IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 3, March 2017

A Survey on Basics of Cryptography

Vaibhav V. Bhujade[1], **Deepak Chaudhary**[2], **Suraj V. Raut**[3]

Asst. Professor, Dept. of CSE, DMIETR, Wardha, Maharashtra, India [1]

Asst. Professor, Dept. of CSE, IET, Alwar, Rajasthan, India [2]

Asst. Professor, Dept. of CE, BDCE, Wardha, Maharashtra, India [3]

Abstract: Cryptography is one of the important and useful technique in which usually a particular file is converted into unreadable format by using public key and private key system called as public key cryptosystem. Then as per the user requirement that file is send to another user for secure data or file transmission between original sender and receiver. In this transmission file of unreadable format is send, after receiving this file receiver used the similar algorithm technique and private key for getting the original file data. In this procedure various algorithms are used as a processing function and depending on that algorithm we used the private key. The power or strength of any algorithm is depending on the secret key used in sender client and receiver side client.

Keywords: Cryptography, authentication, Integrity, Confidentiality.

I. INTRODUCTION

1.1 CRYPTOGRAPHY BASICS

Cryptography is one of the scientific techniques in information security. The word cryptography is derived from the Greek word kryptos, it means hidden thing [1]. Cryptography is very similar to both the disciplines of cryptology and cryptanalysis. This cryptography includes various techniques which involve in hiding any kind of info in storage unit or transmit the data through various ways. However in this computer specific world cryptography is related with the scrambling the normal text which is available in readable form called plaintext into unreadable text called cipher text this process is known to be encryption procedure and the process which is exact oppose of this process used to recover the normal text known as process of decryption and the persons who used this techniques called as cryptographers [2] [3]. Current cryptography worries itself with the following four things

1.1.1 CONFIDENTIALITY:

Student grade information of college or school is the thing which is very important as per their point of view and its confidentiality is also a serious issue. In the country like United States publication this important information is controlled by Family Educational Rights and Privacy Act (FERPA). This student personal and confidential information should be available to the students, their parents and the staff who required such data for their own job only. Enrolled student information will be of specific confidentiality ranking. As FERPA is available to control this information about the students, but still this information is view by many people regularly and there may be possible damaging by using this information. Many time the information like various list that may be list of students, list of faculties or any departmental important lists are assigned as a low confidential rating in worst no rating in some cases. This data is easily available on the college or organization website to the public [4]. This term covers two inter related concepts:

1.1.1.1 DATA PRIVACY:

It means that the private or the information which requests to keep confidential should not be available or disclosed to any other person who have no relation with that data. Confidentiality means the security of transferred data from any type of attack. Several level protections can be identified whenever we are trying for data transmission. The service which user used protects the users transmitted data over a time period. For example when we used the TCP assembly between the two systems, the several level of protection is used to prevent the release of all user data transferred through the TCP connection. This service can also use for the protection of single message or specific fields in the message. This change in the service is not used as large as of original use and many time this use become more complex than the regular use and practically expensive to implement. Also we can protect the analysis of data traffic from any unauthorized person who will get the idea about the type of data transmission by viewing the flow of traffic. And using this attacker may come to know the exact source from where it came and destination to where it goes, of the data transmission [4]. Also frequency, length and many other characteristics of data traffic can be known to attacker.

1.1.1.2 PRIVACY:

It is the promise of an individual or any specific group to preserve information about themselves means express themselves specifically. The boundaries and the information differ among culture to culture or group to group. When something is private it means it is very special and may be sensitive to that person.

DOI10.17148/IJARCCE.2017.63200

1.1.2 INTEGRITY:

Many examples can be given which explain several aspects of the term integrity like medical information about any person stored in the database of that hospital. The doctor must trust all the information kept in the database that it is correct and current. But on the other hand if any person (e.g., a nurse) who have right to see, view and update the information about any patient in database and that person use this to harm the hospital. The database of that hospital must be restored properly and this restoration must be trustable, also this error or mistake needs to be traceable by the responsible person. Patient sensitive data or information is a good example of high level requirement for integrity. Wrong information about the patient may result in harm or death to him and also for that hospital liability [4]. Let us take an example of one of the asset in which a website offer a forum or blog for discussion on any current issue or topic to only the user who are authorized or who are previously registered and verified by the administrator. So in this case a registered user or any hacker might modify or entered some entries on his behalf. If this type of forum exist and used only for the pleasure of the registered user, no used to generate any revenue through this discussion and not that much important for research then there is no major problem in this type of cases but the administrator who manage all this things will experience little loss in terms of financial and time only. Also one case of low integrity is online poll on any specific website. Several websites such as news websites offer this type of poll for their user that may be recorded or not with low level security so we cannot believe on the accuracy of such type of poll for any type of analysis on any topic or current issue. This above term covers two inter related ideas:

1.1.2.1 DATA INTEGRITY:

It gives guarantee that the information and their program are changed only in the specific and official format. As with confidentiality, integrity may be apply to message stream, to any individual message or for certain areas within the message. But the most effective technique for the complete protection is total message stream protection which has high level of data integrity. Whenever we are using the integrity service for connection oriented and if we are using the service for stream of message then this service give guarantees that there will be no modification, duplication or insertion type of things will occur in the process. The data damaging term is also covered in this provision. So we can say that connection oriented integrity service cover both the terms of message stream modification and denial of service. But on the other side in connectionless integrity service it only provide the security against modification of message only so this is not much preferred due to lack of security [4].

There is the modification between the service which provide recovery option and the service who unable to provide the same service. Since integrity service is directly associated to active attacks in the network and usually we focus on the detection as compare to prevention. If an integrity is going to sacrifices then the system give the notification about this violation and then there is a need of human interference or any another software to recover all the loss. Instead of there are some mechanism are available to recover the damage due to loss of integrity of information. But in this case also automated recovery is always preferred.

1.1.2.2 SYSTEM INTEGRITY:

It assured that the system will perform its function properly in a perfect manner, and there will be no unauthorized access or use of that system and system will perform all its operation in logically in correctly and reliable manner so no issue of hacker and same type of intentional operation.

1.1.3 NON-REPUDIATION:

The technical term non-repudiation of beginning denotes a service whereby the recipient is given guarantee of the messages authenticity, in the sense that the receiver can subsequently prove to a third party that the message is trustworthy even if its originator subsequently revokes it. This term is used to avoid the transmitter and recipient form rejecting the transmitted message that this message is sending from them only. So in this case after sending the message, the one user called recipient can prove that suspected transmitter send the message with the proof of sending which is available in sent block. Also on the another side when a message is reached to the destination, the sender can easily show that the suspected recipient received the sent message which is normally present in the inbox section so no one deny for the same [5][3].

1.1.4 AUTHENTICATION

In many environments, it is more essential that transmission be authenticated rather than encrypted. That is, both parties should be convinced of each other's identity. We need to establish identity and verify identity before allowing access to resources.

There are three methods we can use to authenticate someone:

- Use something you have, for example, a key or a card. The problem is that these can be stolen.
- Use something you know. Passwords and PINs (personal ID numbers) fall into these categories. These can be guessed, shared, and stolen by snooping.
- Use something you are. This involves biometrics. For example, a system may examine a user's fingerprint or iris pattern. In general, these types of systems require various types of hardware, can be costly, and are imprecise.

Authentication methods can be collective to toughen the confirmation. Using a single one of these methods is known as one-factor authentication. Using two techniques is two-factor authentication. Withdrawing cash at an ATM machine is an example of two factor authentication. To

DOI10.17148/IJARCCE.2017.63200

authenticate, you present the ATM card (something you have) and enter PIN (something you know). Most of the operating systems maintain a notion of a user identifier (user ID) which is a unique token that identifies each user on a system. Typically, systems employ a user name (a unique alphanumeric string that a user may use to identify himself / herself to the system) as well as well as a numeric user ID. The system uses the user ID to store and verify access permissions. The most frequently used way of authentication is with a simple password authentication scheme. The system prompts us for a user name and then for a password. It then looks up the name in a password table and sees if the passwords match. This is known as a reusable password since the similar password is used for each login. One major weakness here is that if any unauthorized person try and finally manages to break into the system and in this way that person can steal the entire password file [6]. There are two types of authentication facilities:

1.1.4.1 PEER ENTITY AUTHENTICATION:

This authentication provide the validation of identity of peer entity in the association. Two entities are said to be peers if they both are using the same protocol which generally used. For example we are using TCP modules in multiple communication system. This type of validation is given for the use at the foundation or when there is data transfer phase is being active. It give the surety that there is no masquerade attack or any type of unauthorized replay attack of preceding connection when the system is used for communication.

