Jois et al. Journal of Open Innovation: Technology, Market, and Complexity (2015) 1:19
DOI 10.1186/s40852-015-0022-8

Journal of Open Innovation: Technology, Market, and Complexity
a SpringerOpen Journal

RESEARCH Open Access

A 3-d advancement of PythoCrypt for any file type

Harsha S. Jois[1,2*], N. Bhaskar[2,3] and M. N. Shesha Prakash[2,4]

* Correspondence: harsha_jois@hotmail.com
[1]Department of Information Science & Engineering, Vidya Vikas Institute of Engineering & Technology, Mysuru 570028, India
[2]Visvesvaraya Technological University, Belgaum, India
Full list of author information is available at the end of the article

Abstract

This research work discusses a file-type-independent cryptosystem using the conversion of Cartesian and Polar coordinate systems in three dimensions. This is an advancement to PythoCrypt (IJESM, Vol-3,Issue-2, 48-51,2013) where a new algorithm was discussed that can be effectively used to provide security to medical images (IEEE, Proceedings of ICe-HNAS-9, 244 – 247, 2007) or files that have a small region or block of interest (JMBE, Vol- 27, Issue-3,144-149, 2007)(Springer.Link, FAW-1 proceedings, 62-73,2007). The usage of geometrical objects for encryption is still young and an attempt was made to employ an enhanced method of PythoCrypt to encrypt different file types. The present research work explores the inter-operability of the two coordinate systems as a basis for a cryptosystem. The work proposes a crypto-system with considerable security with one requirement; the use of a secure channel for key exchange.

Keywords: Cryptosystem, Cartesian, 3-d coordinates, PythoCrypt, Block cipher

Introduction

Cryptography is the prime requisite for secure communication systems as networks are required to provide the user with privacy and security. The increase in security threats to communication and data has prompted a great leap in research towards development of advanced cryptosystems. The evolution is documented as follows.

Apart from Caesar cipher (Andrew 2006) - a cyclic substitution cipher used by Julius Caesar in 50 B.C., there are many codes that are no longer used. Instead of the codes the security of information is more dependent on higher mathematical algorithms.

The Diffie–Hellman (Diffie and Hellman 1976) key exchange algorithm addressed one of the biggest challenges of all reasonably secure cryptography that was the need for a secret key distributed separately to ensure encrypted communications. In 1976, Whitfield Diffie and Martin Hellman published their method for public exchange of a secret key. This algorithm popularly known as Diffie - Hellman key exchange and was the first known way to distribute keys for secure communication that didn't involve transmitting the key by some private mechanism first.

The RSA (Rivest) algorithm proposed by and named after R. Rivest, A. Shamir and L. Adleman in 1978 is used for both public key encryption and digital signatures. It is the most widely used public key encryption algorithm. The basis of the security of the RSA algorithm is that it is mathematically infeasible to factor sufficiently large integers.

The RSA algorithm is believed to be secure if its keys have a length of at least 1024-bits.

Data Encryption Standard (Meyer and Matyas 1982; Coppersmith 1992) is a block cipher that was selected by the National Bureau of Standards as an official Federal Information Processing Standard (FIPS) for the United States in 1976 and which has subsequently enjoyed widespread use internationally. It is based on a symmetric-key algorithm that uses a 56-bit key. DES consequently came under intense academic scrutiny which motivated the modern understanding of block ciphers and their cryptanalysis. DES is now considered to be insecure for many applications. This is chiefly due to the 56-bit key size being too small

Advanced Encryption Standard (Highland 1992) is an encryption standard which comprises of three block ciphers, AES - 128, AES – 192, AES – 256. Each AES cipher has a 128-bit block size, with key sizes of 128, 192 and 256 bits, respectively. The AES ciphers have been analyzed extensively and are now used worldwide, as was the case with its predecessor, the Data Encryption Standard (DES).

3-DES is similar to the DES. It is three times slower than the actual DES but when used appropriately is more secure than DES. Triple DES enjoys much wider use than DES because DES is so easy to break with today's rapidly advancing technology. Triple DES is simply another mode of DES operation. It takes three 64-bit keys, for an overall key length of 192 bits. It also has the advantage of proven reliability and a longer key length that eliminates many of the shortcut attacks that can be used to reduce the amount of time it takes to break DES.

