A New Coding Technique and Analysis of Trees

D. A. Angel Sherin, V. Maheswari

Abstract: A message is encoded using one-time pad Cipher and Huffman coding with certain algorithm. Encoding process is done using a symmetric key known to sender and receiver. Then we get the encoded message as Ciphertext with a binary tree. Using the binary tree we form a prefix code of Huffman coding and we define a new labeling function of edge and vertex labeling. In this paper, we discuss the two methods of encoding algorithm and investigate Tree, Rooted trees, properties, theorem and Median of Huffman binary tree.

Keywords: One-time pad Cipher, Huffman coding, Trees, Rooted trees, Huffman Vertex, Edge Labeling and Median. 2010 Mathematical subject classification Number: 05C78.

I. INTRODUCTION

One-time pad is first explained by Frank Miller in 1882, and it is again was re-invented in 1917. Frank Miller used this system for securing telegraphy. In the year 1917, Gilbert Vernam used the XOR method to message of an one-time pad. He got one-time pad method from his Vernam cipher. Vernam's system consists of operative key tape, which was again used when the loop form a full cycle. Joseph Mauborgne introduces one-time pad Cipher to the world with random key tapes to crypt the message.

To increase security, the cryptographers print the one-time pads message onto sheets of highly flammable nitrocellulose, so that they could easily be burned after use. The one-time pad Cipher is an encryption technique which cannot be cracked, without knowing key of same size as the message. In this technique we pair the plaintext with a secret key. Then, we get

a pad of character using modular addition. One-time pad cipher belongs to a type of Vignere cipher.

A. Huffman Coding

David A. Huffman tries to solve an assignment using binary code in 1957. But he could not find the accuracy of any codes. One day he got the idea of using a frequency-unorder binary tree and quickly proved the efficient of this method. To continue his project Huffman joined with Information theory analyst Fano and Claude Shannon to develop more code with binary trees. They confirmed the optimality of structured tree which is drawn from the bottom. Huffman coding is a famous Greedy Algorithm and used for the lossless compression of data. It applies to find variable length encoding, frequency of smallest and largest code.

II. LITERATURE REVIEW

M.Senthil Kumar, V.Mathivanan [11] introduces a new concept of analyzing data using Huffman coding and Arthimetic coding. Alyssa Gottshall, Justin Kahn [9], gave detailed project report regarding the Radio number of Biregular paths. K.Sunitha, Dr.C.David Raj and Dr.A.Subramanian [8] talked about the Radio labeling of Hurdle graph and Biregular rooted trees. Inspired by these work on trees we found a new methods of coding by Huffman algorithm.

III. ILLUSTRATION (ONE-TIME PAD CIPHER)

A. Key features of One-time pad Cipher

1. The key is an unbreakable Cipher
2. The key is exactly same as the length of message which is encrypted
3. The key is made up of random symbols
4. The key used once will not be used again

Due to this, encrypted message will not be accessible to attack for a cryptanalyst.

B. Algorithm for One-time pad

1. The message will be given
2. Key with a same string size will be given
3. Addition of each message word with key words
4. Sum of Key and plain text will be further calculated with modulo 10/26/2
5. A new set of Ciphertext will be developed

C. Encryption

Message: We are Happy

Key: MNKLDGHFJR

Revised Manuscript Received on December 5, 2019.

D. A. Angel Sherin*, Research Scholar, Department of Mathematics, Vels Institute of Science, Technology & Advanced Studies (VISTAS), Pallavaram, Chennai-117. India. Email: d.a.angelsherin@gmail.com.

Dr. V. Maheswari, Associate Professor, Department of Mathematics, Vels Institute of Science, Technology & Advanced Studies (VISTAS), Pallavaram, Chennai-117. India.Email: maheswari.sbs@velsuniv.ac.in

Numbering of Alphabet

A	B	C	D	E	F	G	H	I	J	K	L	M
1	2	3	4	5	6	7	8	9	10	11	12	13
N	O	P	Q	R	S	T	U	V	W	X	Y	Z
14	15	16	17	18	19	20	21	22	23	24	25	26

Ciphertext:

W = 23+13=36(mod26)=10
E = 5+14=19(mod26)=19
A = 1+11=12(mod26)=12
R = 18+12=30(mod26)=4
E = 5+4=9(mod26)=9
H = 8+7=15(mod26)=15
A = 1+8=9(mod26)=9
P = 16+6=22(mod26)=22
P = 16+10=26(mod26)=0
Y = 25+18=43(mod26)=17

Ciphertext: JSLDIOIVϕQ

D. Decryption

To decrypt a letter, user subtract key letter from Ciphertext. While subtracting the key if we get negative integer then add 26 and find modular 26. At last we get an original message.