1.1.4.2 DATA ORIGIN AUTHENTICATION:

As its name indicates it is used to provide the validation of source data unit from exactly where the data comes. It never provides the security against the alteration or duplication of the data which is done by the peer entity authentication. This facility is used to provide the service to applications like mailing system where no system interacts between the sender and recipient. In the current time, cryptography becomes the battlefield for the world great mathematicians, world best hackers and the scientist who worked in the computer field. So sending of sensitive and critical information only to authorized person become the most important factor in every area where computer is used, including the secure storage of the same type of secured data [4].

1.2 CRYPTOSYSTEM

Encryption is the technique in which the readable data or original data normally called as plaintext into unreadable data or converted data called as cipher text in the term of cryptography. Plain text is in the form which is understood by any person who know the language or by a computer system. After converting this original data which is in the form of plain text into cipher text, so now no one can understand the meaning of converted data even machine cannot. By using this process we can easily share the confidential information without unauthorized expose through any channel weather the channel is secured or not as shown in figure 1.1. When we are dealing with the stored data, this is protected by many ways that may logical access or physical access control. But whenever critical information has to be shared through network no such control exist and this information open for every type of danger.

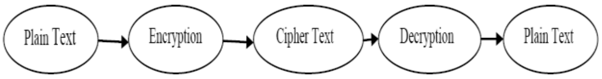

Figure 1.1 Conversion of Plain Text into Cipher

Whenever we are using the security concepts like encryption technique and decryption technique on any data or information for secure transmission or secure storage from any unauthorized access the system is known as cryptosystem. We can apply these security concepts by using any hardware component or by any specific program code which designed to do the same. This system uses the encryption and encryption algorithm to perform both the things, so complexity of the system is also depends on this algorithm. Most of the algorithm contains complex mathematics which is used in a particular sequence on the original form. And in this process this encryption method and decryption method uses a secret code value generally termed as a key (typically a long value is used for more security), which works with both the algorithm of encryption technique and decryption technique also. This key may be same or different reliant on the algorithm we are using. If we are using the larger value of key the complexity increases and it became difficult to encode the original text as shown in the figure 1.2.

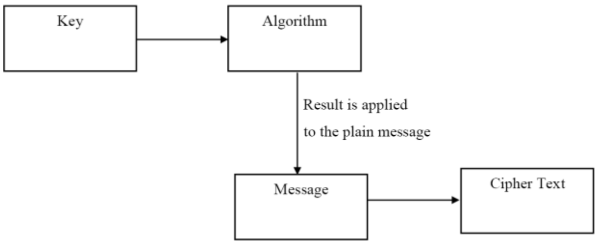

Figure 1.2 Use of Key in Cryptosystem

The algorithm covers the various rules and a set of process shows how encryption and decryption take place. Generally various encryption algorithms techniques are known publically so in some cases there is no secret part or process. As we are dealing with the security issue of critical information so algorithm work should be secret, but in general many security algorithm are known to everyone who worked in this field and it is easily understandable too. If the algorithm is publically available to apply and code then there must be something which has to keep secret. This secret is called as the key which is most vital factor in this process. This key can be anything depending on the algorithm we are using (large prime number or any large string). Every process of this type contains a key space which range of different values that can be used to made the key. The key can be random values but in between the key space provided by the algorithm technique. So large the key space more number of values are available for used in the process and if we are using the random keys then that will be more better option as it become much complex to decode the original available data and in short finding of key become more difficult. So the work of the unauthorized person who is trying to access this secret information increased. So every instance we are try to use the mixture of larger key with random behaviour so it become more and more difficult for everyone than the owner to know or guess the exact key to decipher the received data from sender. As key is the most significant part in security so it"s necessary to use the proper key as per its length and random behaviour.

