International Journal of Modern Agriculture, Vol. No 10, Issue No. 2 (2021)
ISSN: 2305-7246

Securing Distributed Database Using Extended Blowfish Algorithm

[1]Sangeetha Radhakrishnan , Dr.A.Akila[2]
[1]Research Scholar, Department of Computer Science, School of Computing Sciences,Vels Institute of science, Technology & Advanced Studies (VISTAS), Chennai,India
[2]Associate Professor, Department of Computer Science, School of Computing Sciences, Vels InstituteScience, Technology &Advanced Studies (VISTAS), Chennai, India

ABSTRACT

Data is an asset for any dynamically operating organization and their business in today's world. The data to be secured is normally stored in a database. As this data is considered as an asset, it is to be secured from the unauthorised access and the intruders which is the major concern. Cryptography is considered as an efficient way for securing the data in the distributed database specifically in the organisations having a possibility of higher risk factor. In this paper, Blowfish encryption algorithm which is considered as the best algorithm among others is implemented with the educational dataset. The elongated version of the Blowfish algorithm is implemented as the proposed work for securing the data in the distributed database.

Keywords:-Database security, query processing, Encryption, Decryption, Blowfish Algorithm, Elongated Blowfish.

I. INTRODUCTION

Data is one among the foremost important issues for any organisation and the importance to stay the data to be secure for the growth of the corporation efficiently. Organisation's data are usually stored in the databases. The major challenge for the professional is to found out a technique which can keep the information safe and secure away from the intruders. Cryptography is a crucial dimension forsecurity in the database. The conventional security ways are not appropriate for providing the security when comparing with the encryption standards for data security. The security provided by the database encryption is as follows:-

- Cryptography can restrict users from getting the data in an unauthorised way.
- Cryptography can verify the authenticity of the originality of the data.
- When storage mediums like disks, tapes are lost,it prevents the data leakage in the database.

Earlier and even now, the security is an important concern in the databases. The encryption process of DDBMS can be used for securing the data from malicious attacks. Some of the cryptographic algorithms are DES, AES, BLOWFISH, RSA and Twofish algorithm.This article deal with implementation of the Blowfish cryptographic algorithm and the implementation of extended version of Blowfish algorithm with the educational dataset for securing the distributed database.

II. CRYPTOGRAPHY

Cryptography plays a vital partin information security system. It is significant for the data security system from sender to the receiver. It ensures the data confidentiality, reliability, accuracyand the integrity of the data. Presently, cryptography is processed for data security which is saved for longer periods to secure the e-fund transfers and other communications.

Unlike traditional cryptographic techniques, the modern cryptographic methods emphasis on the theoretical and algebraic techniques. Precisely, two kinds of cryptography are there. One is public key cryptography or asymmetric key cryptographic algorithm and the second is Symmetric or private key cryptographic algorithm.

A. Public key Cryptography

In public key cryptographic algorithm, 2 keys are used for encoding and decoding.One is public and the other is private key. The key which is public is used for the encryption and the latter for the decryption. RSA,Merkle's Puzzles, Elgamal and ECC Elliptical Curve Algorithm are few of the public key cryptographic algorithms[2]. The public key cryptographic algorithms are also called as asymmetric key cryptographic algorithms. There is no need of any secured key exchange from sender to receiver at the start. Asymmetric key algorithm is designed in a way in which the receiver can easily

generate the private and public key and decipher the data by the private key generated. The sender can easily encrypt the data by using the public key. It is not easy to get the private key by simply knowing the public key and hence the security is enhanced.

