Using Channel Noise for Information Theoretic Security

CSE540: Research in Information Security, Monsoon 2013
Report by P Vidyadhar Rao

December 2, 2013

Abstract

The fundamental problem in cryptography is secure tranmission of a message between two legitimate users (the sender Alice and the receiver Bob) over an insecure communication channel such that an enemy (Eve) with access to the channel is unable to get useful information about the message being sent. We review some of the most famous results related to this problem from an information theoretic security perspective using channel noise. It turns out that "*Noise is Nice*": even when the eavesdropper has a better channel than the legitimate users, it is possible to have strictly positive secrecy capacity.

1 Introduction

A number of existing and emerging applications require a key distribution mechanism to selectively broadcast confidential messages to legitimate receivers. For example, in pay-TV systems, a content provider wishes to selectively broadcast certain content to a subset of customers who have subscribed to it. An online key distribution mechanism enables the service provider to distribute a decryption key to these legitimate receivers while securing it from potential eavesdroppers. The content can be encrypted via standard cryptographic protocols, so that only customers who have access to the decryption key can view it. In the absence of such a mechanism, current solutions rely on variants of traditional public key cryptography [1]. Apart from the unproven intractability assumptions, the computational cryptography presents some more disadvantages. For example, the security of a cryptographic protocol is measured by whether it survives to a set of attacks or not. It is practically infeasible to design public key cryptosystems which can survive against all attacks. Moreover, because of the nature of limited resources on wireless sensor nodes, many researchers have conducted different techniques to propose different types of key distribution mechanisms [2] [3]. To this end, information theoretic security perspective looks a promising direction to design strong protocols for secure transmission.

Information theoretic security was introduced by Shannon [4] where, Eve has direct access to the insecure channel i.e., receives a perfect copy of the ciphertext C. C is obtained by Alice as a function of the plaintext M and a secret key K, shared by Alice and Bob. Shannon defines the notion of perfect secrecy subject to the condition that $I(M;C) = 0$, where $I(.;.)$ denotes mutual information between its two arguments. Perfect secrecy implies that C received by Eve does not provide any additional information about the source message M. Notice that in this definition of a security, no assumption about the enemy's computational power is made, therefore making the information-theoretic security more desirable in cryptography than computational security. Under this model, it is proved that perfect secrecy can be achieved only when the entropy of the shared private key K is at least equal to the entropy of the message itself (i.e., $H(K) \geq H(M)$, where $H(.)$ denotes the entropy of its argument), making perfect secrecy *unachievable* in practice!

In the remainder of this report, we look into the some of the famous works that are aimed at making perfect secrecy achievable in practice. In particluar, we focus on examples where communication channel is assumed to be noisy. It turns out that information-theoretic security is achievable in a noisy insecure channel even when Alice-Eve channel is less noisy than the Alice-Bob channel.

2 The Wiretap Channel

One of the features in Shannon's model that leads to impossible result is that the enemy Eve has perfect access to the ciphertext C i.e., the channel from Alice to Eve has the same capacity as the channel from Alice to Bob. Wyner [5] and Csiszár and Korner [6] introduced the wiretap channel and established the possibility of creating an almost perfectly secure link without relying on secret keys. In the wiretap channel, the legitimate users communicate over a main channel and an evaesdropper has access to the messages over a wiretap channel (see Fig.1). The main idea is to allow some independent random noise in the transmitted message, and still achieve reliable transmission at some positive rate with perfect secrecy as long as the wiretap channel is a degraded version of the main channel.