Elliptic Curve Cryptography (William et al. 2013) is one of the latest additions to Asymmetric Key Cryptosystems. It uses planar elliptic curves as bases for encrypting data, where points on the plane are used as parameters with infinity (∞) as a reference (identity) point. However, like any asymmetric (public) key cryptosystem, the integrity of this system is not proven.

Quantum Key Cryptography (William et al. 2013) or Quantum Cryptography is one of the most promising new entries into symmetric key cryptosystems. It employs quantum mechanics principles to accomplish cryptographic tasks. Even though currently it is being used only sparingly for key distribution and is extremely costly to install, it may be the future of cryptosystems as it is the only defense against a quantum computing attack (William et al. 2013).

PythoCrypt (Harsha et al. 2013) discussed the application of Pythagorean triplets for encryption and subsequent decryption of medical images in two dimensions. Here a concept that is used in civil engineering; especially survey (Geological Survey Professional Paper, Volumes 1375–1984, Survey marking, Cornell University, Fall 1981), to locate a point using 3 parameters is used. The method is to permanently mark a point before closure of the days survey work with which the commencement of survey can be done without hassle on any subsequent day. The permanent point will be theoretically marked with two basic parameters and one validating parameter. There can be more number of points satisfying that basic parametric condition. The basic parameters are sufficient for encryption but the third dimension parameter or validating parameter is necessary for decryption. Hence this location system is a "3–d coordinate system" that is very efficient for surveys that need to cover very large areas and requires more time to complete the work.

In this paper the method of 3–d system with mapping between Cartesian and Polar coordinates that can be used for encryption and decryption of files of various types and sizes effectively will be discussed. Pythocrypt (Harsha et al. 2013) was specifically developed for medical images where as the present system proposed can be applied for any file type supported by any platform and hence it is termed file-type-independent. The algorithm proposed in this paper shall henceforth be addressed as "3-d Pythocrypt".

Proposed system

3-d Pythocrypt is a new concept which has been developed to encrypt and decrypt files of all types using a file-type independent algorithm proposed in Pythocrypt (Harsha et al. 2013). The algorithm "3-d Pythocrypt" uses the infinite divergence (Ashtiyani M., etal, 2008) of the 3-dimensional coordinate system definition of a sphere. The algorithm is unique compared to all the existing algorithms as it does not need a key for encryption, but generates one for the purpose of decryption. Also the entire algorithm is asymmetric, i.e. the encryption and decryption processes are dissimilar.

Design

The algorithm "3-d Pythocrypt" is independent of platforms and applications. It can be implemented on any system depending on the type of application needed. Any application such as mail, chat and file system can use this algorithm as an add-on without hindering the performance.

As shown in the Fig. 1, the functions use the relationships between the radius, tangent and the angle between them. The equations in the Fig. 1 form the basis of Trigonometry. They were discovered by Pythagoras – a great Greek mathematician (Pythagoras). Figure 1 shows the interrelation between the angles and the distances that subtend those angles on planes. Each point in space can be represented by Cartesian coordinates (x, y) or Polar coordinates (r, θ).

