An Analysis of the Attack on RSA Cryptosystem through Formal Methods

Sachin Upadhyay, Yashpal Singh, Amit Kumar Jain

Abstract:- Communication is the basic process of exchanging information. The effectiveness of computer communication is mainly based on the security aspects whether it is through internet or any communication channel. The aim of this paper is based on analyzing the results given by Wiener's, who says that if the private exponent d used in RSA cryptosystem is less than n^.292 than the system is insecure. We will focus on the result given by Weiner's and try to increase the range of private exponent d up to n^0.5. As n is the product of p & q (which are the relative prime numbers). This paper also aims at considering the different factors that affects the performance of encryption algorithms so as to make our information more secure over the network.

Index Terms:- Conjunctive Normal Form (CNF), Cryptanalysis, RSA Algorithm, , SAT Solver tool.

I. INTRODUCTION To CRYPTOGRAPHY

It is the study of Secret (crypto)-Writing (graph). It is the science or art of encompassing the principles and methods of transforming an intelligible message into one that is intelligible and then transforming the message back to its original form. Cryptography is a complex and mathematically challenging field of study. It involves taking some data or message and obfuscating (make unclear or hard to understand) it so that it is unreadable by parties that the message was not intended. Before the message becomes encrypted it is referred to as the plain text. Once a message becomes encrypted it is then referred to as the cipher text.

A. Encryption

All encryption algorithms are based on two general principles: substitution, in which each element in the plaintext is mapped into another element and transposition, in which elements in plaintext are rearranged. Encryption plays a very vital role for keeping our data secured over the network & also reflects the capability of the sender or the algorithms used for the data encryption.

B. The keys used

If both sender and receiver use the same key, the system is referred to as symmetric, single-key, secret-key, or conventional encryption. On other way the system is referred to as asymmetric, two-keys, or public-key encryption. Symmetric requires that the secret key be known by the party encrypting the data and the party decrypting the data. Asymmetric allows for distribution of your public key to anyone with which they can encrypt the data they want to send securely and then it can only be decoded by the person having the private key.

C. Processing of plaintext

Processing of plaintext mainly depends on the capability of the algorithms used for the encryption of data. A block cipher processes the input one block of elements at a time, producing an output block for each input block. A stream cipher processes the input elements continuously, producing output one elements at a time, as it goes along.

The authors of the accepted manuscripts will be given a copyright form and the form should accompany your final submission.

II. APPROACH

This paper is mainly concerned with the idea of how properly & securely you can use RSA Algorithm for the secure communication over the network. We will convert the RSA problem in sat-satisfibility condition and solve them using sat solver. We are using this approach to analyze the different types of attack on RSA cryptosystem.

A.RSA Algorithm

It is based on the presumed difficulty of factoring large integers, the factoring problem. The RSA scheme is a block cipher in which the plaintext & ciphertext are integers between 0 and n-1 for some n. A typical size for n is 1024 bits, i.e., n is less than 21024 Units.

Select p, q p and q both prime, $p \neq q$

Calculate $n = p \times q$

Encryption

Plain text $M < n$

Cipher text $C = M^e \bmod n$ (1)

Decryption

Cipher text C

Plain text $M = C^d \bmod n$ (2)

Where e is public key

d is private key

So now we need to convert the RSA problem in sat satisfibility condition and then use it as an input for the sat solver.

B. Attacks on RSA

Timing attack: Timing attacks are applicable not just to RSA, but to other public-key cryptography systems. This attack is alarming for two reasons: It comes from a completely unexpected direction and it is a cipher text only attack.

Manuscript received on April 26, 2012.

Sachin Upadhyay, Department of Mathematical Sciences and Computer Applications, Bundelkhand University, Department of Computer Science & Engineering, Bundelkhand Institute of Engineering & Technology Jhansi, India, 09452736650, (Email:sachinupadhyay2010@yahoo.co.in).

Yashpal Singh, Department of Computer Science & Engineering & Technology Jhansi, India, 09415030602, (Email: yash_biet@yahoo.co.in).

Amit Kumar Jain, Department of Mathematical Sciences and Computer Applications, Bundelkhand University, Department of Computer Science & Engineering, Bundelkhand Institute of Engineering & Technology Jhansi, India,09452369851, (Email:diggijain@gmail.com).