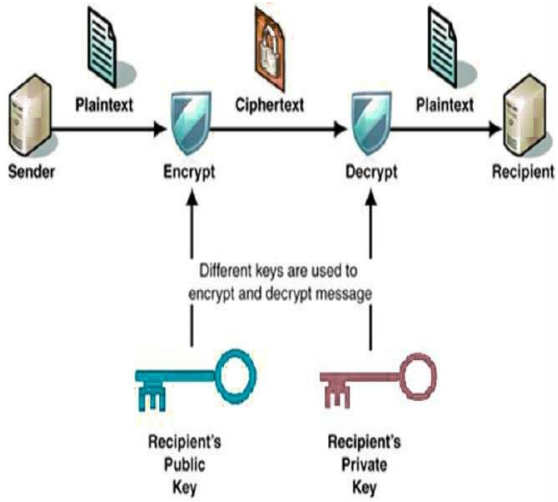


Fig 1:Public key cryptography

B. Private key Cryptography

In private key cryptography, only one key is taken for encoding and decoding process. Private key algorithm can be classified into stream and block cipher. The major distinction from stream to block cipher is that stream cipher encode single bit of plain text at one time.Whereas block cipher picks different number of bits usually 64-bits and encipher it as a single unit.

Some of the known private key algorithms are AES, DES, Blowfish, Two fish, RC4, RC6, 3DES etc. Even though private key cryptography is less intensive when comparing with the public key cryptography, but public key encryption is faster than private key encryption, in practice. Public key algorithm need at least 3000 bit key to acquire the similar security as for a 128 bit private key algorithm. Private key encryption is also called symmetric key encryption.

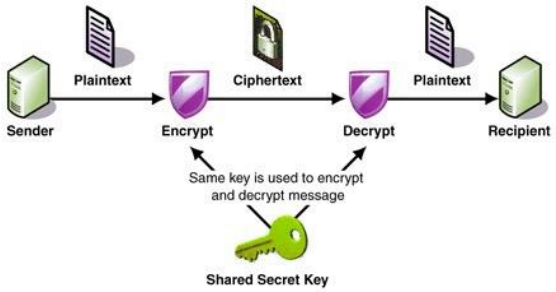


Fig 2: Symmetrickey cryptography

III. EXISTING METHODOLOGY

Blowfish is a private key cryptographic algorithm which can be considered as an alternative for other symmetric algorithms like DES and IDEA. The key is of variable length from 32-bits to 448-bits. The key length which is variable is apt for industrial and domestic usage. Blowfish symmetric algorithm was developed by one of the security expert Bruce Schneier as a free ware in 1993.

This algorithm is a block-cipher algorithm which is having 64 bit length. Because of its compactness, Blowfish can also be used in hardware applications and the algorithm has been not cracked yet [7][8]. Blowfish algorithm consists of two parts. One that focus on the enlargement of the key and the other part with focus on the data encryption.

A. **Key Expansion**

First part of the Blowfish algorithm--key expansion starts with P-array & S-boxes with usage of sub-keys. The sub keys need prior evaluation before the process of data encoding and decoding. The P-array contains 18 four byte sub-keys which is p1,p2,…..,p16,p17,p18.

This algorithm having 448-bit length is transmuted in to various sub-key arrays and all the 4 32-bit S-boxes having 256-entries such as s1,0, s1,1,s1,2….,s1,255 s2,0,s2,1,s2,2…,s2,255

S3,0,s3,1,s3,2….s3,255 s4,0,s4,1,s4,2….s4,255

Following are the step to produce sub-keys:

STEP 1 : Initialise 4 S boxes &P-array with string with fixed length. The string contains hexadecimal digits of π.

STEP 2: P1, P2 are the first and the second elements in P-array.p1 is XOR with initial 32-bits of key and p2 with second 32-bit of key.

STEP 3 : Repeat until every component in P-array XOR with key bit.

STEP 4 : Encode every zero string with blowfish using the keys which is stated in the Step 1 and 2.

STEP 5: Interchange p1 & p2 with result of Step 3.

STEP 6: With the usage of altered sub-keys encode the result of Step 3.

STEP 7: Interchange p3 & p4 with result of Step 5.

The 7 steps should be continued until all the P-array and 4 S-boxes are altered. The procedure continues until the whole P-array & 4 S-boxes are inter changed[3].