IV. ILLUSTRATION (HUFFMAN CODING)

A. Prefix Rule

Huffman Coding is analyzed to get new prefix rule. This rule prevents the uncertainty while decoding. It protects that the code assigned to any character by non repetition. There are two major steps in Huffman Coding

1. Creating a Huffman Tree from the message
2. By traversing the Huffman tree we get an assigned code to each character

B. Algorithm for Huffman tree

1. Create vertices for each character of the Message and calculate the occurring frequency of that character
2. Arrange all the vertices in increasing order of their frequency value
3. Consider the first two vertices of minimum frequency
4. Make the first vertex as a left child and the other vertex as a right child of the newly created vertex
5. Keep repeating the step 4 until all the characters are assigned to form a single tree
6. The tree finally obtained is the desired Huffman tree
7. We assign label to all the edges of the Huffman tree
8. Assign weight '0' to the left edges and label'1' to the right edges

C. Steps for assigning Labels

1. If we assign label '0' to the left edges, then assign label '1' to the right edges
2. If we assign label '1' to the left edges, then assign label '0' to the right edges
3. We can follow any one of the above two steps
4. For decoding the message we need to follow the same convention used in encoding

D. Message:

Fourier series plays a vital role in finding periodic signals in network, propulsion of rocket and light house.

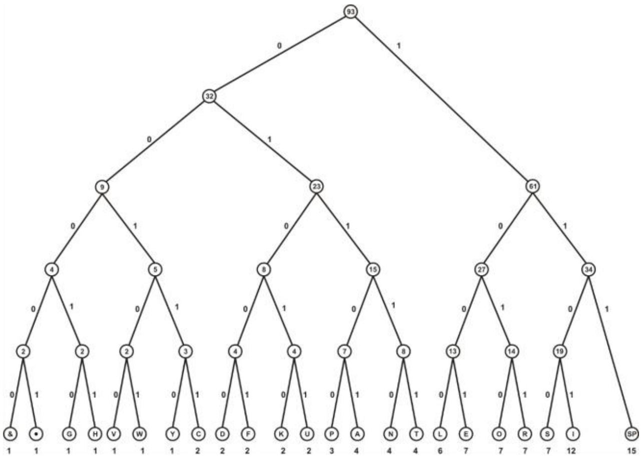

Fig. 1

E. Binary code

& 0	V 100	D 1000	P 1100
L 1000	S 1100	1	W 101
F 1001	A 1101	E 1001	I 1101
G 10	Y 110	K 1010	N 1110
O 1010	SP 111	H 11	C 111
U 1011	T 1111	R 1011	

F. Encoded message

0100110100101110111101100110111100100110111101100111000110010000110100110110001101001001101011110110110001011101010001001110101110010011101011100100011010111000010011001001101111011010010001101001111100110100010011100110110001100110101110011101001011110010110101011010100000101100101110100110001011100011001101101001110101001001101110100011101010100101111000001000110100010000110111100011101001011110010 01.

V. ANALYZED TREE

From the Fig. 1 we can trace the Tree, Rooted trees, Labeling, properties, theorem and Median

***A. Definition** (Tree):*

A graph is said to be a tree, when it contains path of minimally connected graph. It is a special case of graph.

***B. Definition** (Rooted trees):*

A tree which has one distinguished vertex is called a rooted tree.

***C. Definition** (Huffman Edge Labeling) :*

Each edge is labeled with a positive integer value of 0 and 1 from the above vertices.

***D. Definition** (Huffman Vertex Labeling):*

Let T be a tree with a vertex set defined by $f : V(T) \to \{x + y\}$ where x and y are sum vertices got from numbering of alphabets from below.

E. Properties

If a tree contains n vertices, then it has n-1 edges.

To prove this property we consider a tree from the Figure 1. In this diagram we have totally 45 vertices and

45-1 =44 edges.

