Performance Comparison of a DNA based Cryptosystem and Modern Cryptosystems

Paul Emil Ongoco
Patrick Angelo Roderno

Outline

- Introduction
- Objective
- Modern Techniques of Cryptography
  - Triple Data Encryption Algorithm (TDEA)
  - Advanced Encryption Scheme (AES)
  - Rivest-Shamir-Adleman (RSA)
- DNA, DNA Computing, and DNA Cryptography
- Methodology
- Results and Discussion
- Conclusion

Introduction

- Information Security as an issue in the modern society
- 1994 - Leonard Adleman introduced DNA Computing
- DNA Computing - concerned with the use of DNA molecules as tools to perform computations

Reference:
https://www.cs.usc.edu/people/faculty/tenured-tenure-track-faculty/adleman-leonard

Introduction

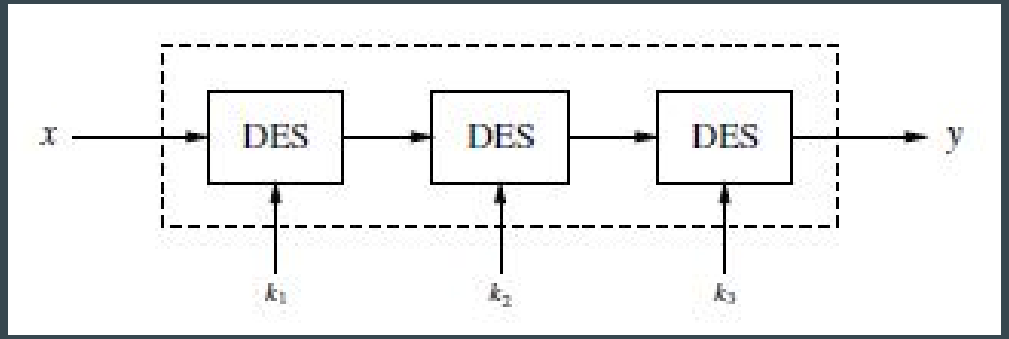
- DNA Cryptography - deals with exploring the potential of the DNA's massive parallelism and information density for cryptographic purposes
- Relatively a new field compared to the other fields of Computer Science
- Not yet applicable in real world scenarios

Objective

- Create a new and improved DNA-based cryptosystem based on an existing DNA-based cryptosystem to contribute to the realization of practical DNA-based cryptosystems
- Compare the developed cryptosystem to the techniques of the cryptography used today

Modern Techniques of Cryptography

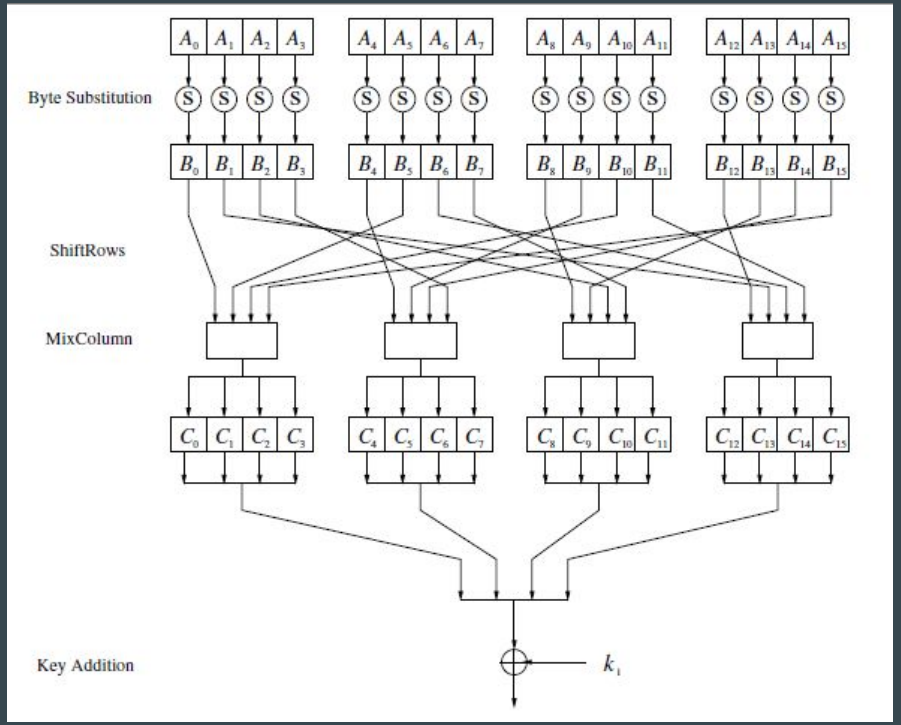
- Triple Data Encryption Algorithm (TDEA)
  - Data Encryption Scheme (DES) done three times
  - DES was broken in mid-1990s
  - Effective key length = 112 bits
  - No practical attacks to break TDEA as of today



