To appear in *Proc. of 4th IMA Conference on Cryptography and Coding*, Cirencester, England, Dec. 13-15, 1993 (invited paper).

The Role of Information Theory in Cryptography

Ueli M. Maurer

Department of Computer Science
ETH Zürich
CH-8092 Zürich, Switzerland

Abstract. This paper reviews the relations between information theory and cryptography, from Shannon's foundation of information theory to the most recent developments in unconditionally-secure key-agreement protocols. For a long time, information theory has mainly been used in cryptography to prove lower bounds on the size of the secret key required to achieve a certain level of security in secrecy and authentication systems. More recent results on a slightly extended model suggest that perfect secrecy is practically possible with only a short secret key, thus apparently contradicting Shannon's lower bound on the key size of a perfect cipher.

1. Introduction

From a scientific point of view, one of the most interesting and challenging problems in cryptography is the design of systems or protocols whose security can be proven rigorously. There exist many approaches to provable security in the literature, but most of these are not truly satisfactory, as will be explained below.

In order to prove the security of a cryptographic system, a definition of security or, alternatively, of breaking the system must be given. Furthermore, the assumptions about the adversary's available information and about his computing power must be stated. Whether a system with provable security is satisfactory from a theoretical and practical viewpoint depends in a crucial manner on three aspects: (1) on the acceptability and generality of the definition of security; (2) on how realistic the two assumptions are; and (3) on the practicality of the system. All previous approaches to provable security fail in at least one of these aspects, and we believe that the approach for generating cryptographic keys discussed in Section 4.2 comes closest to a realistic provably secure system.

For instance, it is trivial to "prove" the security of a cipher if we define security (irrelevantly) to mean that an adversary is unable to square a circle with straightedge and compass ([14], p. 216). It is similarly trivial to prove that an adversary cannot obtain any information about the plaintext for a system in which the legitimate receiver cannot either, or if one assumes that the adversary is unable to even receive the ciphertext.

In order to avoid all possible arguments about the assumptions about the adversary's available information, one normally assumes in cryptography that an adversary has complete information about the design of the system (this is known as Kerkhoff's assumption) and that he can receive all messages transmitted over insecure channels.

There are two possible types of assumptions about the adversary's computing power: A system is called computationally-secure if it is secure against an adversary with reasonably bounded computational resources and it is called information-theoretically secure if it is secure even against adversaries with infinite computing power. There are two problems with the first type of assumption. First, one needs to specify a model of computation and, for instance in view of analog implementations of neural networks or, more severely, of the potential realizability of quantum computers [16], it is not clear whether a Turing machine or any standard discrete computer model is sufficiently general. In other words, one could argue that even the assumption that an adversary has the computing power corresponding to 10^{20} of the newest-generation CRAY computers is not satisfactory. The second problem is that complexity theory, which is (among other things) concerned with proving lower bounds on the difficulty of computational problems, is unable to provide any reasonable lower bound proofs for any reasonable problem and model of computation, let alone for the problem of breaking a cryptographic system.

The second type of assumption, namely that an adversary has infinite computing power, implies no restriction whatsoever and therefore anticipates all arguments about models of computation and realistic estimates of an opponent's computing power. However, if one considers the theoretical possibility of testing all possible keys of a system at once, it appears impossible to prove a system secure under such an assumption. Here is where information theory comes into play. Shannon defined a cipher system to be *perfect* if the ciphertext provides no information about the plaintext or, equivalently, if plaintext and ciphertext are statistically independent. In other words, when a perfect cipher is used to encipher a message, an adversary can do no better than guess the message without even looking at the ciphertext. Shannon gave as a simple example of a perfect cipher the so-called one-time pad previously proposed by Vernam [32] without proof of security: the binary plaintext is concealed by adding modulo 2 (EXOR) a random binary secret key of the same length. Of course, this system is completely impractical for most applications where only a short secret key is available. Unfortunately, Shannon proved the pessimistic result that perfect secrecy can be achieved only when the secret key is at least as long as the plaintext message (cf. Section 3.1).

This pessimistic result led most researchers to believe that perfect secrecy is bound to be impractical. Therefore, information theory has until recently been believed to provide only pessimistic results in cryptography, that is lower bounds on the necessary key size in order to achieve a certain level of unconditional security. It was only recently demonstrated that information theory can also provide optimistic results, showing that perfect secrecy can indeed be achieved in a realistic scenario (cf. Section 4).

In summary, the role of information theory in cryptography can be characterized as that of deriving results on the provable security of a system, even in presence of adversaries with infinite computing power. In view of the fact that no proof of the computational security of a cipher (which is not also information-theoretically-secure) is in sight at the horizon of current research, it appears to be important to investigate the applicability of information theory in cryptography. Note that many systems have been claimed to be computationally-secure in the literature, but all these proofs rely on an unproven intractability assumption. Although many of these proofs are important results from a theoretical point of view, one could nevertheless argue that in most cases the (intractability) assumption is very close to the theorem to be proven.

This paper is organized as follows. Section 2 summarizes the basic concepts of information theory. In Section 3 we review the most important lower-bound results on the size of secret keys for perfect secrecy, unconditionally-secure authentication and secret sharing. Our treatment of authentication is novel and considerably simpler than previous approaches. Section 4 presents some recent optimistic results on secret-key agreement leading to potentially practical systems with perfect secrecy.

2. Information-theoretic preliminaries

A discrete random variable X taking on values from a finite or countably infinite set $\mathcal{X}$ is characterized completely by its probability distribution P_X, a function $\mathcal{X} \to \mathbf{R}$ assigning to every possible value $x \in \mathcal{X}$ the probability $P_X(x)$ that X takes on the value x, and satisfying $\sum_{x \in \mathcal{X}} P_X(x) = 1$. A random vector $[X_1, \ldots, X_n]$ consisting of several random variables can be considered as a single random variable taking on values in the Cartesian product of the individual sets. The random variables $X_1, \ldots, X_n$ are statistically independent if the probability distribution factors, i.e., if $P_{X_1 \ldots X_n}(x_1, \ldots, x_n) = P_{X_1}(x_1) \cdots P_{X_n}(x_n)$ for all $x_1, \ldots, x_n$.