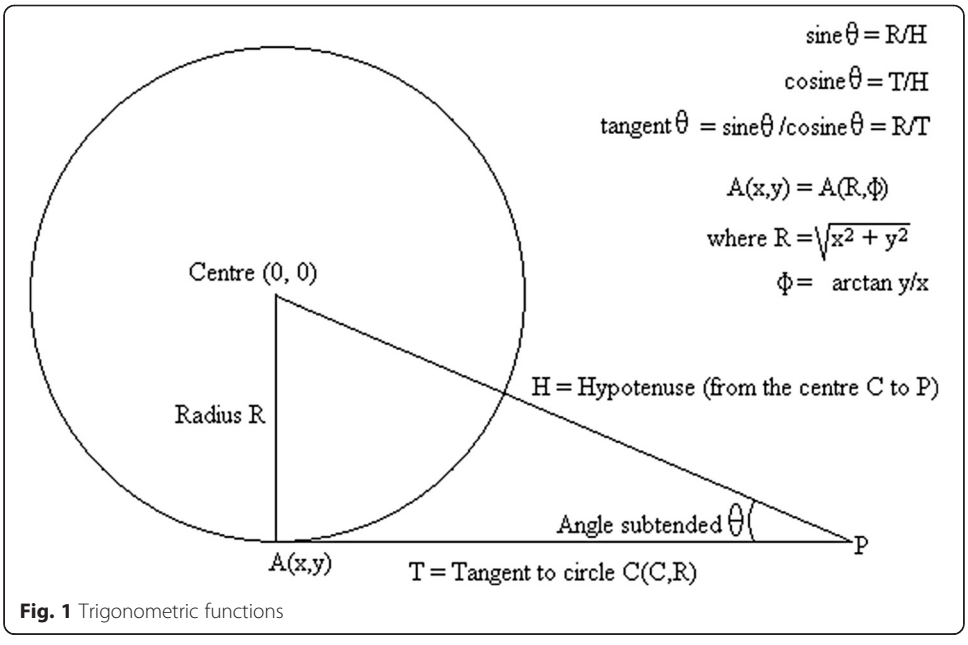

Fig. 1 Trigonometric functions

It is this feature that prompted the use of the inter-operability of the Cartesian and Polar coordinates for 3- dimensional systems.

Similar to the 2-dimensional coordinates, the 3-dimensional coordinates are also infinitely divergent. This unique feature has been exploited in the design of "3-d Pythocrypt".

Algorithm and implementation

The algorithm is designed based on the relations shown in the Fig. 2. "3-d Pythocrypt" uses a mapping function from the plain text to cipher text (Andrew 2006) based on a poly-alphabetic spread mapping (Kahn 1967). The algorithm "3-d Pythocrypt" is based entirely on the divergence of the 3-dimensional coordinate system. The coordinates of a point P are usually given by a set (x, y, z) denoting a point in the space called Cartesian coordinates. These also correspond to the same point using another set of parameters (r, θ, Φ) called the polar coordinates. In the first set, the values of x, y, z are the corresponding coordinates, where as in the second, the values r, θ, and Φ correspond to the radius, x-y planar angle and y-z planar angle of one point of a sphere having the origin O (0, 0, 0) of the coordinate space in focus, as centre.

The algorithm can be broadly divided into 3 phases.

The Encryption phase: Here the plain text i.e. the information is read as byte stream. One set of 3 chunks having 2 bytes each is mapped onto a sphere base using the 3-D coordinate system.

The Key generation phase: Here the destination set, i.e. the parameters defining the sphere are obtained and separated into two subsets.

The sphere's base is combined with one of the subsets and stored as cipher text. The other subset forms the Decryption key.

The Decryption phase: The cipher text and the key are parsed to separate the Sphere base and the definition parameters.

The actual byte sets are obtained by individually operating the sphere base with the necessary parameters. During this process, each set of 6 bytes (48 bits) of plain text

