International Journal of Advanced Computer Technology (IJACT)

ISSN:2319-7900

Encryption using Venn-Diagrams and Graph

Parijit Kedia, Vellore Institute of Technology, Tamil Nadu, India
Sumeet Agrawal, Vellore Institute of Technology, Tamil Nadu, India

Abstract

There are various algorithms that exist for text encryption that are used in various fields like password encryption, exchange of information. In this paper we propose a new algorithm, where in the data is converted to a format consisting of numbers and letters using basic mathematical concepts like Venn-diagram and advanced concepts like graph theory.

Keywords—graph, cryptography, venn-digram, ascii ,binary

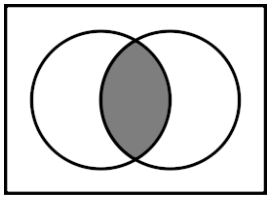
Introduction

Cryptography or Cryptology is the technique of hiding messages in a secure format so that the third party cannot identify what the message is being transmitted. Cryptography involves various techniques from different areas of research like computer science, mathematics and electronics. It basically converts the message to a nonsensical format. The receiver knows the method of decoding the pattern which he then applies to cipher text to get back the plaintext which the sender intended to send him. Cryptography has been used since World War 1 where spies secretly used to send messages to the host country so that the enemy couldn't understand the real meaning of the message. The ciphered text is designed to be long so that its computational power is very large and not possible for a computer to solve it quickly. These are secured algorithms.

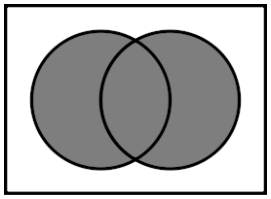
Mathematical concepts used

A. Venn Diagrams

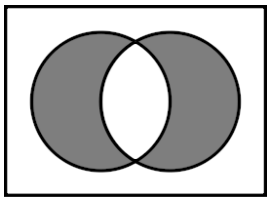
A Venn-Diagram or set diagram shows logical relations among various sets. It was made by John Venn in 1880. The diagrams tell us about how sets are related by showing relations using probability, statistics. The Venn-Diagram consists of various circles drawn in plane overlapping to give a particular region. Each circle represents a class in real world where it is compared to another class or set to determine the relationship between the two sets. The overlapping region represents intersection that is the common of both the sets.



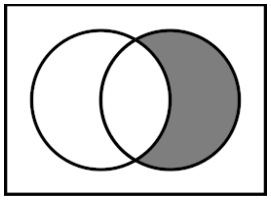
Intersection of 2 sets represented as A ⋂ B



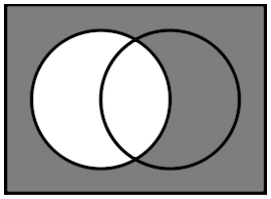
Union of 2 sets represented as A ⋃ B



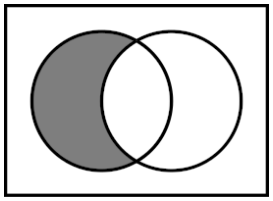
Symmetric Difference of 2 Sets represented as A Δ B



Set Difference of 2 sets represented as B - A



Absolute complement represented as A^c



Set difference of 2 sets represented as A -B

Figure – 1.1, 1.2, 1.3, 1.4, 1.5, 1.6 (top to bottom)

B. Graph Theory

A Graph $G = (V, E)$ consists of a set of objects where $V = \{v_1, v_2, v_3, \dots, v_n\}$ represents vertices and

ISSN:2319-7900

$E = \{e_1, e_2, e_3, \ldots, e_m\}$ represents edges such that edge e_k is identified as an unordered pair of vertices$(v_i v_j)$.

An edge is a connection between two nodes or vertices indicating that a connection is established between the two vertices.

Graph Theory has a wide variety of applications like Electrical network problems. The way the network is connected shows similar to the way a graph is connected. The edges represent resistors and currents in the wire of the electrical network.

They can also be used in seating problems where arrangement of persons determines where exactly a person is located in the graph and how it is connected to other member of the circular meeting arrangement.

In the given figure (2) below e_7 is a self-loop since it's starting and ending vertex is the same.

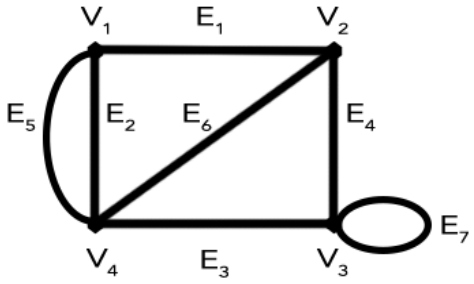


Figure - 2

C. 7-Segment Display

It is a form of electronic display device for displaying numbers in digital format. They are used in digital clocks, basic calculators

They are arranged as shown in the figure (3) below.