Brute force: The attacker tries every possible key on a piece of cipher text until an intelligible translation into plaintext is obtained. This involves trying all possible private keys

C. Classical encryption techniques

The technique enables us to illustrate the basic approaches to conventional encryption today. The two basic components of classical ciphers are substitution and transposition. Then other systems described that combines both substitution and transposition are discussed below

1. Substitution techniques

In this technique letters of plaintext are replaced by numbers or by symbols. If plaintext is viewed as a sequence of bits, then substitution involves replacing plaintext bit patterns with cipher text bit patterns.

a) Caesar Cipher

Caesar Cipher replaces each letter of the message by a fixed letter a fixed distance away e.g. uses the third letter on and repeatedly used by Julius Caesar.

For example:

Plaintext:
IF WE DO NOT PROPERLY PROTECT THE USERS DATA WE CAN SIMPLY HIDE BEHIND THE DMCA IF SOMEONE NOTICES!!

Ciphertext:
RU DV WL MLG KILKVIOB KILGVXG GSV FHVIH WZGZ DV XZM HRNKOB SRWV YVSRMW GSV WNXZ RU HLNVLMV MLGRXVH!!

b) Monoalphabetic Ciphers

With only 25 possible keys, the Caesar cipher is far from secure. A dramatic increase in the key space can be achieved by allowing an arbitrary

Example:

Plain Text: h i j k ….
Cipher Text: K L M N ….
Here a single cipher alphabet is used per message. If, instead the cipher line can be any permutation of the 26 alphabetic characters, then there are 26! or greater than $4x10^{26}$ possible keys, this would seems to eliminate brute-force attack techniques for cryptanalysis.

2. Transposition Techniques

In this technique a different kind of mapping is achieved by performing some sort of permutation on the plaintext letters. The simplest use of transposition is rail fence technique in which the plaintext is written down as a sequence of diagonals and then read off as a sequence of rows.

Cryptanalysis the objective of attacking an encryption system is to recover the key in use rather than simply to recover the plaintext of a single cipher text. There are two general approaches to attacking a conventional encryption scheme:

3. Cryptanalysis

Cryptanalytic is the process of attempting to discover either plaintext or keys with a very little auxiliary information. It is the art of defeating cryptographic security systems, and gaining access to the contents of encrypted messages, without being given the cryptographic key. Cryptanalysis under mathematical analysis also includes the study of side-channel attacks that do not target weaknesses in the cryptographic algorithms themselves, but instead exploit weaknesses in their physical implementation or software implementation.

4. Brute-force attack

The attacker tries every possible key on a piece of cipher text until an intelligle translation into plaintext is obtained. On an average 50% efforts are sufficient to recover the secret key.

III. SAT SOLVER TOOLS

Mostly SAT solvers are based on the Davis-Putnam-Logemann-Loveland (DPLL) algorithm and require the input formula to be in Conjunctive Normal Form (CNF). However, typical formulas that arise in practice are non-clausal, that is, not in CNF. Converting a general formula to CNF introduces overhead in the form of new variables and may destroy the structure of the initial formula, which can be useful to check satisfiability efficiently. Boolean satisfiability (SAT) solvers are used heavily in verification tools as decision procedures for propositional logic. In complexity theory, the satisfiability problem (SAT) is a decision problem, whose instance is a Boolean expression written using only AND, OR, NOT, variables, and parentheses. A formula of propositional logic is said to be satisfeable if logical values can be assigned to its variables in a way that makes the formula true.

IV. CONJUNCTIVE NORMAL FORM

In Boolean logic, a formula is in conjunctive normal form (CNF) if it is a conjunction of clauses, where a clause is a disjunction of literals, where a literal and its complement cannot appear in the same clause. As a normal form, it is useful in automated theorem proving. It is similar to the product of sums form used in circuit theory.

All conjunctions of literals and all disjunctions of literals are in CNF, as they can be seen as conjunctions of one-literal clauses and conjunctions of a single clause, respectively. As in the disjunctive normal form (DNF), the only propositional connectives a formula in CNF can contain are and, or, and not. The not operator can only be used as part of a literal, which means that it can only precede a propositional variable.

V. CONVERSION INTO CNF

Every propositional formula can be converted into an equivalent formula that is in CNF. This transformation is based on rules about logical equivalences: the double negative law, De Morgan's laws, and the distributive law.