The *entropy* (or *uncertainty*) of a random variable X is defined by

$$H(X) \;\stackrel{\triangle}{=}\; -\sum_{x \in \mathcal{X}: P_X(x) \neq 0} P_X(x) \log_2 P_X(x)$$

and for a finite set satisfies

$$0 \;\leq\; H(X) \;\leq\; \log_2 |\mathcal{X}|$$

(where $|\mathcal{S}|$ denotes the cardinality of a finite set $\mathcal{S}$) with equality on the left if and only if $P_X(x) = 1$ for some $x \in \mathcal{X}$ and with equality on the right if and only if $P_X(x) = 1/|\mathcal{X}|$

for all $x \in \mathcal{X}$, i.e., if and only if X takes on all possible values equally likely. $H(X)$ measures the uncertainty of an observer (of the random experiment generating X) about the outcome of X and is a real number that depends only on the set of non-zero values of $P_X(.)$

The entropy of a binary random variable taking on the two values with probabilities p and $1-p$ is

$$h(p) \triangleq -p\log_2 p - (1-p)\log_2(1-p)$$

with the convention that $h(0) = h(1) = 0$. $h(p)$ is a strictly convex-$\cap$ function that takes on its maximum for $p = 1/2$ where $h(1/2) = 1$. The joint entropy $H(X_1, \ldots, X_n)$ of the random variables $X_1, \ldots, X_n$ is defined by the obvious generalization of (), i.e.,

$$H(X_1, \ldots, X_n) = \sum P_{X_1,\ldots,X_n}(x_1, \ldots, x_n)\log_2 P_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$$

where here and in the sequel the summation is understood to be only over those values with nonzero probability. (Alternatively, define $0\log_2 0 \triangleq \lim_{\xi\to 0}\xi\log_2\xi = 0$.)

The conditional entropy of a random variable X when given the random variable Y is defined as

$$H(X|Y) \triangleq -\sum_{(x,y)} P_{XY}(x,y)\log P_{X|Y}(x,y),$$

where $P_{X|Y}(x,y)$ is the conditional probability that X takes on the value x given that $Y = y$. $H(X|Y)$ can equivalently be defined as the expected value (over choices of y) of $H(X|Y=y)$, the entropy of the conditional probability distribution $P_{X|Y}(x,y)$ considered as a function of x only. One can show that

$$0 \leq H(X|Y) \leq H(X)$$

with equality on the left if and only if Y uniquely determines X and with equality on the right if and only if X and Y are statistically independent. An important rule for transforming uncertainties is the so-called chain rule:

$$H(X_1\cdots X_n) = H(X_1) + H(X_2|X_1) + \cdots + H(X_n|X_1\cdots X_{n-1})$$

which for conditional entropies has the form

$$H(X_1\cdots X_n|Y) = H(X_1|Y) + H(X_2|X_1Y) + \cdots + H(X_n|X_1\cdots X_{n-1}Y).$$

The *mutual information* $I(X;Y)$ between two random variables X and Y is defined as

$$I(X;Y) \triangleq H(X) - H(X|Y) = H(Y) - H(Y|X) = I(Y;X)$$

and measures the amount by which the uncertainty about X is reduced by giving Y (and vice versa). Similarly, the mutual conditional information between X and Y, given Z, is defined as

$$I(X;Y|Z) \triangleq H(X|Z) - H(X|YZ)$$

and is also symmetric, i.e., $I(X;Y|Z) = I(Y;X|Z)$, as is easily seen by expanding $H(XY|Z)$ in two different ways.

The interested reader is referred to [4] for a detailed introduction to information theory.

3. Pessimistic results: lower bounds on key size

Information theory has been used in cryptography primarily to derive pessimistic results, i.e., lower bounds on the size of the secret key necessary to achieve a certain level of security. In this section we review the three most important areas for which such bounds have been derived: secrecy, authentication and secret sharing.

Consider the model of a symmetric cryptosystem shown in Figure 1. This is a generalization of Shannon's model: it contains a secret randomizer S known only to the sender of a message X as well as a public randomizer R assumed to be available to everybody, including the eavesdropper.

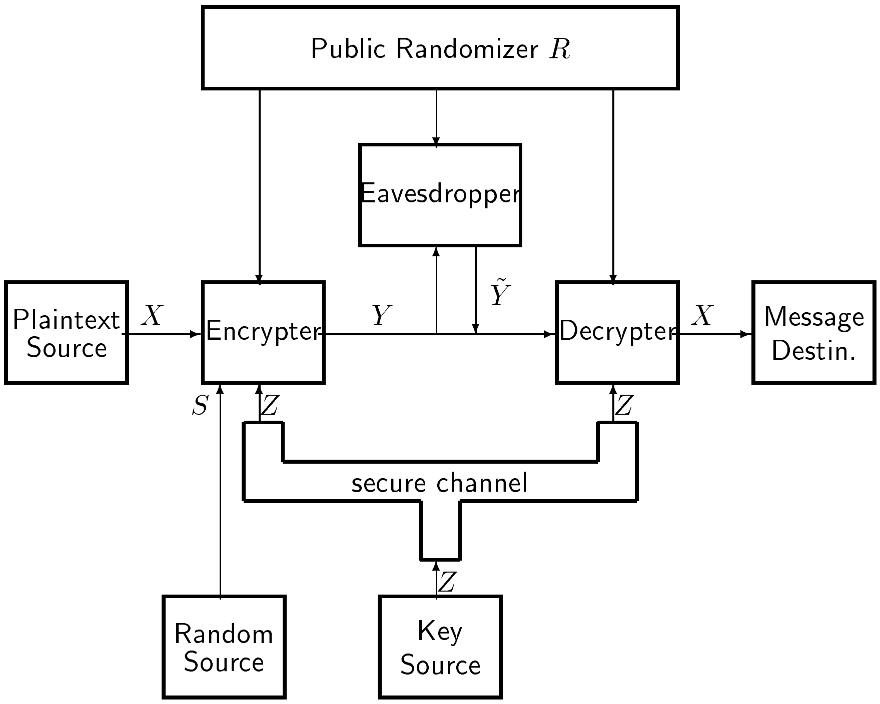

Figure 1: Model of a symmetric cipher with two types of randomizers.

There are two dual and complementary security goals in communication: Confidentiality (or secrecy) and authenticity. Confidentiality means that an eavesdropper cannot obtain any useful information about the plaintext, and authenticity means that an active eavesdropper cannot successfully insert a fraudulent message $\tilde{Y}$ that will be accepted by the receiver.

3.1. Secrecy

A cipher as shown in Figure 1 is defined to be *perfect* [29] if and only if the plaintext X and the ciphertext Y together with the public randomizer are statistically independent, i.e., if and only if $I(X;YR) = 0$. The following theorem is a generalization of Shannon's theorem [29].

