



Information Theory and Public Key Cryptosystems

ANTHONY M. GAGLIONE

*Identification Systems Branch
Radar Division*

May 29, 1987

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE				
1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED		1b. RESTRICTIVE MARKINGS		
2a. SECURITY CLASSIFICATION AUTHORITY		3. DISTRIBUTION / AVAILABILITY OF REPORT		
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE		Approved for public release; distribution unlimited.		
4. PERFORMING ORGANIZATION REPORT NUMBER(S) NRL Report 9031		5. MONITORING ORGANIZATION REPORT NUMBER(S)		
6a. NAME OF PERFORMING ORGANIZATION Naval Research Laboratory	6b. OFFICE SYMBOL (If applicable) Code 5350	7a. NAME OF MONITORING ORGANIZATION		
6c. ADDRESS (City, State, and ZIP Code) Washington, DC 20375-5000		7b. ADDRESS (City, State, and ZIP Code)		
8a. NAME OF FUNDING / SPONSORING ORGANIZATION Naval Air Systems Command	8b. OFFICE SYMBOL (If applicable) APC-209	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER		
8c. ADDRESS (City, State, and ZIP Code) Washington, DC 20361		10. SOURCE OF FUNDING NUMBERS		
		PROGRAM ELEMENT NO. 64211N	PROJECT NO.	TASK NO. W1253
				WORK UNIT ACCESSION NO. DN180-248
11. TITLE (Include Security Classification) Information Theory and Public Key Cryptosystems				
12. PERSONAL AUTHOR(S) Gaglione, * A. M.				
13a. TYPE OF REPORT Final	13b. TIME COVERED FROM 6/15/86 TO 8/14/86	14. DATE OF REPORT (Year, Month, Day) 1987 May 29	15. PAGE COUNT 11	
16. SUPPLEMENTARY NOTATION *Affiliation: Department of Mathematics, U.S. Naval Academy, Annapolis, MD 21402				
17. COSATI CODES		18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)		
FIELD	GROUP	SUB-GROUP	Information theory Public key	
			Cryptosystem Computational complexity (Continues)	
19. ABSTRACT (Continue on reverse if necessary and identify by block number)				
<p>Shannon [1] has defined the unicity distance of a random cipher as the point where there is no uncertainty over which key was used for enciphering. The unicity distance is given as a value N where $N =$ cryptogram length in characters. The usual issue for classical cryptography is: given ciphertext (and possibly corresponding plaintext) under the assumption of a random cipher, is this information sufficient on the average to determine the key. Here, if we let M denote the random variable (defined as the number of keys that will decipher a given intercepted cryptogram into a meaningful message), it turns out that M has a binomial distribution [2]. Meyer and Matyas [2] have expanded Shannon's approach to unicity distance by using information theory. They make no assumption about the distribution of M so their approach applies to cryptosystems in general. This paper applies this method to public key cryptography. In particular, we consider the RSA (Rivest-Shamir-Adelman) [2] cryptosystem which is probably the most widely known public key system.</p> <p style="text-align: right;">(Continues)</p>				
20. DISTRIBUTION / AVAILABILITY OF ABSTRACT <input checked="" type="checkbox"/> UNCLASSIFIED/UNLIMITED <input type="checkbox"/> SAME AS RPT. <input type="checkbox"/> DTIC USERS		21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED		
22a. NAME OF RESPONSIBLE INDIVIDUAL E. Vegh		22b. TELEPHONE (Include Area Code) (202) 767-3481	22c. OFFICE SYMBOL Code 5350	

18. SUBJECT TERMS (Continued)

Entropy
Unicity distance
Plaintext
Ciphertext
RSA system

19. ABSTRACT (Continued)

The motivation for this work was to investigate the complexity of RSA. No new research has been conducted into the information theoretic approach to cryptosystem security since Shannon's work. This report shows that the present state of this theory is inadequate to handle public key cryptosystems like the RSA system. It is hoped that with further research on this topic (e.g., with making proper assumptions), the information theoretic approach can be made to at least give some kind of theoretical bound for the security of public key systems.

