, c. 383, sayı 9, ss. 915–922, 2019.

[42] D. Kumar ve S. Singh, "Secret data writing using DNA sequences", içinde *2011 International Conference on Emerging Trends in Networks and Computer Communications (ETNCC)*, ss. 402–405, 2011.

[43] B. Mondal ve T. Mandal, "A light weight secure image encryption scheme based on chaos & DNA computing", *J ournal of King Saud University - Computer and Information Sciences*, c. 29, sayı 4, ss. 499–504, 2017.

[44] T. Hagras, D. Salama, ve H. Youness, "Anti-attacks encryption algorithm based on DNA computing and data encryption standard", *Alexandria Engineering Journal*, c. 61, sayı 12, ss. 11651–11662, 2022.

[45] J. Li, J. Wang, ve X. Di, "Image encryption algorithm based on bit-level permutation and 'Feistel-like network' diffusion", *Multimedia Tools and Applications*, 2022.

[46] S. Hraoui ve A. JarJar, "Single Feistel lapse acting on reduced ASCII codes

followed by a genetic crossover", *SN Applied Sciences*, c. 4, sayı 4, s. 113, 2022.

[47] L. R. Knudsen ve M. Robshaw. *The Block Cipher Companion* (1. Basım). Springer Science & Business Media, 2011.

[48] M. Szaban ve F. Seredyński, "Designing cryptographically strong S-Boxes with use of 1D cellular automata", *Journal of Cellular Automata*, c. 6, ss. 91–104, 2011.

[49] G. Cui, C. Li, H. Li, ve X. Li, "DNA computing and its application to information security field", *Fifth International Conference on Natural Computation. ICNC 2009*, c. 6, sayı C, ss. 148–152, 2009.

[50] Q. Zhang, L. Guo, X. Xue, ve X. Wei, "An image encryption algorithm based on DNA sequence addition operation", *Proceedings, 4th International Conference on Bio-Inspired Computing: Theories and Applications, BICTA,* ss. 75–79, 2009. Beijing, Çin.

[51] S. Sadeg, M. Gougache, N. Mansouri, ve H. Drias, "An encryption algorithm inspired from DNA", *IEEE 2010 International Conference on Machine and Web Intelligence, ICMWI 2010*, ss. 344–349, 2010. Algiers, Cezayir.

[52] O. Tornea ve M. E. Borda, "DNA Cryptographic Algorithms", *International Conference on Advancements of Medicine and Health Care through Technology,* c. 26, ss. 223–226, 2009. Romanya.

[53] M. Sabry, M. Hashem, ve T. Nazmy, "Three reversible data encoding algorithms based on DNA and amino acids' Structure", *IJCA - International Journal of Computer Applications,* c. 54, sayı 8, ss. 24–30, 2012.

[54] X. Wang ve Q. Zhang, "DNA computing-based cryptography", *Proceedings, 4th International Conference on Bio-Inspired Computing: Theories and Applications, BICTA,* ss. 67–69, 2009. Beijing, Çin.

[55] A. Akanksha, A. Bhopale, J. Sharma, A. Meer Shizan, ve D. Gautam, "Implementation of DNA algorithm for secure voice communication", *International Journal of - Science and Research*, c. 3, sayı 6, ss. 1–5, 2012.

[56] K. Ning, "A pseudo DNA cryptography method", *09032693*, 2009, Erişim: http://arxiv.org/abs/0903.2693.

[57] T. Gautam ve A. Jain, "Analysis of brute force attack using TG - Dataset" 2015 *SAI Intelligent Systems Conference (IntelliSys),* ss. 984-988, 2015. Londra, İngiltere.

[58] S. Roy, U. Rawat, H. A. Sareen, ve S. K. Nayak, "IECA: An efficient IoT friendly image encryption technique using programmable cellular automata", *Journal of Ambient Intelligence and Humanized Computing (JAIHC)*, c. 11, sayı 11, ss. 5083–5102, 2020.

[59] A. Bakhshandeh ve Z. Eslami, "An authenticated image encryption scheme based on chaotic maps and memory cellular automata", *Optics and Lasers in Engineering*, c. 51, sayı 6, ss. 665–673, 2013.

[60] E. Şatır (2013), "Bilgi güvenliği için metin steganografisinde yeni bir yaklaşım", *Doktora Tezi*, Selçuk Üniversitesi, Konya, Türkiye.