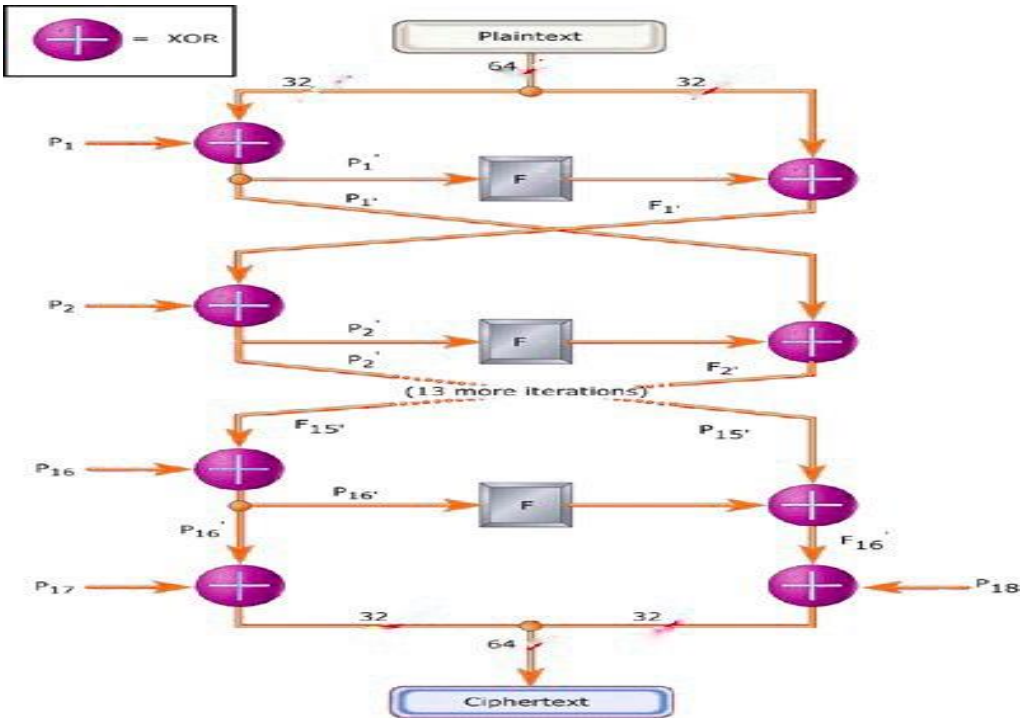


Fig 3:Blowfish Algorithm

B. **Encryption/Decryption**

The process of encryption with the blowfish algorithm of the message 'Hi space' is as follows.

- The block size of the blowfish algorithm is 64 bits.
- The message to be encrypted 'Hi space' contains 7 characters and a space which is 64 bits i.e, 8 bytes.

- Start by dividing the message to be encrypted which is 'Hi space' into two 32-bits.Left 32-bits represents the 'His' XOR with p1 and create the value named p1'.p1 is the initial element of p-array which mentioned in key expansion part.
- Execute the output p1' through the transformation function named F. In this, 32-bit decompose into 4-bytes each having a value in the 4 S boxes.
- The value of the S-box1 and S-box2 are added mutually and the result will be XOR with the value from the S box 3.
- The output is then added to the S box 4 value to generate 32-bit.
- The result of F is XOR with the message right 32-bits which is to be encryptedi.e,'pace' to generate F1.
- Left portion of message is replaced by F1 and right portion replaced by P1'.
- Whole procedure repeats fifteen times more than the P array successive elements.

The output of p16' and f16 after 16thiteration are XOR with final 2 additions in P-arrayi.e, p17 & p18. Next recombine it for generating 64-bit cipher text of 'Hi space' message. The function F is represented in the figure 4 below.