Theorem 1. *A cipher can be perfect only if*

$$H(Z) \;\geq\; H(X).$$

Proof. Every cipher must be uniquely decipherable. Therefore

$$H(X|YZR) \;=\; 0.$$

The definition of perfect secrecy, $I(X;YR) = 0$, can be stated as

$$H(X|YR) \;=\; H(X).$$

Using the basic expansion rule for uncertainties, and the fact that the removal of given knowledge can only increase uncertainty, we obtain

$$\begin{aligned} H(X) \;=\; H(X|YR) \;&\leq\; H(XZ|YR) \\ &=\; H(Z|YR) + H(X|YZR) \\ &=\; H(Z|YR) \\ &\leq\; H(Z). \qquad \square \end{aligned}$$

Note that the condition of Theorem 1 does not guarantee a cipher to be perfect.

It will be demonstrated in Section 4 that virtually-perfect secrecy is possible even when $H(Z) \ll H(X)$, provided that the model of Figure 1 is modified slightly. For instance, if the public randomizer is very large then it is infeasible (although theoretically possible) for an eavesdropper to read the entire string R, and therefore he or she is left with incomplete information, contradicting the Shannon assumption implied by the model. Alternatively, one could assume that accessing the randomizer is not free of errors: for instance R could be broadcast by a satellite with a very low signal power where sender, receiver and eavesdropper can receive the bits only with certain nonzero bit error probabilities. Surprisingly, perfect secrecy can be achieved in such a realistic scenario even without a secret key Z and even when the eavesdropper receives the random bits of R by orders of magnitude more reliably than the legitimate sender and receiver [21].

3.2. Authentication

Consider an active eavesdropper who wants to insert a fraudulent ciphertext $\tilde{Y}$, hoping that it will be accepted by the receiver. There are essentially two different types of attack that the eavesdropper can use. *Impersonation* after seeing $i-1$ ciphertext messages $Y_1, \ldots, Y_{i-1}$ (and letting them pass by) means that the eavesdropper chooses a new message $\tilde{Y}_i$ which she wishes to be accepted by the receiver as the ith message Y_i. Let $P_I(i)$ denote the probability of success for this type of attack. *Substitution* after seeing i ciphertext messages $Y_1, \ldots, Y_i$ (and letting $Y_1, \ldots, Y_{i-1}$ pass by) means that the eavesdropper tries to substitute Y_i with a different message $\tilde{Y}_i \neq Y_i$. Let $P_S(i)$ denote the probability of success for this type of attack.

Many lower bounds on the probability of successful attacks have been derived in the literature (see [30, 18, 34, 31, 27] and the extensive list of references cited therein). The various papers differ somewhat in the model that is adopted: for instance some papers assume that all messages sent by the sender must be different while others point out that this restriction is unnecessary. Furthermore, some papers assume that a fixed encoding rule is used whereas other papers stress a model in which the encoding rule can change with time, depending on a global secret key. Most unsatisfactorily, most papers on authentication are quite lengthy and complicated, and many of them prove essentially the same results.

The purpose of this section is to point out that the right way of looking at authentication is in the framework of classical hypothesis testing (e.g., see [4]). This allows to derive most if not all previously-known lower bounds in a unified and strongly simplified way.

Hypothesis testing is the task of deciding which of two hypotheses, H_0 or H_1, is true, when one is given the value of a random variable U (e.g., the outcome of a measurement). The behavior of U is described by two probability distributions: If H_0 or H_1 is true, then U is distributed according to the distribution $P_{U|0}$ or $P_{U|1}$, respectively. For ease of notation we will write $P_{U|0} = P_U$ and $P_{U|1} = Q_U$. A decision rule assigns one of the two hypotheses to each possible u that U can assume. There are two types of errors possible in making a decision. Accepting hypothesis H_1 when H_0 actually is true is called a type I error, and the probability of this event is denoted by α. Accepting hypothesis H_0 when H_1 actually is true is called a type II error, and the probability of this event is denoted by β. The optimal decision rule is given by the famous Neyman-Pearson theorem which states that, for a given maximal tolerable probability β of type II error, α can be minimized by assuming hypothesis H_0 if and only if

$$\log \frac{P_U(u)}{Q_U(u)} \geq T \tag{1}$$

for some threshold T. (All logarithms in this paper are to the base 2). Note that only the existence of T, but not its value is specified by this theorem. The term on the left of (1) is called the log-likelihood ratio. We refer to [4] for an excellent treatment of hypothesis testing.

Let P_U and Q_U be arbitrary probability distributions over the same finite or countably infinite set $\mathcal{U}$ of values. The expected value of the log-likelihood ratio with respect to P_U is called the discrimination and is defined by

$$L(P_U; Q_U) \triangleq \sum_{u \in \mathcal{U}} P_U(u) \log \frac{P_U(u)}{Q_U(u)}.$$

The discrimination is non-negative and is equal to zero if and only if the two distributions are identical.

A well-known result in hypothesis testing (cf. [4], Theorem 4.4.1) provides a relation between the error probabilities α and β and the discrimination $L(P_U; Q_U)$:

$$L(P_U; Q_U) \geq \beta \log \frac{\beta}{1-\alpha} + (1-\beta) \log \frac{1-\beta}{\alpha} \tag{2}$$

In particular, for $\beta = 0$ we have

$$L(P_U; Q_U) \geq -\log \alpha \tag{3}$$

or, equivalently,

$$\alpha \geq 2^{-L(P_U; Q_U)} \tag{4}$$

Let X and Y be random variables with joint distribution P_{XY} and marginal distributions P_X and P_Y, respectively. It follows immediately from the definition of the mutual information between X and Y,

$$I(X;Y) = H(X) - H(X|Y) = H(X) + H(Y) - H(XY),$$

that

$$L(P_{XY}; P_X P_Y) = I(X;Y). \tag{5}$$

Note that P_{XY} and $P_X P_Y$ are both probability distributions over the set $\mathcal{X} \times \mathcal{Y}$ when X and Y take on values in $\mathcal{X}$ and $\mathcal{Y}$, respectively. We have $L(P_{XY}; P_X P_Y) = 0$ if and only if the two distributions are identical, which is equivalent to saying that X and Y are statistically independent.

Let us now return to the analysis of message authentication. The problem of deciding whether a received message $\overline{Y}$ is authentic or not can be viewed as a hypothesis testing problem. H_0 corresponds to the hypotheses that the message is authentic, and H_1 corresponds to the hypotheses that the message has been generated by a fraudulent opponent.

