Protocols for Secret Key Agreement by Public Discussion Based on Common Information

Ueli M. Maurer

Institute for Theoretical Computer Science
ETH Zürich
CH-8092 Zürich, Switzerland
Email address: maurer@inf.ethz.ch

Abstract. Consider the following scenario: Alice and Bob, two parties who share no secret key initially but whose goal it is to generate a (large amount of) information-theoretically secure (or unconditionally secure) shared secret key, are connected only by an insecure public channel to which an eavesdropper Eve has perfect (read) access. Moreover, there exists a satelite broadcasting random bits at a very low signal power. Alice and Bob can receive these bits with certain bit error probabilities ϵ_A and ϵ_B, respectively (e.g. $\epsilon_A = \epsilon_B = 30\%$) while Eve is assumed to receive the same bits much more reliably with bit error probability $\epsilon_E \ll \epsilon_A, \epsilon_B$ (e.g. $\epsilon_E = 1\%$). The errors on the three channels are assumed to occur at least partially independently. Practical protocols are discussed by which Alice and Bob can generate a secret key despite the facts that Eve possesses more information than both of them and is assumed to have unlimited computational resources as well as complete knowledge of the protocols.

The described scenario is a special case of a much more general setup in which Alice, Bob and Eve are assumed to know random variables X, Y and Z jointly distributed according to some probability distribution P_{XYZ}, respectively. The results of this paper suggest to build cryptographic systems that are provably secure against enemies with unlimited computing power under realistic assumptions about the partial independence of the noise on the involved communication channels.

E.F. Brickell (Ed.): Advances in Cryptology - CRYPTO '92, LNCS 740, pp. 461-470, 1993.

1. Introduction

One of the fundamental problems in cryptography is the transmission of a message M from a sender (referred to as Alice) to a receiver (Bob) over an insecure communication channel such that an enemy (Eve) with access to this channel is unable to obtain useful information about M.

In the classical model of a cryptosystem introduced by Shannon [9], Eve has perfect access to the insecure channel; thus she is assumed to receive an identical copy of the ciphertext C received by the legitimate receiver Bob, where C is obtained as a function of the plaintext message M and a secret key K shared by Alice and Bob. Shannon defined a cipher system to be perfect if the ciphertext is statistically independent of the plaintext or, in information-theoretic terms, if the ciphertext gives no information about the plaintext:

$$I(M;C) = 0.$$

When a perfect cipher is used to encrypt a message M, an enemy can do no better than guess M without even looking at the ciphertext C.

It is assumed that the reader is familiar with the fundamentals of information theory, in particular with the entropy $H(X)$ of a random variable X, the conditional entropy of X given Y, $H(X|Y)$, and the mutual information between X and Y defined as $I(X;Y) = H(X) - H(X|Y)$. We refer to [4] for an introduction to information theory.

Shannon gave as a simple example of a perfect cipher the well-known one-time pad which is completely impractical for most applications where only a short secret key is available. 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 or, more precisely, when

$$H(K) \geq H(M). \tag{1}$$

Almost all presently-used ciphers are based on Shannon's model but have only a short secret key; they can therefore theoretically be broken, for instance by an exhaustive key search. The goal of designing such a practical cipher is to guarantee that there exists no efficient algorithm for breaking it, for a reasonable definition of breaking. However, for no existing cipher can the computational security be proved without invoking an unproven intractability hypothesis.

Perfect secrecy on the other hand is often prejudged as being impractical because of Shannon's pessimistic inequality (1). It is one of the goals of this paper to relativize this pessimism by pointing out that Shannon's apparently innocent assumption that, except for the secret key, the enemy has access to precisely the same information as the legitimate receiver, is much more restrictive than has generally been realized.

The key to perfect secrecy without a shared secret key K satisfying (1) is to modify Shannon's model such that the enemy cannot receive precisely (albeit almost) the same information as the legitimate receiver. Two previous approaches based on this idea are quantum cryptography introduced by Wiesner and put forward by Bennett, Brassard *et al.* [1], and Maurer's randomized cipher [7] which makes use of a public random string that is too long to be read entirely in feasible time. Both these approaches are impractical at present.