Triple Data Encryption Algorithm (TDEA) Structure

Modern Techniques of Cryptography

- Advanced Encryption Scheme (AES)
  - Most used symmetric cipher
  - Encrypts 128 bits, key size = 128/192/256 bits
  - No practical attacks to break AES as of today



Modern Techniques of Cryptography

- Rivest-Shamir-Adleman (RSA)
  - Most used asymmetric cipher
  - Exponentiation in an integer ring
  - Considered secure due to its mathematical difficulty

$$y = e_{k_{pub}}(x) = x^e \bmod n$$

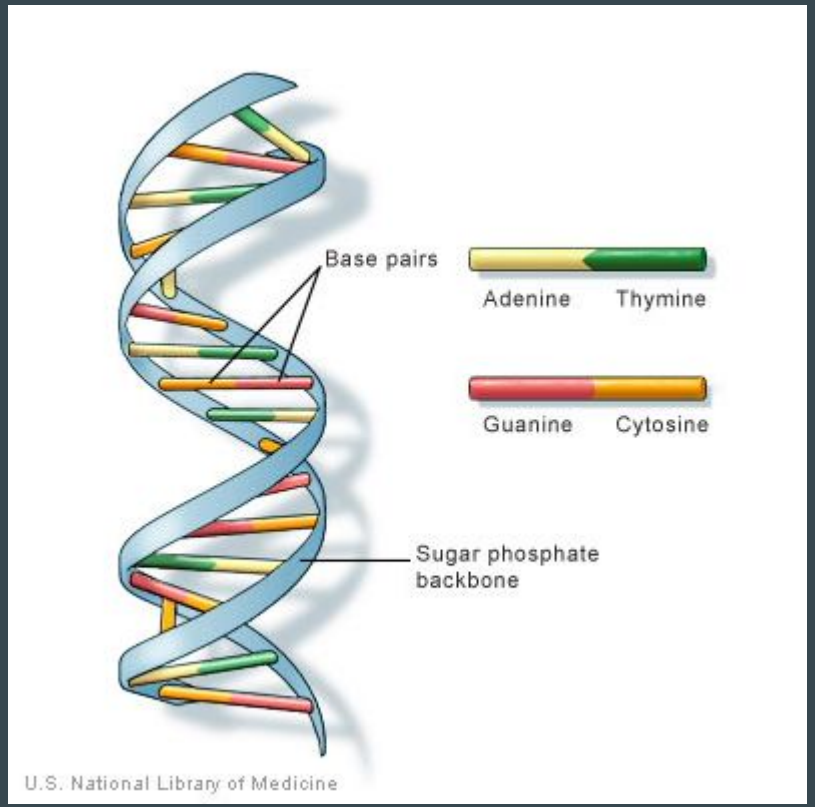
RSA Encryption

$$x = d_{k_{pr}}(y) = y^d \bmod n$$

RSA Decryption

Deoxyribonucleic Acid (DNA)

- Molecule that contains the instructions an organism needs to develop, live and reproduce
- Contains nitrogen bases that naturally pair with each other:
    - Adenine (A) to Thymine (T)
    - Guanine (G) to Cytosine (C)