We first consider an impersonation attack for the first message in which the receiver is given $\overline{Y}$ and the key Z. Under hypothesis H_0, the pair $[\overline{Y}, Z]$ is generated according to $P_{Y_1 Z}$ whereas under hypothesis H_1, $[\overline{Y}, Z]$ is generated according to the distribution $Q_{\overline{Y}} P_Z$ for some distribution $Q_{\overline{Y}}$ because the opponents must choose $\overline{Y}$ independent of the secret key Z. (Note that in a deterministic strategy, $Q_{\overline{Y}}$ is equal to 1 for one particular value and zero otherwise.) One particular strategy for an opponent is to choose $Q_{\overline{Y}} = P_{Y_1}$. In this case it follows from (2) and (5) that

$$I(Y_1; K) = L(P_{Y_1 Z}; P_{Y_1} P_Z) \geq \beta \log \frac{\beta}{1-\alpha} + (1-\beta) \log \frac{1-\beta}{\alpha}.$$

If the receiver is required to always accept legitimate messages, i.e. if $\beta = 0$, then we have

$$P_I(1) \geq 2^{-I(Y_1, Z)} \tag{6}$$

which is Simmons' bound [30].

We now consider a substitution attack for the first message in which an opponent, who is given a first valid message Y_1, tries to substitute it with a different message $\overline{Y} \neq Y_1$. The following bound due to Simmons [30] can also be derived using certain hypothesis testing arguments:

$$P_S(1) \geq 2^{-H(Z|Y_1)}. \tag{7}$$

Combining (6) and (7) one obtains

$$P_I(1) \cdot P_S(1) \geq H(K).$$

Again using hypothesis testing arguments (see [22]) and Jensen's inequality one obtains the following results on multiple authentication which was first derived by Walker [34].

Theorem 2. *The probabilities of impersonation and substitution in authentication for multiple messages using the single key Z satisfy the following inequalities for all n:*

$$P_I(n) \geq 2^{-I(Y_n; Z|Y_1, \ldots, Y_{n-1})},$$

$$P_S(n) \geq 2^{-H(Z|Y_1, \ldots, Y_n)}$$

and

$$P_S(n) \prod_{i=1}^{n} P_I(i) \geq 2^{-H(Z)}.$$

3.3. Secret sharing

Information theory has only recently been applied to derive lower bounds on the size of shares in perfect secret sharing schemes [8]. A perfect secret sharing scheme allows for a secret S to be distributed among n participants in such a way that only qualified subsets of participants can recover the secret while any non-qualified subset has no information about S.

Given a set $\mathcal{P}$ of participants, an access structure $\mathcal{A}$ on $\mathcal{P}$ is a family of subsets of $\mathcal{P}$: $\mathcal{A} \subseteq 2^{\mathcal{P}}$. It only makes sense to consider monotone access structures, i.e., we assume that $T \in \mathcal{A}$ and $T \subset T'$ together imply that $T' \in \mathcal{A}$. For a given access structure we actually consider its monotone closure. We denote the share given to participant $P \in \mathcal{P}$ also by P, hoping that this will not cause any confusion. Similarly we denote the set of shares given to a set $T \subseteq \mathcal{P}$ also by T. The conditions for a perfect secret sharing scheme can be stated as follows:

$$T \in \mathcal{A} \implies H(S|T) = 0$$

and

$$T \notin \mathcal{A} \implies H(S|T) = H(S).$$

Consider now the set $\mathcal{P} = \{A, B, C, D\}$ of four participants and the access structure consisting of the monotone closure of $\mathcal{A} = \{AB, BC, CD\}$, which is $\{AB, BC, CD, ABC, ABD, ACD, BCD, ABCD\}$ Hence we have

$$H(S|AB) = H(S|BC) = H(S|CD) = 0,$$

but

$$H(S|AC) = H(S|AD) = H(S|BD) = H(S).$$

Using these facts one can derive the lower bound $H(BC) \geq 3H(S)$ [8] which implies that at least one of the shares of B and C must be at least 1.5 times the length of S, either $H(B) \geq 1.5H(S)$ or $H(C) \geq 1.5H(S)$ or both. Instead of repeating the argument of [8] we demonstrate a general proof technique.

Consider the set $\mathcal{P}' = \mathcal{P} \cup S$ and the set of all subsets of $\mathcal{P}'$ (which form a partially ordered set). We can define a labeled directed graph associated with $\mathcal{P}'$ naturally as follows: the vertices are the subsets of $\mathcal{P}'$ (including the empty set $\{\}$) and two vertices T and T' are connected by a directed edge $T \to T'$ if and only if $T' \neq T$ and $T = T' \cup P$ for some $P \in \mathcal{P}'$. Each edge $T \cup P \to T$ can be thought of being labeled with $H(T \cup P|T) = H(P|T)$. For a given access structure $\mathcal{A}$, such a labeling implies that an edge $T \cup S \to T$ is labeled $H(S)$ if and only if $T \notin \mathcal{A}$ and labeled 0 if and only if $T \in \mathcal{A}$.

A proof for a lower bound on the size of the shares in a secret sharing scheme usually can be stated in the form of a lower bound on $H(T)$ for some $T \subset \mathcal{P}$. Such a lower bound can be derived by considering an appropriate path from T to $\{\}$ in the described graph, where all entropies assigned to the edges of the path are to be added or subtracted when the edge is traversed in the labeled direction or in the reverse direction, respectively. The art of deriving such a proof is that of finding a path which traverses as many $H(S)$-labeled edges as possible in positive direction and as many 0-labeled edges as possible in the negative direction.