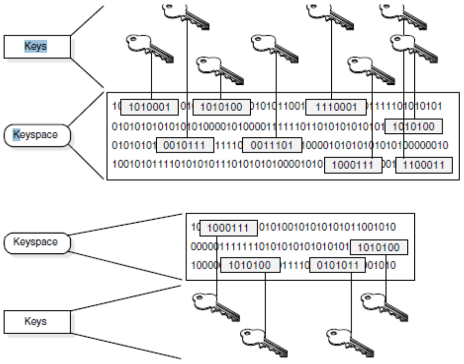

Figure 1.3 Use of Key Space in Cryptosystem [4]

If we are using large key space then our key range increases so more security will be applied. The algorithm we are going to use for encryption must use the complete key space and have to select the random value as probable. On the another hand if we make use of shorter key space then less values are available for selecting when we are going to form the key as shown in the above figure 1.3. But in this case the chances of attacker are increases and maybe he will get the key value in easier way and then he can go for conversion of cipher text to plain text so security may be compromised. Also if sender is using the larger key space then the possibility of getting the right key deceases so there is a good chance that attacker will give up and do not further try for the same process. So using of larger key space or key length for data transmission is advantageous. So many security algorithms have the arrangement to use larger length key for more secure operation. As we know RSA algorithm become very old and nearly everyone in the field of security know about all the process and operation available on this algorithm but still it is use for financial transaction by making the use of larger key and they are using this safely [4]. If in any case while transmission unauthorized person get the message when it was passed between the two authorized persons, he can see the message in unreadable form so no use of such message for this person as shown in the figure 1.4.

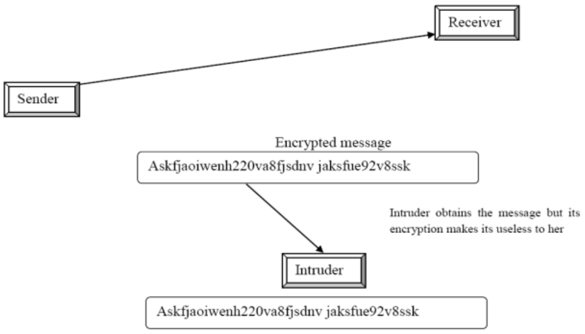
Receiver
Sender
Encrypted message
Askfjaoiwenh220va8fjsdnv jaksfue92v8ssk
Intruder obtains the message but its encryption makes its useless to her
Intruder
Askfjaoiwenh220va8fjsdnv jaksfue92v8ssk

Figure 1.4 Without the Right Key Message is Useless to Attacker

In the worst case if attacker have an idea about the algorithm which used for encryption by both the people for encryption technique and decryption technique in absence of key this data have no use to him.

1.2.1 Symmetric Cipher

- **Plaintext:** This means original plain text message or information which has to put into the algorithm as possible input for the encryption process.

- **Encryption algorithm:** This algorithm implements many substitutions as well as transformations on the original plain text provided to the algorithm.
- **Secret key:** it is another and most important input to the encryption process as per security point of view. Its value is independent of entered plain text and used algorithm. So no matching between the key, plaintext and algorithm and no combination is used for processing. As secret key in independent so the encryption algorithm create different output in the term of cipher text. Every process and operation of algorithm is depends upon the key we are using for specified process.
- **Cipher text:** This is the output of the encryption process technique which will be in unreadable form and it depends upon the original input and secret key used. If we are performing encryption the same message by using two different key then we will get the different output or cipher text. Cipher text is also the data as like plain test but it is not easily understandable.
- **Decryption algorithm:** This is called as a final step of this all operation in which we used the same algorithm which is previously used for encryption process but in the reverse order. This process takes cipher text and secret key as input and generates the original text in the form of plain text. [4] There are couples of requirements for protected use of conventional encryption process as follows:

- If we have to protect our sensitive data then in this case we a sufficient strong algorithm. We need the algorithm such that if the attacker knows which algorithm we used and he manage to get some cipher then also he is unable to decipher the cipher text or unable to find the secret key which will be helpful in further process. This requirement of good and complex algorithm is specified in stronger form. The attacker must not be in the position of decrypting the coded cipher text or secret key in any

DOI10.17148/IJARCCE.2017.63200

condition. Weather he is going to compare the different cipher text and trying to decode by guessing.

- One another important factor in this process is sender and recipient must keep the secret key in secure fashion and must not available except these two persons. If by any means someone gets the secret key and encryption algorithm in any way then everything will be expose to him and all the communication and file are available to him in readable form.

Normally everyone who is transmitting the data in this fashion assume that no one can decrypt the message by using the cipher text and knowledge of encryption algorithm. On in other way there is no need to worry about the secrecy of algorithm, just give the attention to keep the key secret as shown in the figure 1.5. This is the most advantageous to everyone when we are using the symmetric encryption and that"s why this is used globally.