F. Theorem

Let T be a graph of x vertices and y edges. Then

1. T is a tree (connected and acyclic).
2. T is connected and H = y + 1.
3. T is acyclic and if any two nonadjacent points are joined by a line, the resulting graph has exactly one cycle.

Proof:

We intended to prove this theorem from the Huffman binary tree.

We need to show that 1 ⇒ 2 ⇒ 3

1. From the Huffman binary tree any two vertices are joined by path. We state that path to be unique. Every path in the tree does not form a closed path. Therefore we can say it is connected graph of acyclic.
2. If we remove any one of the edge from Huffman binary tree then T will be break into two components. Let the sizes of two components be x_1 and x_2, with $x_1 + x_2 = x$. By the induction hypothesis, $x_1 = y_1 + 1$ and $x_2 = y_2 + 1$; but then $x = x_1 + x_2 = (y_1 + 1) + (y_2 + 1) = (y_1 + y_2) + 2 = y - 1 + 2 =$ $y + 1$.
3. If we join any two nonadjacent points by a line, the resulting graph contains only one cycle. Since T does not have cycle it forms a tree. Suppose if we join the vertices 23 and 61 in Fig. 1 we get a one unique cycle.

VI. MEASURE OF CENTRAL

***A. Definition** (Sum) :*

The sum of a vertex is defined by S(V)= $d(x_1, y_1) + d(x_2, y_2) + \ldots\ldots. d(x_n, y_n)$.

***B. Definition** (Centroid) :*

The centroid of a tree is the set of vertex minimum label.

***C. Definition** (Median) :*

A vertex with a minimum sum is called median vertex.

D. Theorem

A vertex is a median vertex if and only if it is a centroid vertex.

Proof:

Let us first calculate the sum of vertices.

S(V)= $d(x_1, y_1) + d(x_2, y_2) + \ldots\ldots. d(x_n, y_n)$.

S(V)=2 (1)

From the definition of centroid the set of minimum labels is

1+1=2 (2)

Using equations (1) and (2) we can conclude that a vertex is a median vertex if and only if it is a centriod vertex.

E. Theorem

Centroid of a tree always contains one vertex or pair of adjacent vertices.

Proof:

Centroid definition states that, it is the set of vertices minimum labels. From the Fig. 1, the minimum vertex is 2. Above the minimum vertex we have one vertex and below the 2 we have two adjacent vertices. Therefore we conclude that vertex 2 is the centroid of the binary tree.

F. Theorem

Let T be a tree and Z be a geodesic from x to y where y_1 is not a median vertex. Then

$$S(x)=S(y_0) < S(y_1) < \ldots.. S(y_n).$$

Proof:

From the Fig. 1 we assume a particular part (i.e). Now join all the vertices by horizontal edges.

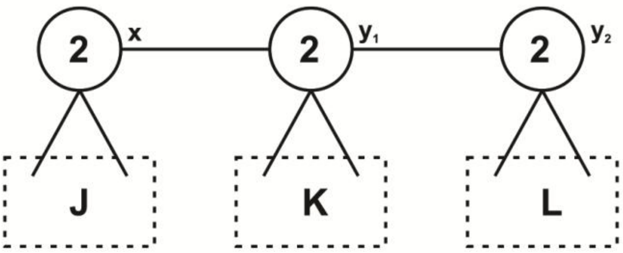

Fig. 2

Clearly we can say that $S(x) < S(y_1)$. Thus the difference between x and y_1 = $C(y_1,x)$ - $C(x, y_1)<0$.

Therefore x is closer to more vertices than y_1. From the above Fig. 2, let us take $C(x,y_1)=J$ and $C(y_1,x)=K+1+L$. Now we find

$S(y_1)$- $S(y_2)$ = $C(y_2,y_1)$- $C(y_1,y_2)$= L-{K+1+J}= L-K-1-J= L-J<0. Therefore $S(y_1)$< $S(y_2)$.

To generalize the theorem we consider the diagram below

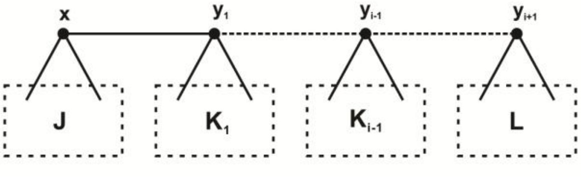

Fig. 3

$S(y_i)$-$S(y_{i+1})$=$C(y_{i+1},y_i)$-$C(y_i,y_{i+1})$=L_i-{J+1+(K_1+1)+(K_2+1)…
…+K_i}=L-J<0 and so $S(y_i)$<
$S(y_{i+1})$.