Another approach is due to Wyner [11] and subsequently Csiszár and Körner [5] who considered a scenario in which the enemy Eve is assumed to receive messages transmitted by the sender Alice over a channel that is noisier than the legitimate receiver Bob's channel. The assumption that Eve's channel is worse than the main channel is unrealistic in general. The results of this paper demonstrate that this unrealistic assumption is unnecessary if Alice and Bob can also communicate over a completely insecure public channel.

In this paper, the broadcast channel scenario is generalized to a scenario where Alice, Bob and Eve know random variables X, Y and Z, respectively, jointly distributed according to some probability distribution P_{XYZ}, and where Alice and Bob can also communicate over a public channel.

Note that the need for a public channel entails no significant loss of practicality in a cryptographic context because the channel need not provide secrecy. It is assumed, however, that all messages sent over the public channel can be received by Eve without error, but that she cannot modify messages or introduce fraudulent messages without being detected. If this last assumption cannot realistically be made, authenticity and data integrity can be ensured by using an unconditionally secure authentication scheme, for instance that of [10] based on universal hashing, which requires that Alice and Bob share a short secret key initially. In this case, the purpose of our protocols is to stretch (rather than to generate) a secret key unconditionally securely. Part of the generated key can be used for authentication in a subsequent instance of the protocol.

The use of a public channel by two parties for extracting a secret key from an initially shared partially secret string was previously considered by Leung-Yan-Cheong [6] and independently by Bennett, Brassard and Robert [3].

This paper is concerned with key distribution as well as encryption. An unconditionally secure shared secret key generated by one of our protocols can be used as the key sequence in the one-time pad, thus achieving (virtually) perfect secrecy of the transmitted messages.

2. Secret Key Agreement by Public Discussion

Consider the following general key agreement problem. Assume that Alice, Bob and Eve know random variables X, Y and Z, respectively, with joint probability distribution P_{XYZ}, and that Eve has no information about X and Y other than through her knowledge of Z. More precisely, $I(XY;T|Z) = 0$ where T summarizes Eve's complete information about the universe. X, Y and Z take on values in some finite alphabets $\mathcal{X}, \mathcal{Y}$ and $\mathcal{Z}$, respectively. 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}. 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.

Alice and Bob use a protocol in which at each step either Alice sends a message to Bob depending on X and all the messages previously received from Bob, or vice versa (with X replaced by Y). Without loss of generality, we consider only protocols in which Alice sends messages at odd steps $(C_1, C_3, \ldots)$ and Bob sends messages at even steps $(C_2, C_4, \ldots)$. Moreover, we can restrict the analysis to deterministic protocols since a possible randomizer which Alice's and/or Bob's strategy and messages might depend on can be considered as part of X and Y, respectively. In other words, Alice and Bob can without loss of generality extend their known random variables X and Y, respectively, by random bits that are statistically independent of X, Y and Z. At the end of the t-step protocol, Alice computes a key S as a function of X and $C^t \triangleq [C_1, \ldots, C_t]$ and Bob computes a key S' as a function of Y and C^t. Their goal is to maximize $H(S)$ under the conditions that S and S' agree with very high probability and that Eve has very little information about S. More formally,

$$H(C_i|C^{i-1}X) = 0 \tag{2}$$

for odd i,

$$H(C_i|C^{i-1}Y) = 0 \tag{3}$$

for even i,

$$H(S|C^tX) = 0 \tag{4}$$

and

$$H(S'|C^tY) = 0, \tag{5}$$

and it is required that

$$P[S \neq S'] \leq \epsilon \tag{6}$$

and

$$I(S;C^tZ) \leq \delta \tag{7}$$

for some specified (small) δ and ϵ.

By Fano's Lemma (cf. [4], p. 156) condition (6) implies that

$$H(S|S') \leq h(\epsilon) + \epsilon \log_2(|\mathcal{S}| - 1) \tag{8}$$

where $|\mathcal{S}|$ denotes the number of distinct values that S takes on with non-zero probability. Note that $H(S|S') \to 0$ as $\epsilon \to 0$.