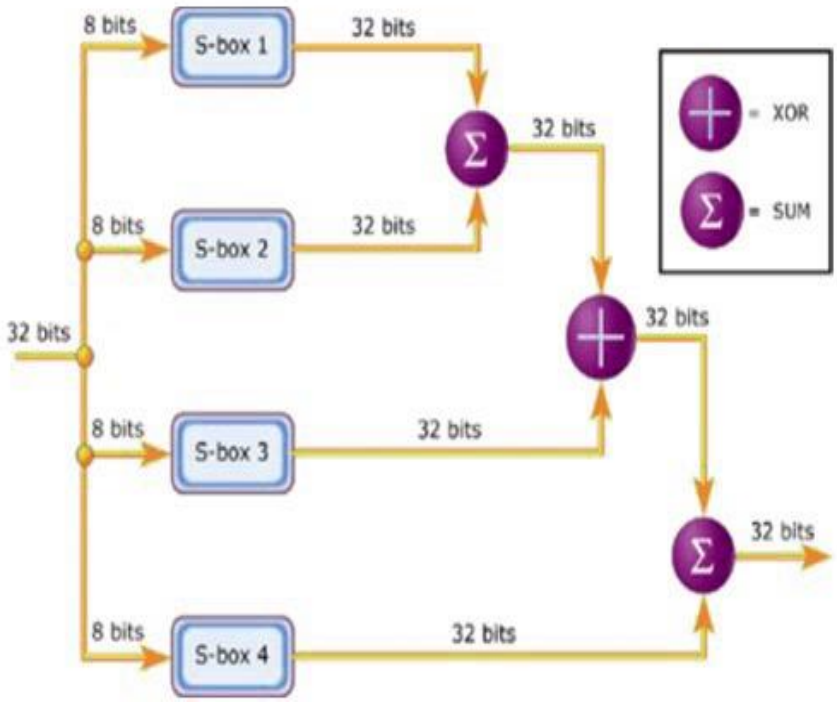


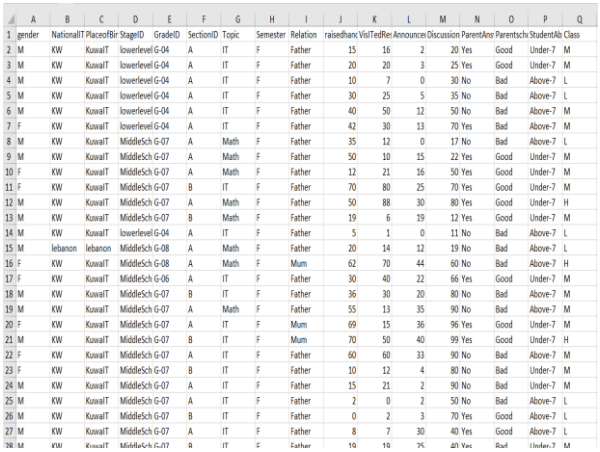
Fig 4: Function F of Blowfish Algorithm

IV. RELATED WORKS

Blowfish is considered as the efficient and fastest among the other symmetric algorithm. It is fastest of its kind except when the change of key happens[3].The security of the blowfish algorithm is tested and evaluated by many researchers but wrapped with the conclusion that it is the secured algorithm when compared to other cryptographic algorithms[1][5][6][4]. This article presents the elongated Blowfish algorithm by altering the F function.When the evaluation of the alterations made, the elongated blowfish is better compare to the traditional blowfish. Additional features are also make it more efficient. The performance evaluation is based on the different performance metrics like accuracy, response time, throughput etc.