Thus we prove the theorem.

DOI:10.35940/ijrte.D1027.1284S519

VII. APPLICATIONS

Applications of trees are used in analysis of new algorithms and finding an algebraic expressions etc. Trees are widely used in computing data structures. Various forms of trees are Binary Search Tree, Hash tree and Huffman coding tree. Binary search tree is used in search algorithm creation. Hash Trees is used in programming a digital image signature. Huffman Coding Tree is used in squeezing data like .jpeg and .mp3.

VIII. CONCLUSION

Huffman coding is used to compress the data in the form binary digits for communicating some message to the receiver. The researchers convinced that in order to code a confidential message the method discussed in this paper can be made use of. In future, more coding method using different encoding algorithm on different graphs are planned to be executed.

REFERENCES

1. http://www.cs.cmu.edu/afs/cs/academic/class/15251/site/current/Maters/b Handouts/lecture20-proofs.pdf
2. Aigner, Martin, Gunter M. Ziegler, Karl H. Hofmann, and Paul Erdos. Proofs from the Book. Vol. 274. Berlin: Springer, 2010.
3. Philippe Flajolet and Robert Sedgewick Analytic Combinatorics, Cambridge University Press, 2009.
4. Alok Shukla, A short proof of Cayley's tree formula : arxiv: 0908.2324v3 [math.co] 4 Oct 2016.
5. http://www.cs.columbia.edu/~cs4203/files/GT-Lec4.pdf
6. http://w3.marietta.edu/~mmm002/Math350Spr06/Lectures/Chapter4.pdf
7. Kohn K M, Dong F M and Tay E G, Graph theory and its applications. Volume 31 no:2 December 2004.
8. K.Sunitha, Dr.C. David Raj and Dr.A. Subramanian, Radio labeling of Hurdle graph and Biregular rooted trees. ISOR journal of Mathematics, e-ISSN:2278-5728, Volme-13, pp 37-44.
9. Alyssa Gottshall, Justin Kahn, The Radio number of Biregular paths, Project code: PRC 3272.
10. F. Harary, Graph Theory Book, Addison-Wesley, 1969.
11. M.Senthil Kumar, V.Mathivanan, Analysis of data compression techniques using Huffman coding and Arthimetic coding. ISSN-2277-128x, Volume 6, 5, May 2016.
12. Latha Pillai, Huffman coding XAPP616 (v1.0) April 22, 2003.
13. D.A.Angel Sherin, V.Maheswari, Encryption and decryption process using edge magic labeling Journal of Physics: Conference series, ISSN-1742-6596/1362/1/01/2024.
14. D.A.Angel Sherin, V.Maheswari, Encoding the Graph using Instant Insanity puzzle and decoding with Hamiltonian cycle, The International Journal of analytical and modal analysis, ISSN-08869 - 9367.P.No: 167-175.
15. S.Rekha and V.Maheswari, Difference Modulo Labeling Journal of Physics: Conference Series, ISSN-1742-6596/1362/1/01/2049
16. 16. G.Uma Maheswari, G.Margaret Joan Jebarani and V.Balaji, Coding through a two star and Super Mean Labeling. Applied Mathematics and Scientific Computing pp 469-478.

AUTHORS PROFILE

D.A.Angel Sherin She is a Research scholar (Part-time) in Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai. She is working at S.D.N.B Vaishnav College for Women since 2017 till date. She completed M.Sc., M.Phil., in Mathematics. She has eight years of teaching experience. Her research interests Graph Labeling and Cryptography. She has published 2 research articles in both National and International journals.

Dr.V. Maheswari She is working at Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai since 2017 till date. She completed M.Sc., M.Phil., in Mathematics and her Ph.D in (Graph theory) Mathematics at Manonmaniam Sundaranar University, Tirunelveli. She has sixteen years of teaching experience. Her research interests Graph Labeling and Cryptography. She has published more than 15 research articles in both National and International journals. She has guided 3 M.Phil scholars and 5 Ph.D scholars.

.