Consider the example discussed above. Edges labeled $H(S)$ are $AS \to A$, $BS \to B$, $CS \to C$, $DS \to D$, $ADS \to AD$, $ACS \to AC$ and $BDS \to BD$, and the edges labeled 0 are $ABS \to AB$, $BCS \to BC$, $CDS \to CD$, $ABCS \to ABC$, $ABDS \to ABD$, $ACDS \to ACD$, $BCDS \to BCD$ and $ABCDS \to ABCD$. Consider now the path

$$BC \to BCS \to ABCS \to ACS \to AC \to ACD \to ACDS \to ADS \to AD \to A \to \{\}$$

which can be interpreted as

$$\begin{aligned} H(BC) \;&=\; -H(S|BC) - H(A|BCS) + H(B|ACS) + H(S|AC) - H(D|AC) \\ &\quad -H(S|ACD) + H(C|ADS) + H(S|AD) + H(D|A) + H(A). \end{aligned} \tag{8}$$

A closed loop in the graph corresponds to a total value of 0. Therefore the loop $B \to AB \to ABS \to BS \to B$ corresponds to the equation

$$-H(A|B) - H(S|AB) + H(A|BS) + H(S|B) = 0 \tag{9}$$

Adding the left side of (9) to the right side of (8), and using

$$\begin{aligned} & H(S|BC) = H(S|AB) = H(S|ACD = 0, \\ & H(S|AC) = H(S|AD) = H(S|B) = H(S), \\ & H(A) - H(A|B) \geq 0, \\ & H(D|A) - H(D|AC) \geq 0, \\ & H(A|BS) - H(A|BCS) \geq 0, \\ & H(B|ACS) \geq 0, \\ \text{and} \qquad & H(C|ADS) \geq 0 \end{aligned}$$

we obtain the desired result $H(BC) \geq 3H(S)$.

This proof technique can be automated and implemented as a graph-theoretic algorithm. It has recently been applied to find the first access structure for which in each secret sharing scheme at least one share is at least twice as long as the secret [6]. Subsequently, Csirmaz [10] improved these results considerably.

4. Optimistic results on perfect secrecy

In this section we focus our attention on the fun part of information theory in cryptography by demonstrating that perfect secrecy defined by Shannon can indeed be achieved in a realistic scenario. An intuitive generalization of Theorem 1 is that in any model in which the adversary can observe the entire communication between two communicating parties, perfect secrecy can only be achieved if the entropy of the secret key is at least equal to the total entropy of the exchanged messages. However, the proof of this intuitive result is not completely trivial [20]. Because of this lower-bound result, perfect secrecy can only be achieved in a model in which an adversary cannot obtain precisely the same information as both legitimate parties. In the following, we will refer to the legitimate parties as Alice and Bob and to the adversary as Eve.

4.1. Wire-tap and broadcast channels

The first such model was the wire-tap channel proposed by Wyner in 1975 [35]. Wyner considered an eavesdropper Eve tapping a telephone over an imperfect wire-tapping channel. Assume for instance that Eve can see all bits transmitted from Alice to Bob, but only with bit error probability ϵ. In this case, Alice could encode each information bit to be sent by generating and sending $N-1$ random bits and sending as the N-th bit of a block the XOR of the $N-1$ random bits and the information bit. Bob could easily recover the information bit because he receives Alices messages without errors, but Eve's bit error probability when guessing the information bit can easily be shown to be $(1-(1-2\epsilon)^N)/2$ which approaches $1/2$ exponentially fast in N.

Wyner's model and results were generalized by Csiszár and Körner [11] who considered a discrete memoryless broadcast channel for which the wire-tapper Eve's received message is not necessarily a degraded version of the legitimate receiver's message. The common input to the main channel and Eve's channel is the random variable X chosen by Alice according to some probability distribution P_X, and the random variables received by the legitimate receiver Bob and by the adversary Eve are Y and Z, respectively. X, Y and Z take on values in some finite or countably infinite alphabets $\mathcal{X}, \mathcal{Y}$ and $\mathcal{Z}$, respectively. The channel behavior is completely specified by the conditional probability distribution $P_{YZ|X}$. Note that in Wyner's original setting [35], X, Y and Z form a Markov chain, i.e., $P_{Z|XY} = P_{Z|Y}$, which implies $I(X;Z|Y) = 0$.

The secrecy capacity $C_s(P_{YZ|X})$ of the described broadcast channel with transition probability distribution $P_{YZ|X}$ was defined in [11] as the maximum rate at which Alice can reliably send information to Bob such that the rate at which Eve obtains this information is arbitrarily small. In other words, the secrecy capacity is the maximal number of bits per use of the channel that Alice can send to Bob in secrecy. Csiszár and Körner [11] proved that

$$C_s(P_{YZ|X}) \geq \max_{P_X}[H(X|Z) - H(X|Y)] \tag{10}$$

where the inequality is satisfied with equality except in very exceptional cases that are not of interest. If equality holds, the secrecy capacity is zero unless $I(X;Y) > I(X;Z)$ for some P_X.

In order to demonstrate that feedback from Bob to Alice over an insecure public channel can increase the secrecy capacity of a broadcast channel, we consider a broadcast channel for which the main channel and Eve's channel are independent binary symmetric channels with bit error probabilities ϵ and δ, respectively, i.e., X, Y and Z are binary random variables and $P_{YZ|X} = P_{Y|X} \cdot P_{Z|X}$ where $P_{Y|X}(y|x) = 1-\epsilon$ if $x = y$, $P_{Y|X}(y|x) = \epsilon$ if $x \neq y$, $P_{Z|X}(z|x) = 1-\delta$ if $x = z$, and $P_{Z|X}(z|x) = \delta$ if

$x \neq z$. Without loss of generality we may assume that $\epsilon \leq 1/2$ and $\delta \leq 1/2$. For ease of notation, we will refer to the described probability distribution $P_{YZ|X}$ as $D(\epsilon, \delta)$. It follows from (10) (see [20]) that

$$C_s(D(\epsilon, \delta)) = \begin{cases} h(\delta) - h(\epsilon) & \textit{if } \delta > \epsilon, \\ 0 & \textit{otherwise}. \end{cases}$$

It follows that secret messages can be sent only if $\delta > \epsilon$. However, if public feedback is allowed, secret messages can be exchanged even when $\delta < \epsilon$. To this end, Bob can conceptually convert the channel scenario into one in which he is the sender and Alice and Eve are the receivers [20]. To achieve this, Alice sends a random bit over the noisy broadcast channel to Bob (and Eve) and Bob XORs the bit he wishes to send to Alice with the received bit and sends the result over the public channel. Of course, Alice can XOR this bit with the random bit she sent and thus "receive" Bob's bit with error probability ϵ. However, one can prove that Eve "sees" Bob's bit as if it had been sent through a cascade of the two noisy channels and hence her bit error probability is $\epsilon + \delta - 2\epsilon\delta$. Thus the secrecy capacity *with feedback* is $h(\epsilon + \delta - 2\epsilon\delta) - h(\epsilon)$ which is positive unless $\epsilon = 1/2$ or $\delta = 0$.

The described broadcast channel scenario is a special case of a much more general secret-key agreement scenario to be described in the following section.