V. DATASET

Educational dataset is collected for the performance evaluation of RSA algorithm in distributed database. 1441 records of dataset are used for the performance evaluation to show the efficiency of the proposed algorithm.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
1	gender	NationalIT	PlaceofBir	StageID	GradeID	SectionID	Topic	Semester	Relation	raisedhanc	VisITedRe	Announce	Discussion	ParentAnn	ParentsSch	StudentAb	Class
2	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Father	15	16	2	20	Yes	Good	Under-7	M
3	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Father	20	20	3	25	Yes	Good	Under-7	M
4	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Father	10	7	0	30	No	Bad	Above-7	L
5	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Father	30	25	5	35	No	Bad	Above-7	L
6	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Father	40	50	12	50	No	Bad	Above-7	M
7	F	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Father	42	30	13	70	Yes	Bad	Above-7	M
8	M	KW	KuwaIT	MiddleSch	G-07	A	Math	F	Father	35	12	0	17	No	Bad	Above-7	L
9	M	KW	KuwaIT	MiddleSch	G-07	A	Math	F	Father	50	10	15	22	Yes	Good	Under-7	M
10	F	KW	KuwaIT	MiddleSch	G-07	A	Math	F	Father	12	21	16	50	Yes	Good	Under-7	M
11	F	KW	KuwaIT	MiddleSch	G-07	B	IT	F	Father	70	80	25	70	Yes	Good	Under-7	M
12	M	KW	KuwaIT	MiddleSch	G-07	A	Math	F	Father	50	88	30	80	Yes	Good	Under-7	H
13	M	KW	KuwaIT	MiddleSch	G-07	B	Math	F	Father	19	6	19	12	Yes	Good	Under-7	M
14	M	KW	KuwaIT	lowerlevel	G-04	A	IT	F	Father	5	1	0	11	No	Bad	Above-7	L
15	M	lebanon	lebanon	MiddleSch	G-08	A	Math	F	Father	20	14	12	19	No	Bad	Above-7	L
16	F	KW	KuwaIT	MiddleSch	G-08	A	Math	F	Mum	62	70	44	60	No	Bad	Above-7	H
17	F	KW	KuwaIT	MiddleSch	G-06	A	IT	F	Father	30	40	22	66	Yes	Good	Under-7	M
18	M	KW	KuwaIT	MiddleSch	G-07	B	IT	F	Father	36	30	20	80	No	Bad	Above-7	M
19	M	KW	KuwaIT	MiddleSch	G-07	A	Math	F	Father	55	13	35	90	No	Bad	Above-7	M
20	F	KW	KuwaIT	MiddleSch	G-07	A	IT	F	Mum	69	15	36	96	Yes	Good	Under-7	M
21	M	KW	KuwaIT	MiddleSch	G-07	B	IT	F	Mum	70	50	40	99	Yes	Good	Under-7	H
22	F	KW	KuwaIT	MiddleSch	G-07	A	IT	F	Father	60	60	33	90	No	Bad	Above-7	M
23	F	KW	KuwaIT	MiddleSch	G-07	B	IT	F	Father	10	12	4	80	No	Bad	Above-7	M
24	M	KW	KuwaIT	MiddleSch	G-07	A	IT	F	Father	15	21	2	90	No	Bad	Under-7	M
25	M	KW	KuwaIT	MiddleSch	G-07	A	IT	F	Father	2	0	2	50	No	Bad	Above-7	L
26	M	KW	KuwaIT	MiddleSch	G-07	B	IT	F	Father	0	2	3	70	Yes	Good	Above-7	L
27	M	KW	KuwaIT	MiddleSch	G-07	A	IT	F	Father	8	7	30	40	Yes	Good	Above-7	L

Fig 5:Dataset

VI. PROPOSED METHODOLOGY

Convolution Operator

Convolution is a mathematical operator which is a base for many processing operators.* represents the convolution operator.It provides a way of multiplying together two arrays of numbers which is of different sizes usually but having the same dimensions. Result is the third array of numbers of the same dimension.

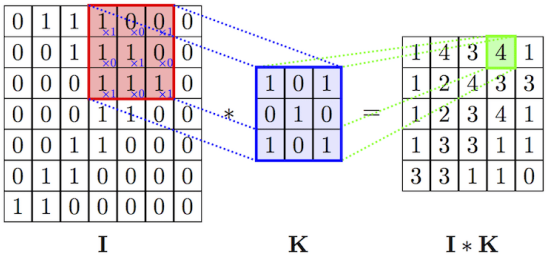


Fig 6: Convolution Operator

Syntax of convolution operator in Mat lab is
W=conv(u,v)
W=conv(u,v,shape)
W=conv(u,v) returns 2 vectors u & v convolution.

If the 2 vectors are of polynomial coefficients, convolution implies the multiplication of the two polynomials.

W= conv(u,v,shape) returns sub section of convolution by shape. For instance, conv(u, v, 'same') returns the convolution mid portion which have the same size of u. conv (u,v,'valid') returns the computed convolution part without 0 padded edges only.