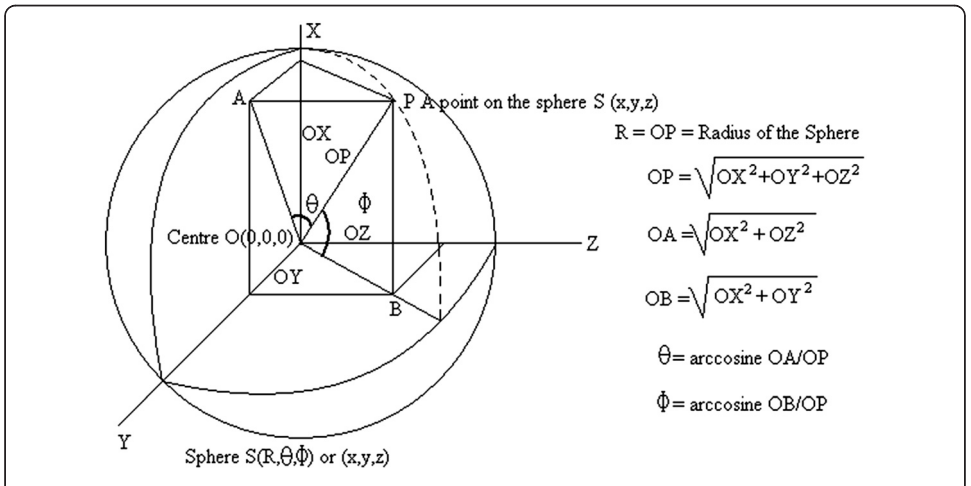

Fig. 2 Parameters of 3D Sphere and the relationship between them

generates 8 bytes (64 bits) of cipher text and key and it has a minimal bloating up of the file which means, lower network transit time (Andrew 2006), Computation Time (CPU time in term of clock cycles) and Optimization of routing Algorithms (Andrew 2006).

The encryption phase

The design of the algorithm uses the inter-convertibility from Cartesian to Polar Coordinates. It is necessary to have at least four of these parameters (Niven and Zuckerman 1972) to evaluate the other four. That is a condition, which can exist only if the four available parameters are independent. Otherwise a subset having five parameters is the minimum requirement for the complete set to emerge.

The algorithm "3-d Pythocrypt" utilizes this to and fro conversion for encrypting and decrypting data.

The plain text is read in sequences of 16 bits (2 bytes) each using a simple parser. 3 such sequences form a block of plain text. This set forms the *x, y, z* coordinates of a point *P* on a sphere of radius *R* given by the equation shown in the Fig. 2. As

$$\mathrm{R} = \sqrt{(\mathrm{OX}^2 + \mathrm{OY}^2 + \mathrm{OZ}^2)} \tag{1}$$

R is a large number in the format ccccccc.eeeeeee. R is floored (rounded down) to get only ccccccc. This forms C (Cipher text Parameter 1). The number of digits (n) in eeeeeee is counted. C is subtracted from R to get 0.eeeeeee, which is then multiplied with 10^n to obtain eeeeeee as another integer. This forms E (cipher text parameter 2).

The planar radii *OA* and *OB* form the parameters *A* and *B* respectively.

Then the supporting parameters θ and Φ are obtained using arccosine functions on *A*, *B* and *R* as shown in Fig. 2.

The key generation phase

The parameters obtained in previous section are separated and put into two files having type suggested by the user (Meyer and Matyas 1982; Bruce et al. 1995). These form the cipher text and the key. The key generated after the encryption is complete and does not play any role in the process of encryption. This is the most important characteristic of "3-d Pythocrypt".

The separation is done as *C, E* and *OA* go to the cipher text file and *OB,* θ and Φ go to the key file. Each of these parameters is 16 bits long and hence the size change from *N* bytes in the plain text to $4N/3$ in the cipher text. The generated key also has the same length as that of the cipher text.

The decryption phase

In this phase, the cipher text and the key text are combined to bring back the original plain text (Sinkov 1966). However, unlike the existing standard algorithms the decryption process in "3-d Pythocrypt" is not a mirror of the encryption process. This is again, a unique characteristic of "3-d Pythocrypt" which is explained as follows.

Once both the files are received, *C, E, OA, OB,* θ and Φ are read from Cipher and Key files using the previously agreed length (16 bits). C is added with $\mathrm{E}/10^n$ to obtain R. Then using the following operations OX, OY and OZ are obtained.

$$OY = \sqrt{(R^2 - OA^2)}, OX = RTan(\theta), OZ = Rtan(\Phi) \quad (2)$$

The numbers (OX, OY and OZ) are written into a file which forms the decrypted plain text file. Since the decryption process uses a separate set of operations compared to encryption unlike existing algorithms, this asymmetry of process is a unique feature.

Experimentation

The algorithm is tested on a number of files of different types. First the program is run on a simple English string. The cipher text and key generated are shown in the table in case 1 (Additional file 1). The last column displays the decrypted text and it is clear that there is no error in the decryption phase.