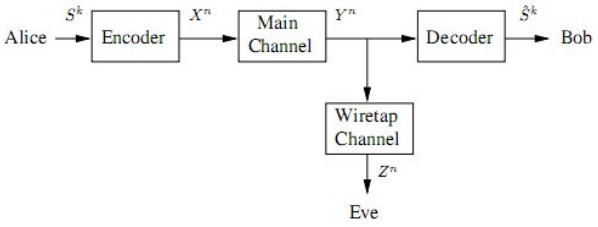

Figure 1: The Wiretap Channel Model

Definition 1: The Wyner's model is defined by the following:

- The source is the sequence $\{S_i\}_{i=1}^{\infty}$, where S_i are i.i.d random variables that take values in the finite set S. Let $H(S_i) = H(S)$;
- The main channel is a discrete memoryless channel(DMC) with finite input alphabet X, finite output alphabet Y and transition probability $Q_M(y|x), x \in X, y \in Y$. Let C_M denote the main channel capacity; For n vectors, $Q_M^{(n)}(y|x) = \prod_{i=1}^{n} Q_M(y_i|x_i)$
- The wiretap channel is a DMC with input alphabet Y, finite output alphabet Z and transition probability $Q_W(z|y), y \in Y, z \in Z$. Let C_W denote the main channel capacity;For n vectors, $Q_W^{(n)}(z|y) = \prod_{i=1}^{n} Q_W(z_i|y_i)$.
- The channel between Alice and Eve is also a DMC, with transition probability $Q_{MW}(z|x) = \sum_{y \in Y} Q_M(y|x)Q_W(z|y)$. the capacity of the channel Q_{MW} is denoted by C_{MW};
- The encode, with parameters (k, n) is a function $e : S^k \to X^n$ and the decoder is a function $d : Y^n \to S^k$.

Definition 2: We define a quantity to measure the ability of Bob to read properly the confidential messages sent by Alice through the main channel.

- For $\hat{S} = (\hat{S_1}, \ldots, \hat{S_k}) = d(Y)$, the error-rate is defined by $P_e = \frac{1}{k} \sum_{i=1}^{k} P(S_i \neq \hat{S}_i)$.

Definition 3: Let Y^n and Z^n be the output of the channels $Q_M^{(n)}$ and $Q_{MW}^{(n)}$, repectively, when the input is X^n. The equivocation of the source at the output of the wiretap channel is defined by: $\Delta = \frac{1}{k} H(S^k|Z^n)$.

We refer to the encoder-decoder described in Definition 1, 2 and 3 as a (k, n, Δ, P_e) encoder-decoder. Ideally, we want the channel to have a small error-rate, while keeping Eve's equivocation high. Thus, the first question that arises is the following.

- *Is it possible to communicate over the main channel at a transmission rate R with small error-rate, while keeping Eve with no significant information about the confidential messages sent through Q_M?*

Wyner characterizes the region of all (R, d) achievable pairs as follows:

Definition 4: For $R > 0$ and $d > 0$, we say that the pair (R, d) is achievable if, for every $\epsilon > 0$, there exists an (k, n, Δ, P_e) encoder-decoder such that:

- $k.\frac{H_S}{n} \geq R - \epsilon$
- $\Delta \geq d - \epsilon$
- $P_e \leq \epsilon$

Let $\Re$ denote set of all achievable pairs. In order to characterise the set $\Re$, we need to study the following quantity.

Definition 5: Let $p_X(x), x \in X$ be a probability mass function and let $P(R)$ denote the set of all distributions p_X such that $I(X;Y) \geq R$. For $0 \leq R \leq C_M$, let

$\Gamma(R) = sup_{p_X \in P(R)} I(X;Y|Z)$. Because, for any distribution p_X on X, the corresponding X, Y and Z form a Markov chain, we have that

$\Gamma(R) = sup_{p_X \in P(R)} [I(X;Y) - I(X;Z)]$.

Lemma 1: For $0 \leq R \leq C_M$, $\Gamma(R)$ satisfies the following:

- $\forall R$, $\exists p_X \in P(R)$ such that $I(X;Y|Z) = \Gamma(R)$;
- $\Gamma(R)$ is a concave function of R.
- $\Gamma(R)$ is nonincreasing in R.
- $\Gamma(R)$ is continuous in R.
- $C_M - C_{MW} \leq \Gamma(R) \leq C_M$

Theorem 1: The Wyner's main result on the set of all acheivable pairs is given by

$\Re = \left\{ (R,d) : 0 \leq R \leq C_M, 0 \leq d \leq H_S, \frac{d}{H_S} \leq \frac{\Gamma(R)}{R} \right\}$

Definition 6: The secrecy capacity of the channel pair (Q_M, Q_W) is defined by $C_S = max_{(R,H_S) \in \Re} R$.