ELONGATED BLOWFISH

In the extended blowfish algorithm the 64 bits are divided into 32 bits and compute the p1',p2' to p16'.The 32 bit is decompose into 4 bytes each having a value in 4 S boxes. The value of S box1 and S box2 is convolute together and is XOR with the value with the S box3.The result is then convolute to the S box4.The user input key K XOR with the result

which reinforce S-box in the process of encryption. Each S-box which is having a K randomly located in the S-box strengthens the quality of encryption.

F(XL)=(((S1, a * S4,d)mod32) XOR ((S2, b * S3,c) mod 32))) XOR K.

S-Substitution box, K- User Input Key, * - Convolution Operator.

PSEUDO CODE

STEP 1: Split x into 2 32-bit half:xL, xR
STEP 2: Then split into four 8-bits
STEP 3: Output of S-box1 and 4 are convoluted
STEP 4: Output of S-box 2 and 3 are convoluted
STEP 5: XOR output of both
STEP 6: Obtain function F(XL) = (((S$_{1,a}$ * S$_{4,d}$) mod2^32) XOR ((S$_{2,b}$ * S$_{3,c}$) mod2^32)))XOR K.
K is the user input and s is the substitution box.

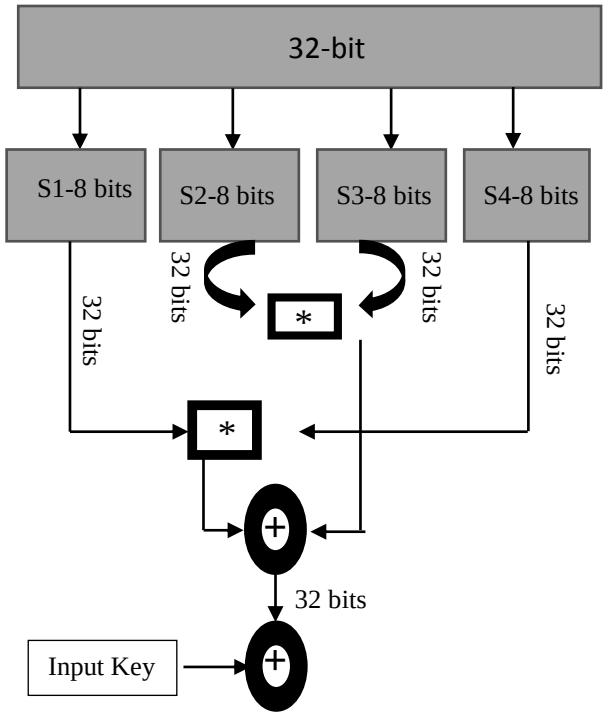


Fig 7:Convolution based F Function of Blowfish Algorithm.

VII. PERFORMANCE METRICS

1)Accuracy:Accuracy is taken for the performance evaluation in distributed environment for the RSA and Blowfish algorithm. Higher the accuracy, efficient the database is.

$$accuracy = \frac{length(encrypted\ data - original\ data)}{length(original\ data)} * 100;$$

2)Encryption Time:Another performance metrics taken into account for the performance evaluation of RSA and Blowfish is time. Time is considered as the important aspect in distributed environment. Lesser the time, efficient the database is.

Encryption Time= Time Taken to encrypt plaintext to ciphertext.

3)Bandwidth:Bandwidth is a challenge in running distributed computations. Higher Bandwidth, efficient the database is.

Maximum capacity of the data encryption.

4)Throughput: Throughput should be higher, the database will be efficient.

$$\text{Throughput} = \frac{\text{Average(Total Plain text)}}{\text{Average(Encryption time)}}$$

5)Response Time:The performance metrics Response time is a key metrics to evaluate the performance of RSA and Blowfish. Lesser the response time, efficient the database is.

$$\text{Response time} = \text{End time of execution} - \text{Start time of execution}$$

VIII. RESULTS AND DISCUSSIONS

Performance of the traditional blowfish algorithm and the extended blowfish algorithm is evaluated using different performance metrics such as throughput, accuracyetc.