The generated formula is:

$(X_1 \vee \ldots \vee X_{n-1} \vee X_n) \wedge (X_1 \vee \ldots \vee X_{n-1} \vee Y_n) \wedge \ldots \wedge (Y_1 \vee \ldots \vee Y_{n-1} \vee Y_n)$

This formula contains 2^n clauses; each clause contains either X_i or Y_i for each i.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

These transformations are guaranteed to only linearly increase the size of the formula, but introduce new variables. For example, the above formula can be transformed into CNF by adding variables Z1,….., Zn as follows:

$(Z_1 \vee \ldots . \vee Z_n) \wedge (\neg Z_1 \vee X_1) \wedge (\neg Z_1 \vee Y_1) \wedge \ldots . \wedge (\neg Z_n \vee X_n) \wedge (\neg Z_n \vee X_n)$

An interpretation satisfies this formula only if at least one of the new variables is true. If this variable is Zi, then both Xi and Yi are true as well. This means that every model that satisfies this formula also satisfies the original one.

VI. POINT AT ISSUE

This paper focus on the work done by DAN BONEH & GLENN DURFEE regarding the range of private exponent d used in RSA cryptosystem. They proved that if d is less than N^.292 that the system is insecure, what we are we doing in this paper is that we are increasing the range of private exponent d up to 0.5 i.e. if d is less than N^0.5 than the system will be insecure.

VII. ASYMMETRIC KEY ALGORITHMS

This type of algorithms is public-key that is to say the key that is used to encrypt the message is different from the key used to decrypt the message. The encryption key, known as the Public key is used to encrypt a message, but the message can only be decoded by the person that has the decryption key, known as the private key. The PKC has the advantage that there is no need for exchange of keys via a secure channel between two users who wish to communicate with each other. Also the recipient can make their public key widely available anyone wanting to send them a message uses the algorithm and the recipient's public key to do so, only the recipient, with the private key can decrypt the message.

RSA Cryptosystem

The RSA scheme makes use of an expression with exponentials. Plaintext is encrypted in blocks, with each block having a binary value less than some number n. That is, the block size must be less than or equal to log2 (n); in practice the block size is I bits, Encryption and decryption are of the following form, for some plaintext block M and cipher text block C:

$$C = M^e \bmod n \qquad (3)$$

$$M = C^d \bmod n = (M^e)^d \bmod n$$
$$= M^{ed} \bmod n \qquad (4)$$

The private key consists of {d , n} and the public key consists of {e , n} . Suppose that user A has published its public key and that user B wishes to send the message M to A. then B calculates C = M^e (mod n) and transmits C. On receipt of this cipher text, user A decrypts by calculating M = C^d (mod n)

RSA Example

A.RSA Encryption

1. Randomly choose two prime numbers: *p* and *q*.
   p = 127 *q* = 211
2. Compute *N* = *pq*
   N = 127 * 211 = 26,797
3. Compute *N'* = (*p* – 1) (*q* – 1)
   N' = (127 – 1) * (211 – 1) = 26,460
4. Choose *e* > 1 such that gcd(*e* , *N'*) = 1
   (any *e* that is relatively prime to *N'*)
   i.e. Any prime *e* = 13,379
5. Compute d as the multiplicative inverse of e, mod N'
   d = 11,099
   Since
   e (13,379) * d (11,099) mod N (26,460) = 1
6. Destroy *p*, *q*, *N'*
7. Use *e* and *N* to encrypt a message. Use *d to* decrypt.

Encryption Example

Encrypt (int M) = Me (mod N)
M is the message to be sent
Encrypt (10,237) = 10,237
13,379(mod 26797) = 8422
8422 is the encrypted M

Decryption Example

Decrypt(*R*) = *Rd*(mod *N*)
R is the received encrypted message
Decrypt (8422) = 842211099(26797) =10,237

Example of attack on RSA

Here we will group the characters into blocks of three and compute a message representative integer for each block.