Then the algorithm is run on a simple image file as shown in Additional file 1: Figure S1. It is a sample image available in all windows™ Computers to be used in MS Office™ applications. The cipher text and key are shown in case 2 (Additional file 1), as text files. The decrypted image is shown in Additional file 1: Figure S2. Once again there is no difference between the decrypted file and original file indicating the success of "3-d Pythocrypt".

As an example to show that "3-d Pythocrypt" also works on formatted files (.doc, .docx, .xls, .ppt etc.), this document (3-d Pythocrypt.docx) is encrypted using "3-d Pythocrypt" and the results are shown in Additional file 1: Figures S3-S6.

Analysis

Any new system needs to be compared against existing standards to bench mark its performance(Feynman, 1989). The algorithm was compared with some standard existing algorithms and Pythocrypt (Harsha et al. 2013) using a small block size of plain text for reference. The results are analyzed and tabulated in Table 1.

The parameters used to obtain the Weighted Space-Time-Complexity Index proposed in this paper and the parameters used to obtain it are detailed below.

The file sizes taken as 1024 bits where as the key sizes are dependent on the algorithms. The encryption and cryptanalysis are done on an AMD™ Phenom™ Octa-core Processor and times are normalized. The time required to cryptanalyze One Time Pad (OTP) (Highland 1992) is taken as infinity (∞) and all other parameters are normalized against the maximum values in their corresponding range. The numerical values are tabulated.

From Table 1, it is evident that "3-d Pythocrypt" fares well against standard algorithms with 2 byte block size. However, it can outperform most algorithms as the cipher text will be too large in size and any sphere has infinite surface points with increase in block size. Further it also has infinite divergence for backtracking. A new radical approach of cryptanalysis titled "Multi agent pattern recognition" (William et al. 2013) is able to obtain parts of possible plain text after multiple iterations on a single file. This approach uses small pieces of codes called agents to try and obtain similar patterns in a packet in transit. They work independent of the source and obtain repeated patterns in the files. Since "3-d Pythocrypt" uses 16 bit blocks for this experiment, in English text, if there are 3 similar 2 letter blocks appearing consecutively, they would generate same cipher text. Once the source program obtains all the repeated patterns and their locations in the cipher text, it can then run a dictionary attack (William et al. 2013, Rabin,

Table 1 Performance comparison of "3-d Pythocrypt" with some standard algorithms and Pythocrypt (Harsha et al. 2013) considering Weighted Space-Time-Complexity (WSTC)

Algorithms Parameters	RSA (Rivest)	DH (Diffie and Hellman 1976)	DES (Coppersmith 1992)	AES (Pythagoras)	3-DES (Biham et al. 1993)	PythoCrypt (Harsha et al. 2013)	3-d Pythocrypt	OTP	Remarks
Type	Asymmetric key	Asymmetric key	Symmetric key	Symmetric key	Symmetric key	N/A	N/A	Symmetric Key	Pythocrypt and 3-d Pythocrypt are unique as the encryption process is independent of the key
Methodology	Modulo-n Arithmetic	Modulo-n Arithmetic	Multiple Iterations of Simple logical operations	Multiple Iterations of Simple logical operations	Multiple Iterations of Simple logical operations	Trigonometry of 2-dimensions	Trigonometry of 3-dimensions	Modular Addition	The conversion between coordinate systems is the key in 3-d Pythocrypt
Cryptanalysis methods that can be applied (Schneier 2000; Bauer 2002)	GNFS	GNFS	Brute Force, Relative Key (William et al. 2013)	Brute Force, Relative Key (William et al. 2013)	Relative Key (William et al. 2013)	Multi-agents (William et al. 2013)	Multi- agents (William et al. 2013)	Impossible	"3-d Pythocrypt" can be partially cryptanalyzed using Multi agent systems for pattern recognition
File Size	1024	1024	1024	1024	1024	1024	1024	1024	Common
Key size	128	64	56	256	192	512	683	1024	Varies depending on implementation
Encryption Time	0.208	0.208	0.402	0.604	1.206	0.278	0.312	0.314	On an octa-core (William et al. 2013) AMD™ Phenom™ Processor
Cryptanalysis Time	3.40282E + 38	1.8447E + 19	7.2058E + 16	1.16E + 77	6.2771E + 57	1.3E + 154	4.013E + 205	∞	Cryptanalysis Time ($2^{Keysize}$) for Brute Force Analysis (Schneier 2000; Bauer 2002)
Normalized Key size	0.125	0.0625	0.0546875	0.25	0.1875	0.5	0.66699219	1	Ratio of Key size and File size
Normalized Cryptanalysis Time	8.4792E-168	4.597E-187	1.796E-189	2.9E-129	1.564E-148	3.34E-52	1	∞	"3-d Pythocrypt" needs the maximum time for cryptanalysis
Normalized Encryption Time	0.172470978	0.17247098	0.33333333	0.500829	1	0.230514	0.25870647	0.260364842	Normalised against 3-DES
Weighted Space–Time-Complexity Index	0.8013	0.6728	0.7691	0.9301	0.928	0.8322	0.9491	1	The proposed index for comparison