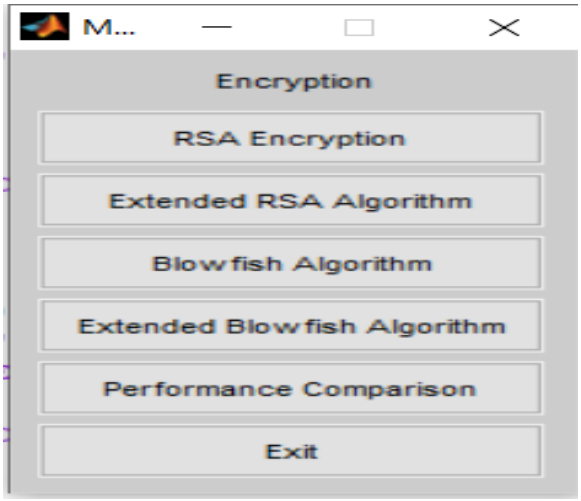


Fig 8:Home Screen

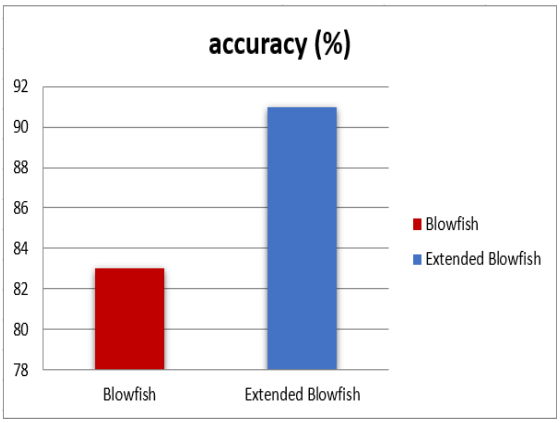


Fig 9:Comparision of accuracy of Blowfish and Elongated Blowfish

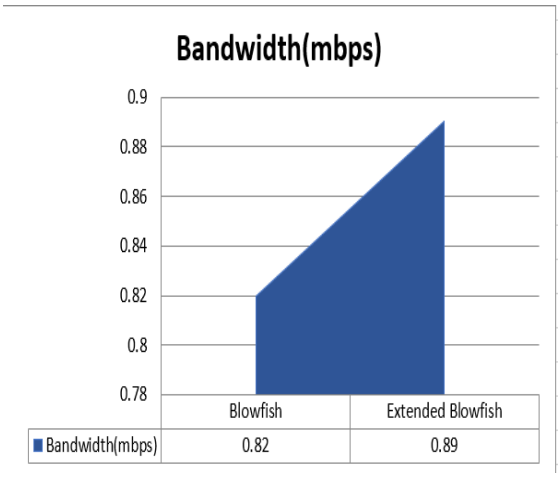


Fig 10:Comparision of Bandwidth of Blowfish and Elongated Blowfish.

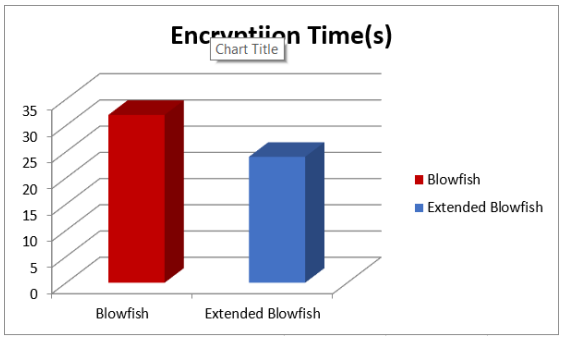


Fig 11:Comparision of Encryption Time of Blowfish and Elongated Blowfish.

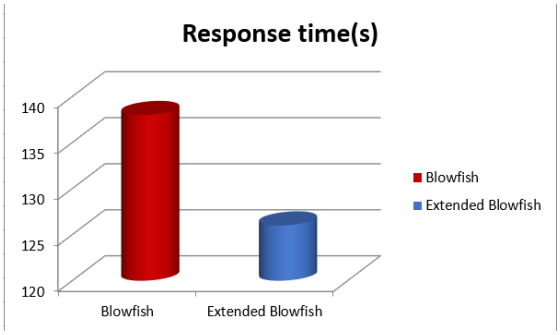


Fig 12:Comparision of Throughput of Blowfish and Elongated Blowfish.