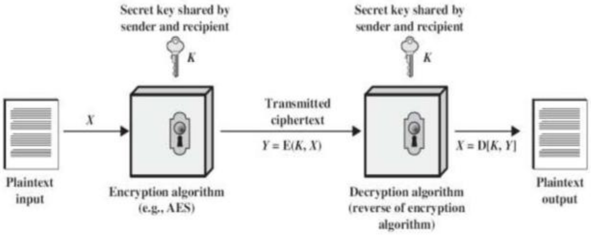
Secret key shared by sender and recipient
K
Secret key shared by sender and recipient
K
X
Transmitted ciphertext
Y = E(K, X)
X = D[K, Y]
Plaintext input
Encryption algorithm (e.g., AES)
Decryption algorithm (reverse of encryption algorithm)
Plaintext output

Figure 1.5 Simplified Model of Symmetric Encryption [3]

This means that there is no need to keep the algorithm secret so manufacturers can make the low cost chips for various data encryption algorithm for general use. This will be available worldwide and can be use jointly used in many products as per the requirement of the people. By using this symmetric algorithm principal we can easily preserve the privacy of key used in the process [4].

1.3 Cryptography

Cryptographic systems are classified into three autonomous dimensions:

- **The category of processes used for converting original plaintext to cipher text.** Nearly all the security algorithm which is used for encryption process is based on couple of principals: substitution in which every element of plain text are converted into another element by using some formulae, and transposition which rearranged the elements. The basic and important requirement is that no data must be loss and after converting whenever we are trying to get back the original data through decryption it must be done properly. It means all the operation of algorithm must be reversible in every situation. Many available systems use many steps of substitutions and transpositions.
- **The number of keys used.** if in the operating sender and receiver are depend on the same secret key then that system is called as symmetric or conventional encryption as it used the single key operation. On the other side if both the user sender and recipient use different keys then the system is said to be asymmetric key or generally known as public key encryption as multiple keys are used.
- **The method by which the original plain text is handled.** A stream cipher method process the input data constantly element by element up to the end of the entire stream. In another way called as block cipher the input of single block element is process at a time and output is produced, the output is also in the form of block itself [4].

1.4 Cryptanalysis & Brute-Force Attack

When an attacker attack on any security system to know about which key in use for encryption rather than the original text in the form of plain text from cipher text. There are two aspects of attacker to attack on the system:

- **Cryptanalysis:** This type of attack called cryptanalysis depend on the algorithm nature and need some general knowledge about the original text in the form of plaintext or if he have some pair of plain text and cipher text then it will be advantageous to him and helpful in further process. This attack makes use of algorithm characteristics for getting the original text and in some cases help to know the secret key.
- **Brute-force attack:** In this type of attack attacker tries to enter every possible key to get the original text. It is like if we have bunch of several keys and don"t know the exact key to open the lock then we are try for every key in the available bunch until the lock is going to open. So in this case attackers try at least half of possible key to get the data. So this is time consuming for attacker as every time he has to try new key if the entered key is not work until original text is not visible. And while performing this attack if that person got the success then the system will be compromised and he will get the exact key and possibly he come to know the all past and future message sent with the help of that key.

In the above we see the cryptanalysis and brute force attack by the person who try to find the secret key. Attacker normally have cipher text when he interfere between the two authorized persons so it is just consideration that he know the cipher text only but we always assume that he have an idea about the algorithm used for encryption. So if this information is available to attackers then he can go for the brute force attack in which he try all possible combination of secret keys. As we discussed previously if we used the larger key space then it is practically and real world it is impossible to try all the keys because range is very large. So attacker depend on the cipher text for every operation of encoding process, try to apply various mathematical test to get the original output. To do these entire things attacker must have some knowledge about the type of plaintext he is searching for. This type can be anything like English sentences, exe files, and financial transaction file of any accounting organization or any type of file which is sensitive for that person [4]. When attacker only attack on the cipher text then it is very easy to defend against this type of attack because in this attack attacker has minimum information about this and so it is not much easy to workout with it. In some of the cases attacker have more information about these things possible he is capable to get the plaintext messages and its encryption process. Also there is a chance of getting some idea about the original text patterns which regularly visible in the message. For example if we consider the C program in that it always start with some common code which can be easily guess or if we are dealing with the electronic fund transfer then this also have some specific pattern which is guessable. So in this way attacker is able to assume the key and go for the further operation to find all things which is required to harm or damage any one [4].