[61] A. Gutub ve M. Fattani, "A novel arabic text steganography method using letter points and extensions", *International Journal of Computer, Electrical, Automation, Control and Information Engineering*, c. 1, ss. 502–505, 2007.

[62] A. M. Garipcan ve E. Erdem, “DESSB-TRNG: A novel true random number generator using data encryption standard substitution box as post-processing”, *Digital Signal Processing*, c. 123, s. 103455, 2022.

6. EKLER

6.1. EK 1: DES S – KUTULARI

S1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	14	4	13	1	2	15	11	8	3	10	6	12	5	9	0	7
1	0	15	7	4	14	2	13	10	3	06	12	11	9	5	3	8
2	4	1	14	8	13	6	2	11	15	12	9	7	3	10	5	0
3	15	12	8	2	4	9	1	7	05	11	3	14	10	0	6	13

S2	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	15	1	8	14	6	11	3	4	9	7	2	13	12	0	5	10
1	3	13	4	7	15	2	8	14	12	0	1	10	6	9	11	5
2	0	14	7	11	10	4	13	1	5	8	12	6	9	3	2	15
3	13	8	10	1	3	15	4	2	11	6	7	12	0	5	14	9

S3	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	10	0	9	14	6	3	15	5	1	13	12	7	11	4	2	8
1	13	7	0	9	3	4	6	10	2	8	5	14	12	11	15	1
2	13	6	4	9	8	15	3	0	11	1	2	12	5	10	14	7
3	1	10	13	0	6	9	8	7	4	15	14	3	11	5	2	12

S4	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	7	13	14	3	0	6	9	10	1	2	8	5	11	12	4	15
1	13	8	11	5	6	15	0	3	4	7	2	12	1	10	14	9
2	10	6	9	0	12	11	7	13	15	1	3	14	5	2	8	4
3	3	15	0	6	10	1	13	8	9	4	5	11	12	7	2	14

S5	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	2	12	4	1	7	10	11	6	8	5	3	15	13	0	14	9
1	14	11	2	12	4	7	13	1	5	0	15	10	3	9	8	6
2	4	2	1	11	10	13	7	8	15	9	12	5	6	3	0	14
3	11	8	12	7	1	14	2	13	6	15	0	9	10	4	5	3

S6	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	12	1	10	15	9	2	6	8	0	13	3	4	14	7	5	11
1	10	15	4	2	7	12	9	5	6	1	13	14	0	11	3	8
2	9	14	15	5	2	8	12	3	7	0	4	10	1	13	11	6
3	4	3	2	12	9	5	15	15	10	11	14	1	7	10	0	8

S7	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	4	11	2	14	15	0	8	13	3	12	9	7	5	10	6	1
1	13	0	11	7	4	9	1	10	14	3	5	12	2	15	8	6
2	1	4	11	13	12	3	7	14	10	15	6	8	0	5	9	2
3	6	11	13	8	1	4	10	7	9	5	0	15	14	2	3	12

S8	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	13	2	8	4	6	15	11	1	10	9	3	14	5	0	12	7
1	1	15	13	8	10	3	7	4	12	5	6	11	10	14	9	2
2	7	11	4	1	9	12	14	2	0	6	10	10	15	3	5	8
3	2	1	14	7	4	10	8	13	15	12	9	9	3	5	6	11

DES S – Kutuları [62]

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Oğuzhan KENDİRLİ

Yabancı Dili : İngilizce

ÖĞRENİM DURUMU

Derece	Alan	Okul/Üniversite	Mezuniyet Yılı
Doktora	Elektrik – Elektronik ve Bilgisayar Müh.	Düzce Üniversitesi	2022
Y. Lisans	Bilgisayar Müh.	Düzce Üniversitesi	2014
Lisans	Elektrik Elektronik Müh.	Düzce Üniversitesi	2021
Lisans	Elektronik Öğretmenliği	Selçuk Üniversitesi	2008

YAYINLAR

E. Şatir ve O. Kendirli, "A symmetric DNA encryption process with a biotechnical hardware", *Journal of King Saud University. - Science*, c. 34, sayı 3, s. 101838, 2022, doi: https://doi.org/10.1016/j.jksus.2022.101838

E. Şatır ve O. Kendirli, "A hybrid steganographic approach via web adresses", *İleri Teknoloji Bilimleri Dergisi*, c. 2, sayı 3, s. 53-60, Eylül 2013.