Theorem 2: If $C_M > C_{MW}$, there exists a unique solution C_S of $C_S = \Gamma(C_S)$.

Further C_S is the maximum R s.t $(R, H_S) \in \Re$ and verifies $0 < C_M - C_{MW} \leq \Gamma(C_M) \leq C_S \leq C_M$. Here, it is required that $C_M > C_{MW}$ to have strictly positive secrecy capacity. This means that, in order to be able to communicate with perfect secrecy, Alice and Bob must have a better channel than the wiretap channel.

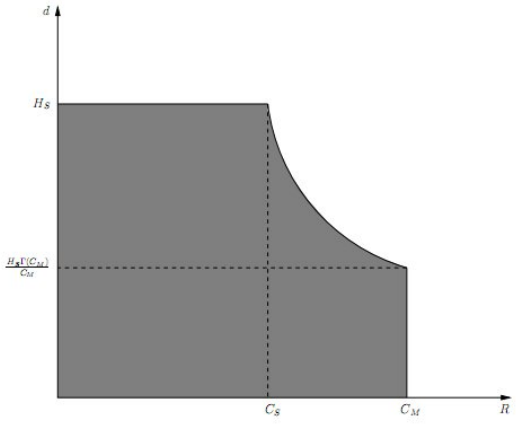

Figure 2: Region of (R, d) achievable pairs.

3 Using Public discussion to achieve Perfect Secrect Key

Wyner's work is limited as it demands that Alice and Bob have significant advantage over Eve. Maurer [7], proposed a new model and showed that a strictly positive secrecy capacity is possible, even if Eve's channel is stronger than the legitimate user's channel. The main feature about the Mauer's model is that a public insecure channel (yet authenticated) is used to generate a secret key.

Definition 7: The broadcast channel of interest given in Fig.3 is defined as:

- The source is the sequence $\{S_i\}_{i=1}^{\infty}$, where S_i is a binary random variable, $\forall i$;
- The main channel has a finite input X, and a finite output alphabet Y.
- The wiretap channel has the same input as the main channel, and a finite output alphabet Z.
- The channel behaviour is completely specified by the conditional probability $P(Y = y, Z = z|X = x)$, which we refer to as $P_{YZ|X}$;
- The encode, with parameters (k, n) is a function $e : \{0,1\}^k \rightarrow X^n$, where R is the rate and $k = nR$; and the decoder is a function $d : Y^n \rightarrow \{0,1\}^k$.

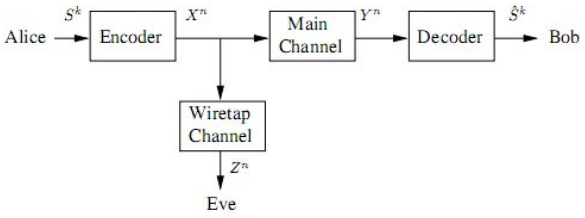

Figure 3: Broadcast channel without a public channel.

Definition 8: The secrecy capacity of a broadcast channel specified by $P_{YZ|X}$ is the maximum rate R for which, for every $\epsilon > 0$, for all sufficiently large n, there exists an encoder-decoder such that for S uniformly distributed over $\{0,1\}^k$ the following two conditions are satisfied:

- $P(d(Y) \neq S) < \epsilon$, where $X = e(S)$
- $\frac{1}{k} H(S|Z^n) > 1 - \epsilon$

Consider a broad channel for which both the main and the wiretap channel are independent binary symmetric channels, i.e, $P_{YZ|X} = P_{Y|X} P_{Z|X}$ and

$$\begin{aligned} P_{Y|X}(y|x) &= 1 - \epsilon \qquad \text{if } x = y \\ &= \epsilon \qquad \text{if } x \neq y \end{aligned}$$

$$\begin{aligned} P_{Z|X}(y|x) &= 1 - \delta \qquad \text{if } x = z \\ &= \delta \qquad \text{if } x \neq z \end{aligned}$$

Without loss of genereality, consider the case $\epsilon < \frac{1}{2}, \delta < \frac{1}{2}$. Denote this channel by $D(\epsilon, \delta)$. The secrecy capacity of this channel is only strictly positive if the legitimate user‘s channel is better than Eve‘s channel.