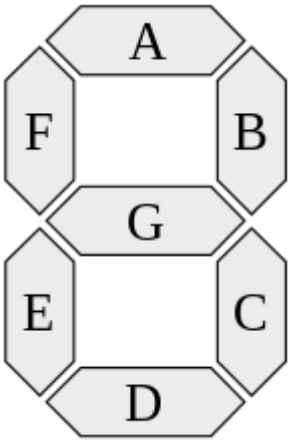


Figure - 3

The following table (1) shows which sides are on each indicating which number will be displayed.

Hexadecimal encodings for displaying the digits 0 to F

Digit	gfedcba	abcdefg	a	b	c	d	e	f	g
0	0×3F	0×7E	on	on	on	on	on	on	off
1	0×06	0×30	off	on	on	off	off	off	off
2	0×5B	0×6D	on	on	off	on	on	off	on
3	0×4F	0×79	on	on	on	on	off	off	on
4	0×66	0×33	off	on	on	off	off	on	on
5	0×6D	0×5B	on	off	on	on	off	on	on
6	0×7D	0×5F	on	off	on	on	on	on	on
7	0×07	0×70	on	on	on	off	off	on	off
8	0×7F	0×7F	on	on	on	on	on	on	on
9	0×6F	0×7B	on	on	on	on	off	on	on
A	0×77	0×77	on	on	on	off	on	on	on
b	0×7C	0×1F	off	off	on	on	on	on	on
C	0×39	0×4E	on	off	off	on	on	on	off
d	0×5E	0×3D	off	on	on	on	on	off	on
E	0×79	0×4F	on	off	off	on	on	on	on
F	0×71	0×47	on	off	off	off	on	on	on

Table - 1

Algorithm

Now we discuss our algorithm. Take any input as password and find each letters equivalent ASCII code. Then each ASCII code is converted to its binary representation. The 8-bit binary number is then split into 2 parts consisting of 4 bits each and then converted to the hexadecimal form where now graph theory and Venn-diagrams are applied.

ISSN:2319-7900

The numbers so obtained are in the 7-segment format. We can now know which edges are to be selected to form the number on the 7-segment display. Two such numbers are superimposed and the edge difference is calculated. The reason behind this is that the edge difference so obtained will result in a new graph formation. This edge difference will be unique and the resulting graph will tell us what numbers were used to form the new graph. The set difference will result in a key that will keep on changing after each operation to the passphrase which will generate a new cipher text. So the cipher text is converted to non-sensible number format plus the set difference.

THE ALGORITHM

1. $X :=$ *array containing letters of plaintext.*
2. $Y :=$ *empty array for storing coded value of the letters.*
3. $Z :=$ *empty array for storing arithmetic difference between two hexadecimal digits*
4. **for each** letter in X:
5. $i := 1$
6. Convert ASCII code of the letter to hexadecimal.
7. **If** both digits of hexadecimal are equal, **then**
8. $Y[i] :=$ "1234"
9. $Z[i] := 0$
10. **Else**
11. $A :=$ *set of edges corresponding to the less significant digit (LSD).*
12. $B :=$ *set of edges corresponding to the more significant digit (MSD).*
13. $D := A - B$
14. **If** $D = \emptyset$, **then**
15. $D := B - A$
16. $Y[i] :=$ *ascending vertex sequence that makes up D*
17. $Z[i] := abs(MSD - LSD)$
18. $i := i + 1$
19. $n :=$ *length of X*
20. $mid := floor(\frac{n}{2})$
21. $key :=$ *numerical value of Z*
22. $Y' :=$ *array for storing second level of encoded value of letters*
23. $Y'[mid] := Y[mid] \oplus key$
24. **while** all letters have not been XOR-d with key, **do**
25. $j := 1$
26. $Y'[mid - (-1)^j \cdot j] := Y[mid - (-1)^j \cdot j] \oplus Y'[mid]$
27. $mid := mid - (-1)^j \cdot j$
28. $j := j + 1$
29. Concatenate all the elements of Y followed by the key to give us the final encrypted cipher text.
30. The final cipher text is then converted to hexadecimal format.

Proposed Approach

We are taking "JET" as a password which is going to be encrypted by the above proposed scheme.