Plain text attack is also referred as word attack. If attacker is operating with encryption process of some particular type message then by experience or by working with that environment style he will get some knowledge about that message means which type of it is. Also id attacker get specific information about that message then some portion of the message may know by some processing and guessing too. For example if one user transmitted one C program code to another user and this code is critical or sensitive for both these person. And while transmitting this code, if attacker got that code and he make some research and work for hours to decode then after some work he get the knowledge that this code contain some specific header, some comment line, some specific special character and many such things which are normally available in the program code then he will come to know that this is some program code. Also if while transmitting the message between sender and receiver, sender uses the same technique, same style for encryption and similarity between the keys then structure of message will be same and attacker will see that there is same structure in generating messages. It means if pattern of message is repeated then the security will be compromised as attacker will not find more problem while decoding it [4]. The above case is never happened as only weak algorithm fail when cipher text attack happen and in general all the encryption algorithm are designed in specific manner that they must stand whenever such attack take place. The encryption method we are using will be totally secure if and only if the generated output in the form of cipher text does not contain the sufficient information to define the original plaintext in any possible way. In this case size and how much is the output text does not matter. Also there is no meaning how much span attacker has for the decryption; it is not possible to decode the information as required information is not present at that instance. There is some exception of the above method called as one time pad (OTP), no algorithm is totally secure in today‟s world. So every user who is using the algorithm for encryption process attempt to match the following principal:

- The price of decoding the cipher beats the value of the encrypted information.
- The span needed to decode the cipher surpasses the beneficial lifetime of the information.

When algorithms apply both the above principals then that algorithm is called as sufficiently secure. But in fact it is not easy to follow the above completely and difficult to calculate the effort required finding the correct solution. Everyone who worked in the security system accept the fact that if attacker identify for the specific pattern then there is more possibility that all secret information will be exposed to attacker and this will affect the past and future operation of that user. As user will never know that attacker cracked the secret key and he is getting all the information in original form. And the original user will transmitting his critical information to the sender continuously, sender will always thought that his information is secure and method is sufficient strong enough to protect all his important work [4].

The brute force attack try for every possible combination of the key by which the coded information will be decoded and in this process the attacker who is using this attack nearly half of the possible combination for getting success. Data encryption standard usually called as DES use 56 key size and on the other hand triple DES uses 168 bit key size. And at later stage in the case of advanced encryption standard called as AES uses minimum 128 bit key size which may vary as per the user. This all are the key size only but in the operation of internal part there are various permutation and substitution process are available which make the encryption process more and more complex. Many times some encryption technique uses multiple key with proper combination. After getting everything it is not easy to decrypt the data as after some time span the knowledge of such information become meaningless. But now many people uses parallel microprocessor so there may be a possibility of getting success [4].

II. CONCLUSION

In this paper the basic of cryptographic technique are explained in detail which covers the basics of data security and its characteristics. Also it gives the details about the system security and the process of encryption and decryption and gives the basic idea about the possible attacks. Later part gives introduction about more complex algorithm and techniques.

REFERENCES

[1] Alfred J. Menezes and Paul C. van Oorschot and Scott A. Vanstone: “Handbook of Applied Cryptography”, pp. 1-2, August 1996

[2] Rafael pass and Abhi shelat: “A course in cryptography”, pp. 1-2, January 2010

[3] Darrel Hankerson, Alfred Menezes and Scott Vanstone: “Guide to Elliptic Curve Cryptography”, ISBN 0-387-95273-X, Springer, 2004

[4] William Stallings: ”Cryptography and Network Security Principles and Practice”, ISBN 13:978-0-13-609704-4, 5th Edition, Pearson Publication, pp. 10-13, 2011

[5] Bart Preneel: “Analysis and Design of Cryptographic Hash Functions”, pp. 24-26, February 2003

[6] Paul Krzyzanowski: “Cryptographic communication and authentication”, Rutgers University – CS 417: Distributed Systems, pp. 1-2, 1997

[7] Vishwa Gupta, Gajendra Singh and Ravindra Gupta : “Advance cryptography algorithm for improving data security” ISSN: 2277 128X, International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE), Volume 2, Issue 1, January 2012

[8] Atul Kahate: “Cryptography and Network Security”, second edition, Tata McGraw-Hill, pp.43-53, 2006.

DOI10.17148/IJARCCE.2017.63200