CONTENTS

INTRODUCTION	1
PUBLIC KEY CRYPTOGRAPHY	2
INFORMATION THEORY AND PUBLIC KEY	2
CONCLUSIONS	6
REFERENCES	6

INFORMATION THEORY AND PUBLIC KEY CRYPTOSYSTEMS

INTRODUCTION

In 1949, C.E. Shannon [1] laid the foundations for a general, theoretical analysis of secrecy systems. This paper establishes (as far as the author knows) the first published formalization of the intuitive notion of a secrecy system, hereinafter called a cryptosystem. At the same time, Shannon introduces the concept of an information theoretic analysis of cryptosystems to evaluate the theoretical security of such systems.

Shannon showed that the plaintext can be completely and unambiguously recovered if and only if the redundancy of the plaintext is at least as large as the sum of the noise and the information content of the key. In the present report, the redundancy and information content of the key have to be taken in the quantitative sense defined in Ref. 2. Reference 2 emphasizes that this is a theoretical lower limit. Where a cryptosystem is breakable in principle according to this condition, it could still be practically secure. This is possible because the number of operations required to determine the key might be excessively large, i.e., a cryptosystem might have good practical security but not necessarily good theoretical security.

Public key cryptosystems represent an extreme case in the relationship between key size and computational complexity for the cryptoanalyst. There is no secret key information to detect whatsoever, but instead there is a formidable computational problem. This paper applies these information theoretical results to public key cryptosystems, in particular to the RSA (Rivest-Shamir-Adelman) system. Information theory measures the amount of information concerning plaintext or key that is contained in a cryptogram. This report does not consider the computational complexity. A cryptosystem that is safe from the information theoretic analysis is safe if the amount of computational resources available to the cryptoanalyst is assumed to be bounded.

A general treatment of the complexity theory approach to cryptology seems to be very difficult to handle. Some relevant references and an attempt to address these problems are contained in Refs. 3 and 4. One of the main objectives of this research is to try to address this problem. More specifically, it is hoped that with further research the present treatments of information theory can be suitably modified so as to form a basis for the complexity problem for public key cryptosystems. This report shows that although the current approach of information theory does give some security bounds for classical cryptosystems, it is not suitable for public key systems.

It is interesting to note that almost 40 years have elapsed since the publication of Ref. 1 and there has been little additional research on the information theory approach to cryptosystems [2]. Moreover, we note that treatments from both Refs. 1 and 2 lack the precision necessary to be considered rigorous mathematical analyses. For example, let ud be the unicity distance of a given cryptosystem. Then it is suggested that ud is a sort of threshold separating cryptograms that can be cryptanalyzed from those that cannot. Meyer [2] points out that this is not exactly correct.

Here, we follow the notation and terminology of Ref. 2. We assume that the reader is familiar with the results from Ref. 2.

PUBLIC KEY CRYPTOGRAPHY

Public key cryptosystems do not require a priori distribution of keying material to the two parties who wish to communicate; this is their main advantage. As an example of a public key scheme, we consider the RSA system. Briefly, the method works as follows: Two large primes p and q are chosen, and $n = p \cdot q$ is computed. Letting $m = \phi(n) = (p - 1)(q - 1)$, a large random number y is chosen such that $\gcd(y, m) = 1$. This step guarantees the existence of a unique integer x , $0 < x < m$, such that $x \cdot y \equiv 1 \pmod{m}$. If we let X and Y represent the plaintext and the corresponding ciphertext respectively, then the enciphering process consists of computing $E_{PK}(X) \equiv X^y \pmod{n}$ where $E_{PK}(X) = Y$ is such that $0 \leq Y \leq n - 1$. The encryption scheme is made public by announcing n and y . Using Meyer's terminology, let the public key $PK = y$ and the secret key $SK = x$. To decode the ciphertext $Y = E_{PK}(X)$, one computes

$$D_{SK}(Y) \equiv Y^{SK} \pmod{n}$$

where

$$0 \leq D_{SK}(Y) \leq n - 1.$$

Thus,

$$D_{SK}(Y) \equiv X^{xy} \pmod{n}$$

$$\equiv X \pmod{n}$$

by Euler's Theorem.