Lemma 2: The secret capacity of the binary broadcast channel $D(\epsilon, \delta)$ is given by:

$$\begin{aligned} C_S(D(\epsilon, \delta)) &= h(\delta) - h(\epsilon) \qquad \text{if } \delta > \epsilon \\ &= 0 \qquad \text{otherwise} \end{aligned}$$

where $h(p)$ is the binary entropy function, i.e. $h(p) = -p \log(p) - (1-p) \log(1-p)$.

To overcome the need of an advantage of the legitimate users over the eavesdropper, we use a public channel (see Fig.4), insecure but with unconditional secure authentication. It is assumed that Eve can listen to the communication over the public channel, but cannot perform an identity spoofing attack.

Definition 9: The secrecy capacity with public discussion, denoted $\hat{C}(P_{YZ|X})$ is the secrecy capacity of the broadcast channel defined in Definition 7 with the additional feature that Alice and Bob can communicate over an insecure (yet authenticated) public channel.

Theorem 3: The secrecy capacity with public discussion of a broadcast channel is given by

$\hat{C}(P_{YZ|X}) = h(\epsilon + \delta - 2\delta\epsilon) - h(\epsilon)$.

Moreover, $\hat{C}(P_{YZ|X})$ is strictly positive unless $\epsilon = 0.5, \delta = 0$ or $\delta = 1$. i.e., unless X and Y are statistically independent or Z uniquely determines X.

The idea is to construct a conceptual broadcast channel similar to the model of Wyner [5], such that the conceptual main channel is equivalent to the real main channel between Alice and Bob, and the conceptual wiretap channel is a cascade of the real main channel and the real wiretap channel.

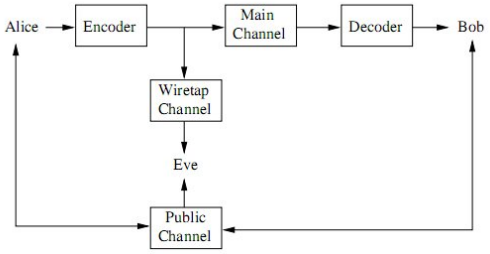

Figure 4: Broadcast channel with a public channel.

Consider the following general key agreement problem. Alice, Bob and Eve know random variables X,Y and Z, respectively, with joint probability distribution P_{XYZ}. Assume that Eve has no information about X and Y other than through her knowledge of Z, i.e. if T represents all the information that Eve has, then $I(XY;T|Z) = 0$. Alice and Bob share no secret key initially (other than a short key required for authentication in the public channel), but they are assumed to know P_{XYZ} , or at least an upper bound on the quality of Eves channel. Assume also that Eve knows the protocol and the codes used.

Without loss of generality, consider only protocols in which Alice sends messages at odd steps $(C_1, C_3, ...)$ and Bob sends messages at even steps $(C_2, C_4, ...)$. At the end of the t-step protocol, Alice computes a key S as a function of X and $C^t = [C_1, ..., C_t]$, and Bob computes a key $S^{'}$ as a function of Y and C^t.

Definition 10: A secret key agreement protocol as described above is (ϵ, δ)-secure if, for some specified (small) ϵ and δ, the following conditions hold:

1. For odd i, $H(C_i|C^{i-1}X) = 0$; and for even i, $H(C_i|C^{i-1}Y) = 0$;
2. $H(S|C^tX) = 0$; and $H(S^{'}|C^tY) = 0$;
3. $P(S \neq S^{'}) \leq \epsilon$;
4. $I(S; C^tZ) \leq \delta$;

Conditions 1-2 guarantee that Alice and Bob have no uncertainty regarding the protocol procedures. Condition

3 guarantees that Alice and Bob agree on the same key with probability $1-\epsilon$. Finally, condition 4 guarantees that, given that Eve knows all the messages exchanged between Alice and Bob over the public channel during the protocol and also the output of her channel, the information on the key that Eve has is upperbounded by δ.