If one requires that $P[S \neq S'] = 0$ and $I(S;C^t) = 0$ (i.e., that $\epsilon = 0$ in (6) and $\delta = 0$ in (7)) it appears obvious that $I(X;Y)$ is an upper bound on $H(S)$. It appears to be similarly obvious that $H(S) \leq I(X;Y|Z) = I(XZ;YZ) - H(Z)$ because even under the assumption that Alice and Bob could learn Z, the remaining information shared by Alice and Bob is an upper bound on the information they can share in secrecy. The following theorem, which is proved in [8], summarizes these results.

Theorem 1. *For every key agreement protocol satisfying (2)-(5),*

$$H(S) \leq I(X;Y|Z) + H(S|S') + I(S;C^tZ).$$

In particular,

$$H(S) \leq I(X;Y) + H(S|S') + I(S;C^t).$$

The following corollary follows from Theorem 1, inequality (8) and from $I(S;C^t) \leq I(S;C^tZ)$. It should be pointed out that $I(X;Y) < I(X;Y|Z)$ is possible.

Corollary 2. *For every key agreement protocol satisfying (2)-(7),*

$$H(S) \leq \min[I(X;Y), I(X;Y|Z)] + \delta + h(\epsilon) + \epsilon \log_2(|\mathcal{S}| - 1).$$

3. The Secret Key Rate

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

$$P_{X^NY^NZ^N} = \prod_{i=1}^{N} P_{X_iY_iZ_i}$$

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

For such a scenario of independent repetitions of a random experiment, which is well motivated by models such as discrete memoryless sources and channels previously considered in information theory, the quantity that appears to be of most interest from an information-theoretic point of view is defined below.

Definition. The *secret key rate of X and Y with respect to Z*, denoted $S(X;Y||Z)$, is the maximum rate at which Alice and Bob can agree on a secret key S while keeping the rate at which Eve obtains information arbitrarily small, i.e., it is the maximal R such that for every $\epsilon > 0$ there exists a protocol for sufficiently large N satisfying (2)-(6) with X and Y replaced by X^N and Y^N, respectively, satisfying

$$\frac{1}{N} I(S; C^t Z^N) \leq \epsilon,$$

and achieving

$$\frac{1}{N} H(S) \geq R - \epsilon.$$

Before deriving lower bounds on $S(X;Y||Z)$ we state the following theorem, which is an immediate consequence of Corollary 2.

Theorem 3. *The secret key rate of X and Y with respect to Z is upper bounded by*

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

The following theorem (cf. [8] for a proof) states a nontrivial lower bound on the secret key rate. If it is either the case that Eve has less information about Y than Alice or, by symmetry, less information about X than Bob, then such a difference of information can be exploited.

Theorem 4. *The secret key rate of X and Y with respect to Z is lower bounded by*

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

Theorem 4 demonstrates that the upper bound in Theorem 3 is tight if either $P_{YZ|X} = P_{Y|X} \cdot P_{Z|X}$ or $P_{XZ|Y} = P_{X|Y} \cdot P_{Z|Y}$. The lower bound of Theorem 4 is not tight in general as will be demonstrated in the next section. In particular, the lower bound of Theorem 4 is 0 for the situation described in the abstract of the paper. There exist protocols with several rounds of interaction between Alice and Bob which are superior to single-round protocols like the one used in the proof of Theorem 4 (cf. [8]).

4. Binary Symmetric Random Variables

In this section the case of symmetrically distributed binary random variables is considered. One way of generating such a set X, Y, Z is by generating a random bit R according to

$$P_R(0) = P_R(1) = 1/2 \tag{9}$$

and "sending" R over three *independent* binary symmetric channels C_A, C_B and C_E with error probabilities ϵ_A, ϵ_B and ϵ_E, respectively, i.e., P_{XYZ} is defined by

$$P_{XYZ|R} = P_{X|R} \cdot P_{Y|R} \cdot P_{Z|R} \tag{10}$$

where $P_{X|R}(x, r) = 1 - \epsilon_A$ if $x = r$ and ϵ_A else, $P_{Y|R}(y, r) = 1 - \epsilon_B$ if $y = r$ and ϵ_B else and $P_{Z|R}(z, r) = 1 - \epsilon_E$ if $z = r$ and ϵ_E else.