INFORMATION THEORY AND PUBLIC KEY

Information theory applies a numerical measure of information to a message. This measure is usually given in terms of bits. Following Meyer, we let $X = \{x_1, x_2, \dots, x_r\}$ denote the message space and associate with each message a probability $P(x_i) = p_i$ where $\sum_{i=1}^r p_i = 1$. The information associated with $x_i \in X$ is $-\log_2(p_i)$ bits. If each message is equally likely, every x_i has information value $\log_2 r$. For the set X , the average information per message is defined to be the *entropy of X* , denoted by $H(X)$ and defined by

$$H(X) = \sum_{i=1}^r -(p_i) \log_2(p_i).$$

Here, $H(X)$ can be interpreted as a measure of the uncertainty over which message the sender will select and transmit to the receiver. If each message is equally likely, there is a maximum uncertainty concerning which message will be transmitted, and $H(X)$ assumes its maximum $H(X) = \log_2 r$. On the other extreme, if there is no uncertainty over which message will be transmitted, $H(X) = 0$. Thus $H(X)$ assumes values in the interval 0 to $\log_2 r$.

If we let K denote a set of keys each having an associated probability of occurrence, we can then define $H(K)$ in the same manner as $H(X)$ above. Thus in cryptanalysis, $H(X)$ and $H(K)$ can be interpreted as the analyst's prior information regarding which message and key are selected for encipherment.

Next, we list some information measures used in this study. Here U , V , and W are finite sets whose elements have been assigned probabilities such that

$$\sum_u P(u) = \sum_v P(v) = \sum_w P(w) = 1.$$

(Here \sum_u means the summation is over all $u \in U$.)

1. Conditional entropy of U given $v \in V$:

$$H(U | v) = - \sum_u P(u | v) \log_2 P(u | v)$$

2. Equivocation of U given V :

$$H(U | V) = - \sum_v P(v) H(U | v)$$

3. Entropy of U and V :

$$H(U, V) = - \sum_{u, v} P(u, v) \log_2 P(u, v)$$

4. (a) Equivocation of U given V and W :

$$H(U | V, W) = - \sum_{u, v, w} P(u, v, w) \log_2 P(u | v, w)$$

- (b) Equivocation of U and V given W :

$$H(U, V | W) = - \sum_{u, v, w} P(u, v, w) \log_2 P(u, v | w)$$

5. Entropy of U , V , W :

$$H(U, V, W) = - \sum_{u, v, w} P(u, v, w) \log_2 P(u, v, w).$$

We also need the identity

$$(*) \quad H(U | V, W) + H(V | W) = H(V | U, W) + H(U | W).$$

To prove (*), let us first show

$$(*)' \quad H(U, V) = H(U | V) + H(V).$$

By definition (3 above),

$$\begin{aligned}
 H(U, V) &= - \sum_{u, v} P(u, v) \log_2 P(u, v) \\
 &= - \sum_{u, v} P(u | v) P(v) [\log_2 P(u | v) + \log_2 P(v)] \\
 &= - \sum_v P(v) H(U | v) - \sum_v P(v) \log_2 P(v) \\
 &= H(U | V) + H(V)
 \end{aligned}$$

by 2 above. This proves (*').

Next we claim

$$(*)'' \quad H(U, V, W) = H(U | V, W) + H(V, W).$$

Proceeding in a manner similar to the proof of (*'), by definition (5 above)

$$\begin{aligned}
 H(U, V, W) &= - \sum_{u, v, w} P(u, v, w) \log_2 P(u, v, w) \\
 &= - \sum_{u, v, w} P(u, v, w) \log_2 (P(u | v, w) \cdot P(v, w)) \\
 &= - \sum_{u, v, w} P(u, v, w) \log_2 P(u | v, w) - \sum_{v, w} P(v, w) \log_2 P(v, w) \\
 &= H(U | V, W) + H(V, W)
 \end{aligned}$$

by 4(a) and 3 above. This proves (*'').