4.2. Unconditionally-secure secret-key agreement

Unconditionally-secure secret-key agreement [3], [20] takes place in a scenario where Alice and Bob are connected by an insecure channel to which a passive eavesdropper Eve has perfect access, and where Alice, Bob and Eve know the correlated random variables X, Y and Z, respectively, which are distributed according to some joint probability distribution P_{XYZ}.

Alice and Bob share no secret key initially (other than possibly a short key required for guaranteeing authenticity and integrity of messages sent over the public channel), but are assumed to know P_{XYZ} or at least an upper bound on the quality of Eve's channel. In particular, the protocol and the codes used by Alice and Bob are known to Eve. Every message communicated between Alice and Bob can be intercepted by Eve, but it is assumed that Eve cannot insert fraudulent messages nor modify messages on this public channel without being detected.

Possible attacks by Eve other than passive wire-tapping can be detected when an unconditionally secure authentication scheme with a short initially shared secret key is used. If only a computationally secure authentication scheme were used, the unconditional security would only be retained against passive, but not against active wire-tapping.

A broadcast channel as described in the previous section is one of several possible realizations for the distribution of random variables X, Y and Z. An alternative for Alice and Bob to acquire random variables X and Y is to receive the signal of a satellite broadcasting random bits at a very low signal power (so that even if Eve uses a much better and larger receiving antenna she cannot avoid at least a small bit error probability). Quantum cryptography (see Section 4.4) is another example of such a scenario where, according to the laws of quantum physics, Eve cannot obtain complete information.

A key agreement protocol for such a scenario generally consists of at least two phases. In the first phase, often referred to as *information reconciliation* [1, 5], Alice and Bob exchange redundant information and apply error-correction techniques in order to generate a shared string W which both of them know with very high probability while Eve has only incomplete information about W. In the second phase, called *privacy amplification*, Alice and Bob distill from W a shorter string S about which Eve has only a negligible amount of information. Privacy amplification will be discussed in the following section.

It was proved in [20] that the size of the secret key S that can be generated by any protocol, not necessarily one of the two-phase type described above, is upper bounded by

$$H(S) \ \leq \ \max(I(X;Y|Z), I(X;Y)) + I(S;CZ),$$

where C summarizes the total communication between Alice and Bob over the public channel. In other words, if $I(S;CZ)$ must be negligible (which is the goal of such a key agreement protocol), then Alice and Bob cannot generate a key that is longer than the mutual information between X and Y. Moreover, because if Eve revealed her random variable Z for free, this could only help Alice and Bob to generate a secret key. Therefore, the remaining mutual information between X and Y when given Z, $I(X;Y|Z)$, is also an upper bound on $H(S)$. Note that both $I(X;Y|Z) < I(X;Y)$ or $I(X;Y|Z) > I(X;Y)$ is possible.

In order to be able to prove lower bounds on the achievable size of a key shared by Alice and Bob in secrecy we need to make more specific assumptions about the distribution P_{XYZ}. One natural assumption is that the random experiment generating XYZ is repeated many times independently: Alice, Bob and Eve receive $X^N = [X_1, \ldots, X_N]$, $Y^N = [Y_1, \ldots, Y_N]$ and $Z^N = [Z_1, \ldots, Z_N]$, respectively, where

$$P_{X^N Y^N Z^N} \ = \ \prod_{i=1}^{N} P_{X_i Y_i Z_i}$$

and where $P_{X_i Y_i Z_i} = P_{XYZ}$ for $1 \leq i \leq N$.

This model of independent repetitions of a random experiment is well motivated by the satellite scenario described above. Moreover, it is consistent with standard information-theoretic models such as discrete memoryless sources and channels. The natural quantity of most interest is defined as follows.

Definition 1 [21]. The *secret key rate of X and Y with respect to Z*, denoted $\bar{S}(X;Y||Z)$, is the maximum rate at which Alice and Bob can agree on a secret key S while keeping the amount of information about S available to Eve arbitrarily small.

It should be pointed out that the original definition of secret-key rate of [20], which was motivated by the definition of secrecy capacity [35, 11], was considerably weaker in that only the rate at which Eve obtains information about S, rather than the total amount, had to be arbitrarily small.

The following lower bound on the secret-key rate was proved in [21].

Theorem 3. $\bar{S}(X;Y||Z) \;\geq\; \max[I(Y;X) - I(Z;X),\; I(X;Y) - I(Z;Y)]$.

This theorem states that if either $I(X;Y) > I(X;Z)$ or $I(Y;X) > I(Y;Z)$, i.e., if Bob has more information than Eve about X (or Alice has more information than Eve about Y), then secret-key agreement is possible. Although the proof of the theorem is quite involved and makes use of privacy amplification techniques discussed in the following section, the result is quite intuitive. For instance if $I(X;Y) > I(X;Z)$, it appears reasonable that Alice can send redundant information to Bob at a rate of $H(X|Y)$ which allows Bob to determine X^N. At the same time, Eve is left with uncertainty about X^N at a rate $H(X|Z)-H(X|Y)$ which is equal to $I(X;Y)-I(X;Z)$.

It is quite surprising that even when neither $I(X;Y) > I(X;Z)$ nor $I(Y;X) > I(Y;Z)$ is satisfied, secret-key agreement is nevertheless possible, i.e., the lower bound of Theorem 3 is not tight. This was first illustrated in [20] for the described satellite scenario in which Eve receives the bits (much) more reliably than both Alice and Bob.

4.3. Privacy amplification

A basic tool for applying information theory to cryptography is privacy amplification originally introduced by Bennett, Brassard and Robert [3]. However, this technique was only recently shown to be applicable to a wide range of scenarios by Bennett, Brassard, Crépeau and Maurer [2] who provided a generalized analysis of privacy amplification. These results are summarized briefly in this section. An important technique used is universal hashing.

Definition 2. A class G of functions $\mathcal{A} \longrightarrow \mathcal{B}$ is *universal* if, for any distinct x_1 and x_2 in $\mathcal{A}$, the probability that $g(x_1) = g(x_2)$ is at most $1/|\mathcal{B}|$ when g is chosen at random from G according to the uniform distribution.