ATTACK*AT*SEVEN= ATT ACK *AT *SE VEN

In the same way that a decimal number can be represented as the sum of powers of ten, e.g.
$135 = 1 \times 10^2 + 3 \times 10^1 + 5$
We could represent our blocks of three characters in base 26 using A=0, B=1, C=2, ..., Z=25

$ATT = 0 \times 26^2 + 19 \times 26^1 + 19 = 513$
$ACK = 0 \times 26^2 + 2 \times 26^1 + 10 = 62$
$XAT = 23 \times 26^2 + 0 \times 26^1 + 19 = 15567$
$XSE = 23 \times 26^2 + 18 \times 26^1 + 4 = 16020$
$VEN = 21 \times 26^2 + 4 \times 26^1 + 13 = 14313$

For this example, to keep things simple, we'll not worry about numbers and punctuation characters, or what happens with groups AAA or AAB.

In this system of encoding, the maximum value of a group (ZZZ) would be 26^3-1 = 17575, so we require a modulus n greater than this value.

1. We generate primes p=137 and q=131
2. n = p.q = 137.131 = 17947
   phi = (p-1)(q-1) = 136.130 = 17680
3. Select e = 3
   check gcd(e, p-1) = gcd(3, 136) = 1, and
   check gcd(e, q-1) = gcd(3, 130) = 1
4. Compute $d = e^{-1} \bmod phi = 3^{-1} \bmod 17680 = 11787$.
5. Hence Public Key, (n, e) = (17947, 3) Private Key (n, d) = (17947, 11787)

To encrypt the first integer that represents "ATT", we have
$c = m^e \bmod n = 513^3 \bmod 17947 = 8363$.
We can verify that our private key is valid by decrypting
$m' = c^d \bmod n$
$=8363^{11787} \bmod 17947 = 513$.

Overall, our plaintext is represented by the set of integer m

(513, 62, 15567, 16020, 14313)

We compute corresponding ciphertext integers $c = m^e \bmod n$,

(8363, 5017, 11884, 9546, 13366)

The security of RSA algorithm depends on the ability of the hacker to factorize numbers. New, faster and better methods for factoring numbers are constantly being devised. The Trent best for long numbers is the Number Field Sieve. Prime Numbers of a length that was unimaginable a mere decade ago are now factored easily. Obviously the longer a number is, the harder is to factor, and so the better the security of RSA.

VIII. CONCLUSION

In this paper we give the substantial improvement to Wiener's, Dan Boneh & Glenn Durfee results. Our results are based on the seminal work of Coppersmith Wiener describes a number of clever techniques for avoiding his attack while still providing fast RSA signature generation. One such suggestion is to use a large value of e. Indeed, Wiener's attack provides no information as soon as e > N^1.5. In contrast, our approach is effective as long as e < N^1.875. Consequently, larger values of e must be used to defeat the attack. Results are successfully tested for some of the sets of prime numbers not on very large prime numbers. A conclusion can be drawn that the selection of prime numbers plays a very vital role regarding the security of RSA algorithm; most importantly a proper implementation of RSA algorithm is definitely mandatory.

REFERENCES

1. D. Boneh. Twenty Years of Attacks on the RSA. Notices of the American Mathematical Society, vol 46(2):203–213, 1999.
2. R. L. Rivest, A. Shamir, and L. Adleman. A method for obtaining digital signatures and public key cryptosystems. Commun. of the ACM, 21:120-126, 1978.
3. Brown, Lawrie. "Classic Cryptography". 22 Feb 1996.
4. SANS Institute. "SANS GIAC Training and Certification".URL:http://www.sans.org/giactc/GIAC_certs.htm (24 Nov 2001)
5. Cryptography & Network Security by William Stallings fourth edition.
6. Cryptography & Network Security by Kumar Manoj, Krishna's Prakashan Media (P) Ltd.

AUTHORS PROFILE

Mr. Sachin Upadhyay, received a degree in Master of Computer Applications from Bundelkhand University, Jhansi in 2008. He is a research scholar & working as a Lecturer at Bundelkhand University, Jhansi (U.P). His area of interest is operating system, network security.

Dr. Yashpal Singh, received a Master degree in ME from Allahabad University, Allahabad in 2000. Currently he is Associate Professor working as a Head in the Department of Information Technology at Bundelkhand Institute of Engineering and Technology, Jhansi. He received Ph.D (CS) degree in from Bundelkhand University, Jhansi in 2008 His area of interest are OOPS, Data Structure, Image Processing, Neural Networks.

Mr. Amit Kumar Jain received a degree in Master of Computer Applications from Bundelkhand University, Jhansi in 2009. He is a research scholar & working as a Lecturer at Bundelkhand University, Jhansi (U.P). His area of interest is operating system, network security.