Finally, we may now verify (*).

$$H(U | V, W) + H(V | W) = H(U, V, W) - H(V, W) + H(V | W)$$

by (*''). But $H(U, V, W)$ clearly equals $H(V, W, U)$, so we get

$$H(U | V, W) + H(V | W) = H(V, W, U) - H(V, W) + H(V | W).$$

By (*'') applied to $H(V, W, U)$ we get,

$$H(V, W, U) = H(V | W, U) + H(W, U).$$

Thus

$$H(U | V, W) + H(V | W) = H(V | W, U) + H(W, U) - H(V, W) + H(V | W).$$

But (*)' gives $H(V, W) = H(V | W) + H(W)$, so

$$\begin{aligned} H(U | V, W) + H(V | W) &= H(V | W, U) + H(W, U) - H(W) \\ &= H(V | W, U) + H(U | W) \end{aligned}$$

by (*)' again. This verifies (*).

To apply these information measures of theoretical secrecy to the case of public key cryptosystems, we need a definition of *unicity distance* for the case when both plaintext and corresponding ciphertext are available for analysis since this is the case in such systems. But Meyer [2, pp. 631-632] has given such a definition.

Following Meyer, we rewrite equation (*) as

$$H(K | Y, X) + H(Y | X) = H(Y | K, X) + H(K | X)$$

where Y is the cyptogram space. But since a knowledge of $k \in K$ and $x \in X$ determines $y = E_K(x) \in Y$,

$$H(Y | K, X) = 0.$$

Also keys and messages are selected independently, so $H(K | X) = H(K)$. Thus (*) becomes

$$H(K | Y, X) = H(K) - H(Y | X).$$

So there is no uncertainty regarding the key that is used, we must have $H(K | Y, X) = 0$. Of course in doing this, the usual interpretation is that we are dealing with a classical cryptosystem. But here we will also allow public key systems; even though the public key, PK , is known, there is still uncertainty about the secret key, SK . Thus the unicity distance ud of a cryptosystem in which both plaintext and ciphertext are available is defined as the value of N (= cryptogram length) for which

$$H(K) - H(Y | X) = 0,$$

provided such an N exists.

We turn now to the special case of public key cryptosystems like the RSA system. In such systems given the plaintext $x \in X$, the cryptogram $Y = E_K(x) \in Y$ is determined thus: $H(Y | X) = 0$. So the defining equation for the ud becomes (in the case of public key systems)

$$H(K) = 0.$$

This is not completely incorrect because the public key, PK , is known. However, it does not consider that the secret key, SK , is *not* determined, so $H(K)$ should not be 0.

CONCLUSIONS

The preceding discussion shows that the present state of information theory is not adequate to handle public key cryptosystems; further research should be conducted in this area. It would be valuable to have a more mathematically precise treatment of this information theoretic approach to cryptosystems. Neither Shannon's nor Meyer's formulations are sufficiently precise. A more precise treatment itself may help to handle the case of public key cryptosystems.

More specifically, if we are going to adapt these methods to public key cryptosystems, we should determine what assumptions about the set of keys K are appropriate and realistic. Although theoretically the set of all possible keys for a public key system like RSA is infinite, given one's computing capability only a finite number can actually be used. Moreover, further assumptions about the probabilities associated with these keys can be made. Clearly, if the key is too small, the system will not have sufficient cryptographic strength. Thus such keys could be assigned low probabilities (or probability 0). Also, the fact that the key for enciphering (PK) is known but the key for deciphering (SK) is unknown should be taken into account. This information theoretically says that if K_1 is in the set of PKs and K_2 is in the set of SKs, then $H(K_1) = 0$ but $H(K_2) \neq 0$. (It may be that K_1 and K_2 are the same set, but they need to be distinguished for application to public key systems.) As mentioned earlier, it is hoped that a definition of *unicity distance* can be formulated that could be suitable for handling public key cryptosystems. Moreover, such a concept could serve as a theoretical lower bound for describing the complexity of such systems.

REFERENCES

1. C.E. Shannon, "Communication Theory of Secrecy Systems," *Bell Syst. Tech. J.*, **28**, 656-715 (1949).
2. C.H. Meyer and S.M. Matyas, *Cryptography: A New Dimension in Computing Data Security* (John Wiley and Sons, New York, 1982).
3. A.M. Gaglione, "Some Complexity Theory for Cryptography," NRL Report 9024, Jan. 1987.
4. H.C. Williams, "Computationally 'Hard' Problems as a Source for Cryptosystems," *Secure Communications and Asymmetric Cryptosystems*, AAAS Selected Symposium, (1982), Vol. 69, pp. 11-39.