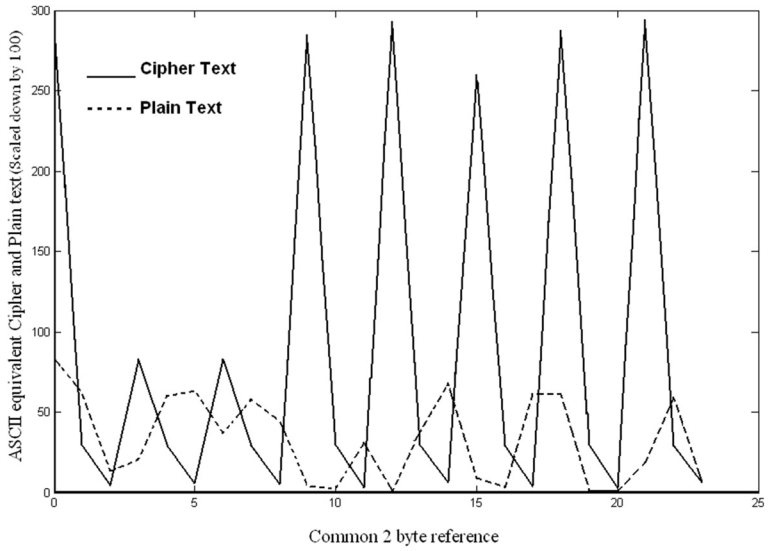

Fig. 3 Comparison plot of Cipher and Plain text for case 1

1976) on the patterns to generate possible sets of plain text. However this works only on unformatted English text and since "3-d Pythocrypt" encrypts the entire file including the header block, even this form of attack can only partially succeed after multiple iterations of dictionary attack. This is because each file has parts of its own data as the key similar to OTP (Highland 1992).

For the plain text shown in case 1 (Additional file 1), the plain text against the cipher text generated was plotted (Fig. 3) and the distance between the two was calculated (Fig. 4). The Figs. 3 and 4 indicate a very low correlation (Gujarati 2003) between plain text and cipher text. The coefficient of correlation (computed using Pearson Product moment correlation coefficient formula (Highland 1992)) is 0.128.

Conclusion

With the results of the analysis of the algorithm "3-d Pythocrypt", it is concluded that:

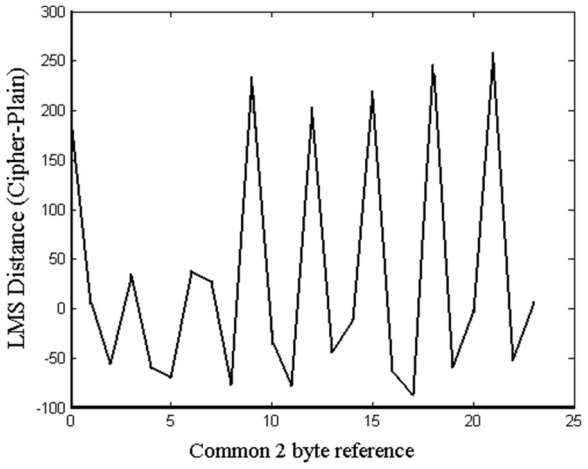

Fig. 4 Least Mean Square Distance plot for Graph in Fig. 3

- The File-type, size and Platform independent nature makes it a unique cryptosystem
- It provides better security against most of the existing cryptanalysis methods
- The only susceptibility to Cryptanalysis is partial in nature (only pieces of numerical values can be obtained and not the actual file data) using a multi agent approach for pattern recognition
- It surpasses most of the cryptosystems that are in use presently as shown in Table 1 using the proposed Weighted Space-Time-Complexity index.