Consider functions from $\{0,1\}^n$ to $\{0,1\}^r$. For $r = n$, the class consisting only of the identity function is trivially a universal class. Consider the more interesting case $r < n$. The class of all functions from $\{0,1\}^n$ to $\{0,1\}^r$ is obviously universal, but it is not useful because there are too many functions in that class (it takes $r2^n$ bits to specify a function). A more useful universal class is that of all *linear* functions from $\{0,1\}^n$ to $\{0,1\}^r$ [9]. These functions can be described by $r \times n$ matrices M over $GF(2)$, i.e., by rn bits. Other universal classes, which are more economical in terms of the number of bits needed to specify them, are discussed in [9, 33]. A very small such class requiring only n bits to specify a function is given in the following example.

Example. Let a be an element of $GF(2^n)$ and also interpret x as an element of $GF(2^n)$. Consider the function $\{0,1\}^n \to \{0,1\}^r$ assigning to an argument x the first r bits of the element ax of $GF(2^n)$. The class of such functions for $a \in GF(2^n)$ with $a \neq 0$ is a universal class of functions for $1 \leq r \leq n$.

We further need the following definition of an alternative measure of information.

Definition 3. Let X be a random variable with alphabet $\mathcal{X}$ and distribution P_X. The *collision probability* $P_c(X)$ of X is defined as the probability that X takes on the same value twice in two independent experiments, i.e.,

$$P_c(X) = \sum_{x \in \mathcal{X}} P_X(x)^2.$$

The *collision entropy* of X, also known as the Renyi entropy of order two [25], is defined as the negative logarithm of its collision probability:

$$H_c(X) = -\log_2 P_c(X).$$

For an event $\mathcal{E}$, the *collision entropy of X conditioned on $\mathcal{E}$*, $H_c(X|\mathcal{E})$, is defined naturally as the collision entropy of the conditional distribution $P_{X|\mathcal{E}}$.

In order to contrast collision entropy with the normal entropy measure defined by Shannon, we will in the sequel refer to the latter as "Shannon entropy". Note that collision entropy (like Shannon entropy) is always positive. $H_c(X)$ can equivalently be expressed as $H_c(X) = -\log E[P_X(X)]$, where $E[\cdot]$ denotes the expected value. Shannon entropy $H(X)$ can be expressed similarly as $H(X) = -E[\log P_X(X)]$. It follows from Jensen's inequality (see [4], p. 428) that collision entropy is upper bounded by the Shannon entropy:

$$H_c(X) \leq H(X),$$

with equality if and only if P_X is the uniform distribution over $\mathcal{X}$ or a subset of $\mathcal{X}$.

We now can state the main theorem on privacy amplification.

Theorem 4. *Let P_{VW} be an arbitrary probability distribution where W is known to both Alice and Bob and V summarizes the complete information known to Eve. If Eve's collision entropy $H_c(W|V=v)$ about W is known to be at least t and Alice and Bob choose $S = G(W)$ as their secret key, where G is chosen at random from a universal class of hash functions from $\mathcal{W}$ to $\{0,1\}^r$, then*

$$H(S|G, V=v) \;\geq\; r - \frac{2^{r-t}}{\ln 2}.$$

This theorem states that if for some reason one knows that the probability distribution about W seen by Eve, $P_{W|V=v}$, satisfies a certain global constraint, namely

$H_c(W|V = v) \geq t$, then Eve's information about the secret key S, $H(S) - H(S|G, V = v)$, is provably exponentially small in the excess compression $t - r$. Note that a statement about the particular value $V = v$ known to Eve is stronger that a statement about an average over all values of V. Theorem 4 can be applied in many scenarios like quantum cryptography, the satellite scenario mentioned earlier as well as if Eve can obtain arbitrary t bits of deterministic information about W (see corollary below) because in these scenarios one knows a lower bound on Eve's collision entropy.

Theorem 4 also implies the following result on privacy amplification against deterministic information first stated in [3]. (For a proof see [2].)

Corollary 5. *Let W be a random n-bit string with uniform distribution over $\{0,1\}^n$, let $V = e(W)$ for an arbitrary eavesdropping function $e : \{0,1\}^n \to \{0,1\}^t$ for some $t < n$, let $s < n - t$ be a positive safety parameter, and let $r = n - t - s$. If Alice and Bob choose $S = G(W)$ as their secret key, where G is chosen at random from a universal class of hash functions from $\{0,1\}^n$ to $\{0,1\}^r$, then Eve's total information about the secret key S, given G and V, satisfies*

$$I(S; GV) \;\leq\; 2^{-s}/\ln 2.$$

It should be pointed out that averaging over the values of V is necessary here. Note also that Alice and Bob learn nothing about the particular function e selected by Eve. Equivalently, Eve could be allowed to obtain W and to perform an arbitrary computation on W, as long as she is guaranteed to keep at most t bits of the result of her computation.

4.4. Quantum cryptography

Quantum cryptography (see [1] and references therein) can be viewed as a special case of the model of secret-key agreement discussed above. The laws of quantum physics (provided they are correct) guarantee that an eavesdropper can measure at most one bit of information about the real-valued polarization angle of a photon. This fact can be exploited by sending single photons whose polarization encodes two bits of information, i.e., whose polarization takes on any one of four different values equally likely.

The major difference between quantum cryptography and secret-key agreement as discussed above, which requires an extension of the model, is that an eavesdropper can influence the distribution P_{XYZ} by her measurement. In other words, she can choose from a collection of such distributions. However, quantum physics implies that all of these distributions leave Eve with sufficient collision entropy to be exploited by privacy amplification.

4.5. Randomized Ciphers

A quite different approach to beating Shannon's pessimistic bound (Theorem 1) on the key length was described in [19]. In this model it is assumed that the public randomizer R (cf. Fig. 1) is a very large array of random bits which is publicly available and can be accessed by everybody, but which is infeasible to read entirely. For instance, one could (somewhat unrealistically) think about the surface of the moon which could be scanned to provide random bits. It is assumed that reading these bits is possible without error.

What distinguishes Alice and Bob from Eve, however, is that they know a *short* secret key which specifies which bits need to be accessed and how they must be combined in order to compute a key stream to be used as a one-time pad. Unless Eve reads essentially all the public random bits, which is completely infeasible, she is left with zero information about the key stream with overwhelming probability.

5. Conclusions

This paper has illustrated many relation between information theory and cryptography. While the results of Section 3 were pessimistic in the sense that they only showed what is impossible to achieve, a slight modification of the classical Shannon model of a cipher system allows to derive constructive results on perfect secrecy, some of which were described in Section 4. We hope that this paper stimulates further research in unconditionally-secure secret-key agreement protocols.