Step 1: ASCII of each letter of the password is
J = 74
E = 69
T = 84
Step 2: Binary equivalent of each letter is
J = 01001010
E = 01000101
T = 01010100
Step 3: Grouping each binary number into 4 and then converting to hexadecimal we get
J = 4A
E = 45
T = 54

Step 4: Converting each number into digital format

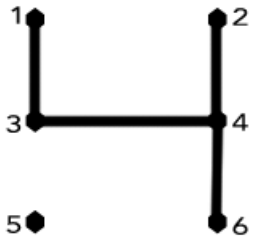


Figure - 4

The graph connections of 4 are 1-3-4-6, 2-4

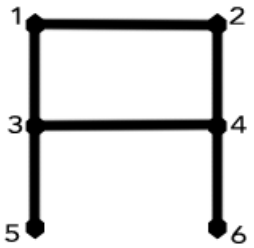


Figure - 5

The graph connections of A are 1-2-4-6, 1-3-4 and then applying Venn diagram concepts we find A – B which is

ISSN:2319-7900

null so we go for B – A and get graph connections to be 1-2,3-5 which may or may not represent a disconnected graph.

Figure - 6

Similarly for E we get 2-4
And for T we get 1-2, 5-6
Also store each words difference like for J we get 6 i.e.
|4 - A| = 6

Step 5: the code for now is 1235241256611

Step 6: the key is 611
Step 7: applying key with middle element i.e. XOR-ing we get
24 Θ 611 = 635
Step 8: The resulting number is XOR-d with next symbol i.e. $(k + i)^{th}$ and then with $(k - i)^{th}$ position with value of 'i' changing as per the formula given in the algorithm and the process is repeated.

i.e. 635 Θ 1256 = 1683
and 1683 Θ 1235 = 576

So the final cipher text we get is 5766351683611.

So converting the final cipher text into hexadecimal we get 53e9558001b

Finally the keyword "JET" is encrypted as 53e9558001b.

Similarly, when we take other words even with 80% of the letter similar, no words will have similar cipher text because the key operation performed on them will always generate a new cipher text. So no cipher words get repeated.

Since the algorithm uses ASCII numbers of the words for conversion and since the uppercase and lowercase letters have different ASCII, so the final cipher text of the same password for both uppercase and lowercase will be totally different. Also we can use all the special characters in the encryption because of the useful nature of ASCII.

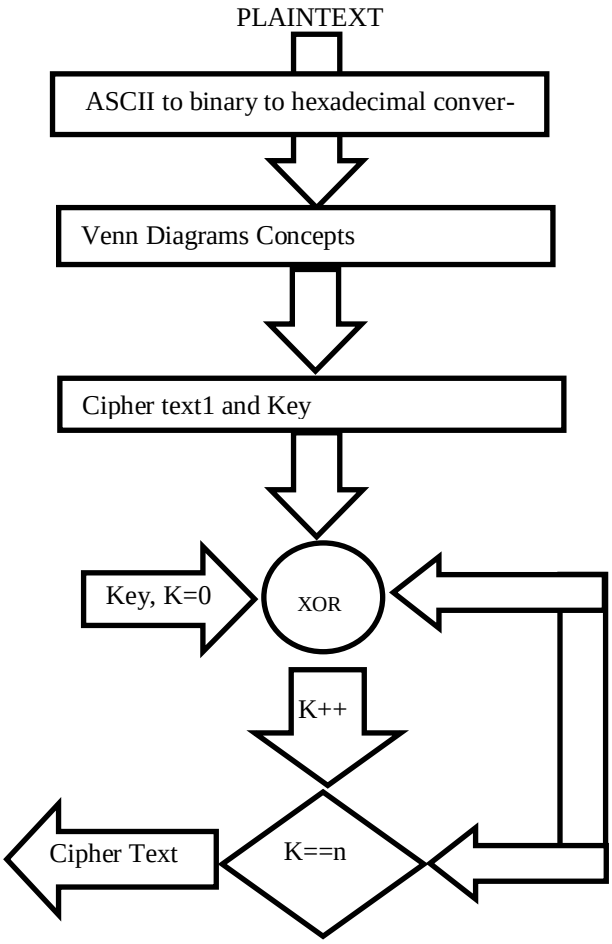


Figure – 7

FLOW CHART OF THE ALGORITHM

Optimization

The code can be optimized using paging techniques along with Dynamic Programming.