Consider now an arbitrary probability distribution P_{XYZ} over $\{0, 1\}^3$ satisfying the symmetry condition

$$P_{XYZ}(x, y, z) \;=\; P_{XYZ}(\overline{x}, \overline{y}, \overline{z}) \tag{11}$$

for $x, y, z \in \{0, 1\}$, where $\overline{c}$ denotes the complement of a binary variable c. Note that condition (11) implies that X, Y and Z are symmetrically distributed. One can prove (see [8]) that every set X, Y and Z of random variables satisfying (11) and for which not exactly for one of the pairs $[X, Y]$, $[X, Z]$ and $[Y, Z]$ the two random variables are statistically independent, can be generated according to (9) and (10) for some ϵ_A, ϵ_B and ϵ_E.

As one realistic scenario where X, Y and Z with probability distribution P_{XYZ} satisfying (11) are available for two parties and an enemy, consider a satellite broadcasting random bits at a very low signal-to-noise ratio such that even an enemy Eve with a receiving antenna that is much larger and more sophisticated than Alice's and Bob's antenna cannot receive the bits without error. Note that P_{XYZ} satisfies the given condition also when the channels C_A, C_B and C_E are dependent, as one would realistically have to assume. The following theorem has been proved in [8].

Theorem 5. *Let X, Y and Z be binary random variables generated according to (9) and (10). Then*

$$S(X; Y||Z) \;\geq\; \max[h(\epsilon_A + \epsilon_E - 2\epsilon_A\epsilon_E),\; h(\epsilon_B + \epsilon_E - 2\epsilon_B\epsilon_E)] - h(\epsilon_A + \epsilon_B - 2\epsilon_A\epsilon_B).$$

The lower bound of Theorem 5 vanishes unless either $\epsilon_A < \epsilon_E$ or $\epsilon_B < \epsilon_E$, i.e., unless either Alice's or Bob's channel is superior to Eve's channel. It is somewhat surprising that even when Eve's channel is much more reliable that both Alice's and Bob's channel, secret key agreement is possible.

The proof of Theorem 4 in [8] illustrates that by sending $X_i + V_i$ over the public channel, where X_i is the ith random bit received by Alice and where addition is modulo 2, Alice can send the bit V_i over a conceptual broadcast channel to Bob and Eve such that Bob receives V_i as if it were sent over a cascade of Alice's and Bob's channel (bit error probability $\epsilon_A + \epsilon_B - 2\epsilon_A\epsilon_B$) and Eve receives V_i as if it were sent over a cascade of Alice's and Eve's channel (bit error probability $\epsilon_A + \epsilon_E - 2\epsilon_A\epsilon_E$).

In order to share a secret key with Bob, Alice randomly selects a codeword V^N from the set of codewords of an appropriate error-correcting code $\mathcal{C}$ with codewords of length N and sends it to Bob (and also to Eve) over the described conceptual broadcast channel. The key to achieving a positive secret key rate even if both $\epsilon_A > \epsilon_E$ and $\epsilon_B > \epsilon_E$ is for Bob to accept a received word only if he can make a very reliable decision about the codeword sent by Alice, i.e., if it is very close to some codeword of the code $\mathcal{C}$, i.e., if the Hamming distance to a codeword is much smaller than the number of errors correctable by an optimal decoder for the code. For each received block Bob announces over the public channel whether he accepts or rejects it.

The key observation in the above protocol is that although Eve receives codewords V^N more reliably than Bob on the average, her conceptual channel may nevertheless be worse (for appropriate choices of a code $\mathcal{C}$ and for an appropriate reliability decision) than Bob's channel, if one averages only over those instances accepted by Bob. Because consecutive uses of the channel are independent, the words discarded by Bob are also useless for Eve.