Acknowledgements

It is a pleasure to thank Charles Bennett, Gilles Brassard, Christian Cachin and Claude Crépeau for interesting discussions.

References

[1] C. H. Bennett, F. Bessette, G. Brassard, L. Salvail and J. Smolin, Experimental quantum cryptography, *Journal of Cryptology*, Vol. 5, No. 1, 1992, pp. 3–28.

[2] C.H. Bennett, G. Brassard, C. Crépeau and U.M. Maurer, Generalized privacy amplification, submitted to *IEEE Trans. on Information Theory*, 1994.

[3] C.H. Bennett, G. Brassard and J.-M. Robert, Privacy amplification by public discussion, *SIAM Journal on Computing*, Vol. 17, No. 2, 1988, pp. 210–229.

[4] R.E. Blahut, *Principles and Practice of Information Theory*, Reading, MA: Addison-Wesley, 1987.

[5] G. Brassard and L. Salvail, "Reconciliation of a secret key through public discussion," to appear in *Advances in Cryptology – Eurocrypt '93*, Springer Verlag.

[6] C. Cachin, personal communication, 1994.

[7] C. Cachin and U.M. Maurer, Linking information reconciliation and privacy amplification, to appear in *Advances in Cryptology – EUROCRYPT '94*, Lecture Notes in Computer Science, New York, NY: Springer Verlag.

[8] R. M. Capocelli, A. De Santis, L. Gargano and U. Vaccaro, On the size of shares for secret sharing schemes, *Advances in Cryptology – CRYPTO 91*, Lecture Notes in Computer Science, No. 196. New York, NY: Springer Verlag, 1992, pp. 101–113.

[9] J.L. Carter and M. N. Wegman, Universal classes of hash functions, *Journal of Computer and System Sciences*, Vol. 18, 1979, pp. 143–154.

[10] L. Csirmaz, The size of a share must be large, to appear in *Advances in Cryptology – EUROCRYPT '94*, Lecture Notes in Computer Science, New York, NY: Springer Verlag.

[11] I. Csiszár and J. Körner, Broadcast channels with confidential messages, *IEEE Transactions on Information Theory*, Vol. 24, No. 3, 1978, pp. 339–348.

[12] W. Diffie and M.E. Hellman, "New directions in cryptography," *IEEE Transactions on Information Theory*, Vol. 22, No. 6, 1976, pp. 644–654.

[13] M. Gander and U.M. Maurer, On the secret-Key rate of binary random variables, to appear in *Proc. 1994 IEEE Int. Symp. on Information Theory*, Trondheim, Norway, June 27 – July 1, 1994.

[14] I.N. Herstein, *Topics in Algebra*, New York: John Wiley & Sons, 2nd ed., 1975.

[15] R. Johannesson, A. Sgarro, Strengthening Simmons' bound on impersonation, *IEEE Trans. on Information Theory*, Vol. 37, No. 4, July 1991, pp. 1182–1185.

[16] S. Lloyd, A potentially realizable quantum computer, *Science*, Vol. 261, 1993, pp. 1569–1571.

[17] J.L. Massey, A simplified treatment of Wyner's wire-tap channel, *Proc. 21st Annual Allerton Conf. on Comm., Control, and Computing*, Monticello, IL, Oct. 5-7, 1983, pp. 268–276.

[18] J.L. Massey, Contemporary cryptology – an Introduction, in *Contemporary cryptology – the science of information integrity*, G.J. Simmons (Ed.), IEEE Press, 1992.

[19] U.M. Maurer, Conditionally-perfect secrecy and a provably-secure randomized cipher, *Journal of Cryptology*, Vol. 5, No. 1, 1992, pp. 53-66.

[20] U.M. Maurer, Secret key agreement by public discussion from common information, *IEEE Transactions on Information Theory*, Vol. 39, No. 3, 1993, pp. 733–742.

[21] U. M. Maurer, The strong secret key rate of discrete random triples, to appear in *Proc. Symp. on Communications, Coding and Cryptography*, R. Blahut et al. (eds.), Ascona, Switzerland, Feb. 10–13, 1994, Kluwer.

[22] U. M. Maurer, Authentication theory and hypothesis testing, in preparation.

[23] U.M. Maurer and J.L. Massey, Local randomness in pseudo-random sequences, *Journal of Cryptology*, Vol. 4, No. 2, 1991, pp. 135–149.

[24] U.M. Maurer and J.L. Massey, Cascade ciphers: the importance of being first, *Journal of Cryptology*, Vol. 6, No. 1, 1993, pp. 55–61.

[25] A. Renyi, "On Measures of Entropy and Information," *Proc. 4th Berkeley Symp. Math. Statist. Prob.*, Vol. 1, 1961, pp. 547–561.

[26] R.L. Rivest, A. Shamir, and L. Adleman, A method for obtaining digital signatures and public-key cryptosystems, *Communications of the ACM*, Vol. 21, No. 2, 1978, pp. 120–126.

[27] U. Rosenbaum, A lower bound on authentication after having observed a sequence of messages, *J. of Cryptology*, Vol. 6, No. 3, 1993, pp. 135–156.

[28] C.E. Shannon, A mathematical theory of communication, *Bell Syst. Tech. J.*, Vol. 27, No. 3, 1948, pp. 379-423 and 623-656.

[29] C.E. Shannon, Communication theory of secrecy systems, *Bell System Technical Journal*, Vol. 28, Oct. 1949, pp. 656-715.

[30] G.J. Simmons, Authentication theory/coding theory, in *Advances in Cryptology – CRYPTO 84*, G.R. Blakley and D. Chaum (Eds.), Lecture Notes in Computer Science, No. 196. New York, NY: Springer, 1985, pp. 411–431.

[31] B. Smeets, Bounds on the Probability of Deception in Multiple Authentication, to appear in *IEEE Transactions of Information Theory.*

[32] G.S. Vernam, Cipher printing telegraph systems for secret wire and radio telegraphic communications, *J. Amer. Inst. Elec. Eng.*, Vol. 55, pp. 109–115, 1926.

[33] M.N. Wegman and J.L. Carter, "New hash functions and their use in authentication and set equality," *Journal of Computer and System Sciences*, Vol. 22, 1981, pp. 265–279.

[34] M. Walker, Information-theoretic bounds for authentication schemes, *J. of Cryptology*, Vol 2, No. 3, 1990, pp. 131–143.

[35] A.D. Wyner, The wire-tap channel, *Bell System Technical Journal*, Vol. 54, No. 8, 1975, pp. 1355–1387.