A. Paging Technique

It is a memory management technique where computer uses data for its main memory from secondary storage. They retrieve data in a format called pages. It allows physical ad-

ISSN:2319-7900

dress space to be non-contiguous. It is also used in virtual memory implementation. It is used when the data that is to be used is not present in RAM and so it accesses its secondary storage for retrieval. If the data is not present in RAM, then it must

1. Locate data in secondary storage
2. Obtain the page in RAM
3. Loading of data
4. Updating of page table

In our algorithm there will be storage of passwords that the user enters. This database will keep a count of the number of times the same password is occurring. The new password will keep on stacking up the database and the passwords that are not occurring or not commonly used are removed from the database. So, if the word exists it is searched from the database and returned with the ciphered text thus saving encryption time.

B. Genetic Algorithm

Genetic Algorithm or GA is one of the best optimization heuristics techniques to optimize computational complexity. It uses frequency analysis to find the word that has occurred the most number of times thereby reducing the number of times the cipher text of the word is calculated from scratch. It conducts a direct random search in the database for the word that is already converted. They may also produce optimum keys for conversion to cipher text.

C. Particle Swarm Optimization

It is used to find the candidate solution. The Algorithm considers particles moving around in a search space with given position and velocity. The particle is influenced by local best position and is moved towards the solution depending on the best known position which are closer to the final solution. Thus, the swarm moves towards the best solution. This method is computationally faster as compared to the above mentioned methods.

Applications

The Applications of our proposed scheme of encryption lies in

1. Secrecy in Transmission where the man in the middle attack is not possible.
2. Secrecy in Storage where user gives the key to open the system like in computer
3. Integrity in Transmission which ensures information is safe and does not cause erroneous action to take place.
4. Integrity in Storage which involves cryptographic values to be compared to expectations.
5. Authentication of Identity is done when user logs into his particular account or system.
6. Credential Systems which involves verification of user electronically. The most common example being the smart card system.
7. Electronic Signatures involves physical components where the physical presence of a person is recorded as a cryptographic value which will later be compared and so becomes difficult to crack.
8. Electronic Cash.
9. Threshold Systems involves distributing the keys into different parts so that different equations result into different cipher texts.
10. Systems Using Changing Keys involves changing the key over a period of time so that the hacker doesn't know what will be the password because the cipher texts will always keep on changing whenever the hacker makes a move.
11. Hardware to support Cryptography involves building super computers which can computationally convert the key like smart card, debit card.

References

[1] wikipedia.org

[2] Graph Theory with Application to Engineering and Computer Science by Narsing Deo, Eastern Economy Edition.

[3] Understanding Cryptography, A Textbook for Students and Practitioners by Christof Paar and Jan Pelzl.

[4] Graph Theory by Frank Harary, Addison-wesley publishing company

[5] Cogweels of the mind - The Story of Venn Diagrams by A.W.F. Edwards, John Hopkins University press.

[6] https://www.coursera.org/course/crypto

[7] Making Venn Diagrams by Therese Harasymiw, Gareth Stevens Publishing

[8] I. Mantin. Predicting and Distinguishing Attacks on RC4 Keystream Generator. Eurocrypt Vol. 3494 of LNCS, pp. 491-506, Springer-Verlag, 2005.

[9] G. Gong, K. C. Gupta, M. Hell, andY. Nawaz, Towards a General RC4-like Keystream Generator, SKLOIS Conference on Information Security and Cryptology (CICS05), December15-17, Beijing, China. Springer-verlag, 2006.

[10] eSTREAM - The ECRYPT Stream Cipher Project, http://www.ecrypt.eu.org/stream/

ISSN:2319-7900

[11] Y. Nawaz and G. Gong, WG: A family of stream ciphers with designed randomness properties, Information Sciences, Vol. 178, No. 7, April 1, 2008, pp. 1903-1916

[12] National Bureau of Standards, Data Encryption Standard, FIPS Publication 46, U.S. Department of Commerce, 1977

[13] National Institute of Standards and Technology, Advanced Encryption Standard, FIPS-197, http://csrc.nist.gov/archive/aes/index.html, 2000

[14] A Handbook of Applied Cryptography by Alfred J.Menezes, Paul C.Van Oorschot and Scott A. Vanstone, CRC Press Series on Discrete Mathematics and its Applications

[15] A course in Number Theory and Cryptography, Neal Koblitz, Springer 1987

[16] Modern Cryptography, Probabilistic Proofs and Pseudorandomness, Oded Goldreich, Spring-Verlag 1998

Biographies

PARIJIT KEDIA pursuing B.TECH degree in Computer Science and Engineering from Vellore Institute of Technology, Vellore, Tamil Nadu.

SUMEET AGRAWAL pursuing B.TECH degree in Computer Science and Engineering from Vellore Institute of Technology, Vellore, Tamil Nadu.