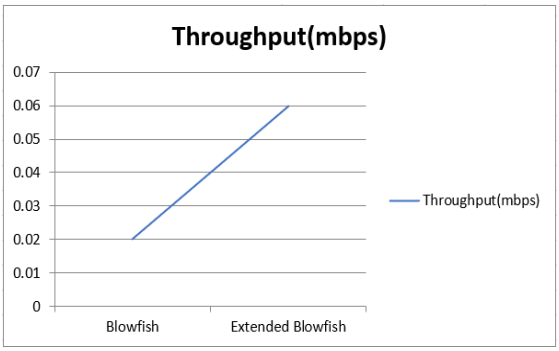


Fig 13:Comparision of Response Time of Blowfish and Elongated Blowfish.

IX. CONCLUSION

Cryptographic algorithm plays an inevitable part in security data system. These encryption algorithm uses different techniques to improve the user's data confidentiality and privacy by encrypting data which can be decrypt only by the authorized person who possesses the associated key. This paper deals with the implementation of blowfish algorithm in the distributed database using the educational dataset.

The proposed work is the implementation of the elongated version of Blowfish algorithm in distributed environment to secure the data in the database. In the elongated version,the 32 bits are decomposed into 4 S boxes with 8 bits. The S box2 is convoluted with the S box3 and the S box1 is convoluted with the S box4. The result of both is XOR ed to obtain the 32 bit. For further strengthening the S boxes, we have XOR ed the result with the User input key. The comparison of performance of traditional and the elongated Blowfish algorithm with various metrics is done and hence conclude that the extended Blowfish algorithm is better in performance compared to the original Blowfish algorithm for the dataset.

REFERENCES

[1] Joan Daemen,Vincent Rijmen,Fast Software Encryption, 9th International Workshop,FSE 2002,Leuven,Belgium,February 4-6,2002,Revised Papers, Springer 2002, ISBN 3-540-44009-7
[2] William Stallings, "Cryptography and Network Security", Third Edition,Pearson Education 2003.
[3] B. Scheier, "Description of a New Variable Length key,64 bit Block cipher (Blowfish)",FastSoftware Encryption,Cambridge security workshop proceedings Dec 1993,Springer-Verlag,pp 191-204,1994.
[4] Ashwaq T Hashim "Type 3 Feistel Network of the 128 bits block size Improved Blowfish Cryptographic Encryption" IJCSNS international Journal of Computer Science and Network Security Vol 8 No 12,pp 280-286,Dec 2008.
[5] G.Chen,Y. Mao,C.K.Chui, "A symmetric image encryption based on 3D chaotic maps", Chaos solutions and fractals, vol 21,pp 749-761,2004.
[6] Vincent Rijmen, Bart Preneel,Erik De Win, "On weakness of Non surjective round functions designs codes and cryptography", springer,vol 12(3),pp 253-266,1997.
[7] Nie T Song and Zhi X, April 2010,Performance evaluation of DES and Blowfish algorithms in biomedical Engineering and computer Science (ICBECS) pp 1-4,IEEE.
[8] Oukili S and Bri S, 2016, High Throughput Parallel Implementation of Blowfish Algorithm, Applied mathematics & Information Sciences 10(6),pp 2087-2092.
[9] Dr.G.Suresh, Dr.A.Senthil Kumar, Dr.S.Lekashri, Dr.R.Manikandan. (2021). Efficient Crop Yield Recommendation System Using Machine Learning For Digital Farming. International Journal of Modern Agriculture, 10(01), 906 - 914. Retrieved from http://www.modern-journals.com/index.php/ijma/article/view/688
[10] Dr.A.Senthil Kumar, Dr.G.Suresh, Dr.S.Lekashri, Mr.L.Ganesh Babu, Dr. R.Manikandan. (2021). Smart Agriculture System With E – Carbage Using Iot. International Journal of Modern Agriculture, 10(01), 928 - 931. Retrieved from http://www.modern-journals.com/index.php/ijma/article/view/690