Theorem 4: For every, (ϵ, δ)-secure key agreement protocol, we have that $H(S) \leq min[I(X;Y), I(X;Y|Z)] + \delta + h(\epsilon) + \epsilon \log(|S| - 1)$

To be able to provide a lower bound on the key size, we need to make further assumptions. Consider the case when Alice, Bob and Eve receive $X^N = [X_1, \ldots, X_N]$, $Y^N = [Y_1, \ldots, Y_N]$ and , $Z^N = [Z_1, \ldots, Z_N]$, where $P_{X^N Y^N Z^N} = \prod_{i=1}^{N} P_{X_i Y_i Z_i}$. We define the secret key rate as follows:

Definition 11: The secret key rate of X and Y with respect to Z, denoted $S(X;Y||Z)$, is the maximum rate R such that, for every $\epsilon > 0$, there exists a protocol, for sufficiently large n, satisfying conditions 1-3 in Definition 10 (with X and Y replaced by X^n and Y^n, respectively) and also the two following conditions:

- $\frac{1}{n} I(S; C^t Z) \leq \epsilon$;
- $\frac{1}{n} H(S) \geq R - \epsilon$

Theorem 5: The secret key rate $S(X;Y||Z)$ verifies

- $S(X;Y||Z) \leq min[I(X;Y), I(X;Y|Z)]$;
- $S(X;Y||Z) \geq max[I(Y;X) - I(Z;X), I(X;Y) - I(Z;Y)]$.

The upper bound for the secret key rate in the previous theorem shows that if Eve has less information about Y than Alice or less information about X than Bob, then such a difference of information can be exploited. The secrecy capacity with public discussion of a general broadcast channel is given below.

Theorem 6: The secrecy capacity with public discussion, $\hat{C}(P_{YZ|X})$, of a broadcast channel specified by $P(YZ|X)$ verifies:

$$max_{P_X} S(X;Y||Z) \leq \hat{C}(P_{YZ|X}) \leq min[max_{P_X} I(X;Y), max_{P_X} I(X;Y|Z)].$$

The secrecy capacity with public discussion, shows that, even if the eavesdropper has a better channel than the legitimate users, perfect secure communication can still be performed.

4 Outlook

Practical interest of information-theoretic security has two striking benefits over conventional cryptography: (1) no computational assumptions, and (2) no keys and hence no key distribution. (1) is attractive to governments who are concerned with long-term security and worried about quantum computing. (2) is attractive in the world where vulnerable, low-power wireless devices are proliferating and key-distribution and key-management are insurmountable obstacles to security.

The practical challenge is to realize a secrecy capacity, meaning ensure by physical means that the evaesdropper channel is noisier than the receiver one. As per my knowledge, thus far, I have come across only one such attempt: Government-sponsored Ziva Corporation [8] is using optical techniques to build a receiver channel in such a way that wiretapping results in a degraded channel. It makes a strong case that we need definitions that yield information theoretic security in applications, and also that we need constructive results yielding practical schemes achieving secrecy under these definitions.

References

[1] Diffie, Whitfield, "The first ten years of public-key cryptography," 1988.

[2] Chen, Chi-Yuan and Chao, Han-Chieh, "A survey of key distribution in wireless sensor networks," *Security and Communication Networks*, 2011.

[3] Singhal, Mukesh and Bai, Rendong and Lin, Yun and Wang, Yongwei and Yang, Mengkun and Zhang, Qingyu, "Key management protocols for wireless networks."

[4] Shannon, Claude E, "The mathematical theory of communication," 1949.

[5] Wyner, Aaron D, "The wire-tap channel," 1975.

[6] Csiszár, Imre and Korner, Janos, "Broadcast channels with confidential messages," 1978.

[7] Maurer, Ueli M, "Secret key agreement by public discussion from common information," 1993.

[8] Ziva Corporation, "http://www.ziva-corp.com/."