The special case of a repeat code was considered in [8]. Alice sends each bit N times over the conceptual channel, and Bob accepts a received word if and only if all the bits are equal. Although this scheme demonstrates that secret key agreement is possible even if $\epsilon_A > \epsilon_E$ and $\epsilon_B > \epsilon_E$, it is extremely inefficient when ϵ_E is considerably smaller than both ϵ_A and ϵ_B. The reason is that in order to arrive at a situation where Bob's channel is better than Eve's channel if averaged over those instances accepted by Bob, a large block length N must be used in which case the probability that no error occurs within a block and thus the block is accepted by Bob can be extremely small. It is one of the purposes of this paper to describe protocols that are much more efficient than the protocol discussed in [8].

An important observation towards improving the key agreement rate is that several rounds of a protocol as described above can be used by Alice and Bob to continuously increase the reliability of the shared string at the expense of shrinking it. In a first step, and even in some subsequent steps, it is not required that Bob knows Alice's bits more reliably than Eve; it is sufficient that Eve's advantage is reduced in every step. Hence using several protocol steps with short blocks allows to achieve comparable bit error probabilities for the finally shared

string as if a long repeat code were used, but with a much larger rate.

Consider as an example a simple $N = 3$ repeat code. Bob accepts a received block of length 3 if and only if all three bits agree, and announces which blocks he accepts. The probability of accepting a block is $\geq 1/4$; hence the strings held by Alice and Bob are shrunk by this step by at most a factor 12. Alice and Bob can use the same step on the resulting string repeatedly, each time decreasing its length by at most a factor 12 while increasing the bit agreement probability. It is straight-forward to verify that when k steps are used, Bob's and Eve's bit error probabilities when guessing the bits of Alice's final string are precisely the same as if a repeat code of length 3^k had been used in the above described basic protocol, but that the expected rate at which random secret key bits are extracted is exponentially larger in the new protocol.

Example. Let $\epsilon_A = \epsilon_B = 0.47$ and let Eve's channel be 100 times less noisy, i.e., have 100 times greater capacity. From $1-h(\epsilon_E) = 100\cdot(1-h(\epsilon_A))$ we obtain $\epsilon_E = 0.2093$. A repeat code of length 243 yields bit error probabilities 0.148 and 0.193 for Bob and Eve, but the probability that a block is accepted by Bob is not significantly larger than 2^{-242}. On the other hand, 5 consecutive applications of the described step with a code of length 3 allow to achieve the same bit error probabilities, but only an expected number of at most $12^5 < 250.000$ (actually much less) bits are required for generating one bit shared with the mentioned bit error probabilities.

Of course, additional protocol steps are required for exploiting the advantage over Eve achieved by this protocol and reducing the bit error probability of the final shared string. For example, error correcting codes can be used to remove the errors between Alice's and Bob's string, and universal hashing as described in [3] can be used to reduce Eve's information.

It should be pointed out that for given assumed ratios of the noise power on the three channels, the signal power is a free parameter; thus ϵ_A can be chosen arbitrarily. The larger ϵ_A, the smaller is the signal power and hence the larger can the satelite's bit transmission rate be chosen.

The use of repeat codes as described above, and more generally of linear error-correcting codes, is equivalent to the exchange of parity checks of the stored string over the public channel, without generating and encoding random bits, and using as a new string some orthogonal parity checks. Reconciliation protocols based on the exchange of parity checks were also discussed in [2].

A further improvement over the basic use of repeat codes described above is for Bob to also accept instances for which a decision about the bit sent by Alice is less reliable than if N identical bits were received. In such a scenario, Bob informs Alice (and Eve) about the number of errors he has received in a block, assuming that his majority decision is correct.

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 and J.-M. Robert, How to reduce your enemy's information, *Advances in Cryptology - Crypto '85*, Springer Verlag, New York, pp. 468-476.

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

[6] S.K. Leung-Yan-Cheong, Multi-user and wiretap channels including feedback, Tech. Rep. No. 6603-2, Stanford University, Information Systems Lab., July 1976.

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

[8] U.M. Maurer, Secret key agreement by public discussion from common information, to appear in *IEEE Transactions on Information Theory*.

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

[10] 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.

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