DNA Computing

- Concerned with the use of DNA molecules as tools to perform computations
- The operations used in DNA computing are limited to the capabilities of molecular biology
  - Merge
  - Anneal
  - Melt
  - Separation by length
  - Separation by sequence
  - Copying/Amplification
  - Detect
- Massive parallelism
- High information density

DNA Cryptography

- Recent efforts by researchers
  - Leier, et al (2000) - two approaches in cryptography that makes use of DNA binary strands
    - First approach - uses the DNA binary strands in steganography
    - Second approach - uses the method of graphical subtraction and binary gel images to constitute a molecular checksum
  - Chen (2003) - cryptosystem that uses carbon nanotubes to transform data between DNA and conventional binary storage media
  - Thiruthudavoss (2012) - DNA-based cryptosystem that uses one-time pads (OTP) for encryption
  - Fasila, et al (2014) - a new hybrid cryptosystem that uses Red, Green, and Blue (RGB) colors to further improve the security of data

Methodology

- Simulate a new DNA-based cryptosystem based from the works of Thiruthuvadoss
- Compare the new cryptosystem with the original as well as some modern cryptosystems:
  - Triple Data Encryption Algorithm (TDEA)
  - Advanced Encryption Scheme (AES)
  - Rivest-Shamir-Adleman (RSA)
- Metrics used in comparing:
  - Key Length
  - Attack Steps
  - Rounds
  - Running Time
  - Time Complexity
  - Algorithm Strength

Methodology

- New DNA-based cryptosystem
  - Essentially added steganography over the original system after encryption.
    - Store the first and last 10 nitrogenous bases of the ciphertext (primers).
    - Mix the ciphertext with randomly generated DNA strands to form one long strand of approximately 3,000,000 nitrogenous bases.
  - Simulate Polymerase Chain Reaction (PCR) to decrypt.
    - Actual PCR process not necessary (very slow).
    - Search for DNA segment that starts and ends with the primers
    - Decrypt using OTP.

Methodology

- Tests were run on three cases
  - Single sentence
  - Sample SMS
  - Sample email
- Specifications of the computer used
  - Processor: Intel i5-6500 4-core @ 3.2 GHz
  - RAM: 2x4 GB DDR4 2133
  - OS: Windows 10 64-bit

Results and Discussion

n - length of string in bits

Metrics	Thiruthuvadoss	New System	TDEA	AES	RSA
Key Length	20n	20n + 40	112	128/192/256	1024/2048
Attack Steps	2^{20n}	$2^{20n} * 2^{40}$	2^{112}	$2^{128}/2^{192}/2^{256}$	$2^{1048}/2^{2048}$
Rounds	n/a	n/a	48	10/12/14	n/a
Time Complexity	O(n)	$O(n^2)$	O(n)	O(n)	O(n)
Algorithm Strength	US	US	CS	CS	CCS

Results and Discussion: Running Time, Encryption

Test Case	Thiruthuvadoss	New System	TDEA	AES	RSA
Single Sentence	14.07 ms	2324 ms (2788.61 ms*)	8.27 ms	2.64 ms	13.17 ms
140-character message	30.16 ms	2719.65 ms (7695.05 ms*)	10.42 ms	2.71 ms	24.52 ms
Sample email	290.94 ms	2669.36 ms (135114.05 ms*)	11.25 ms	3.44 ms	264.99 ms

* without iteration limit

Results and Discussion: Running Time, Decryption

Test Cases	Thiruthuvadoss	New System	TDEA	AES	RSA
Single Sentence	1.55 ms	78 ms (66.89 ms*)	1.77 ms	1.21 ms	2.2 ms
140-character message	5.53 ms	171.89 ms (77.919 ms*)	3.43 ms	2 ms	4.29 ms
Sample email	92.04 ms	170.76 ms (3375 ms*)	3.18 ms	2.27 ms	7.446 ms

* without iteration limit

Conclusion

- The new DNA-based cryptosystem provides an additional layer of security through a steganographic approach but proves to be impractical because of its time complexity of $O(n^2)$ which is mainly due to the generation of random strands of DNA used to hide the ciphertext.