Additional file

Additional file 1: Experimental data. (DOC 390 kb)

Competing interests
The authors declare that they have no competing interests.

Authors' contributions
HSJ designed the algorithm and developed the system. NB participated in the development of the mathematical model for the system and its validation. MNS provided the insight to the survey methods of reference points and the complexity index. HSJ and MNS drafted the manuscript. All authors have read and approve the final manuscript.

Author details
[1]Department of Information Science & Engineering, Vidya Vikas Institute of Engineering & Technology, Mysuru 570028, India. [2]Visvesvaraya Technological University, Belgaum, India. [3]Department of Mathematics, Vidya Vardhaka College of Engineering, Mysuru 570002, India. [4]Department of Civil Engineering, Vidya Vikas Institute of Engineering & Technology, Mysuru 570028, India.

Received: 29 March 2015 Accepted: 29 November 2015
Published online: 18 December 2015

References
Andrew S. Tenenbaum, "Computer Communication Networks", McGraw Hill, Revised 4th edition. 2006.
Ashtiyani, M. Electr. Eng. Dept., IHU, Tehran, Birgani, P.M. ; Hosseini, H.M, Chaos-Based Medical Image Encryption Using Symmetric Cryptography.
Bauer FL. "Decrypted Secrets Methods and Maxims of Cryptology". Springer; 2002. 978-3-540-48121-8 ISBN.
Biham E, Shamir A. "Differential Cryptanalysis of the Data Encryption Standard". Newyork:Springer Verlag; 1993.
Bruce Schneier, Applied Cryptography, 2nd edition, Wiley, 1995
Coppersmith D. "The Data Encryption Standard (DES) and its strength against attacks". IBM. 1992; 38(3):243–50.
Diffie W, Hellman M. New directions in cryptography. IEEE transaction. 1976;22(1):644–54.
Feynman RP. Lectures in Applied Mathematics and Theoretical Physics, Princeton University Press, 8th reprint. 1989.
Geological Survey Professional Paper, Volumes 1375–1984, Survey marking, Cornell University, Fall 1981
Gujarati D. Basic Econometrics, 4th edition. McGraw-Hill Publications. 2003. 0072335424, 9780072335422 ISBN.
Harsha S, Bhaskar N, Chandan CM. "PythoCrypt-A cryptosystem for Medical Images". IJESM. 2013;3(2):48–51.
Highland HJ. "Perspectives in Information Technology Security", Proceedings of the 1992 IFIP Congress, Education and Society. 1992.
Kahn D. The Code breakers - The Story of Secret Writing. 1967.
Meyer CH, Matyas SM. "Cryptography: A New Dimension in Computer Data Security", John Wiley and Sons. 1982.
Niven I, Zuckerman H.S. An Introduction to the Theory of Numbers. Newyork U.S.A:Wiley; 1972.
Pythagoras, Pythagorean Theorem, 570–495 BC.
Rabin MO. "Probabilistic algorithms: Algorithms and Complexity", J. F. Traub Education, New York:Academic Press; 1976.
Rivest RL, Shamir A, Adleman L. "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems". Communications of the ACM. 1978;21(2):120–126.
Schneier B. Self-Study Course in Block Cipher Cryptanalysis. Cryptologia. 2000;24(1):18–34.
Sinkov A. "Elementary Cryptanalysis: A Mathematical Approach", Mathematical Association of America. 1966. e-book.
William S. Cryptography and Network Security. 5th Edition. USA:Pearson Publications; 2013.
Yang Ou, Department of Information Security, Chul Sur, Kyung Hyune Rhee, Division of EC and TE, Pukyong National University, Region-based selective encryption for medical imaging, Busan, Republic of Korea.