On Information Theoretic Security: Mathematical Models and Techniques

Imre Csiszár

Rényi Institute, Budapest

ICITS 2011

Outline

Overview of some models of information theoretically secure transmission over insecure channels and of secret key generation taking advantage of public communication.

Based on new chapter of Csiszár-Körner book, 2nd edition, to appear June 2011.

Attention concentrated on fundamental limits. Emphasis on mathematical techniques.

Brief history

Shannon 1949: first applied information theory to cryptology

Wyner 1975: first studied secure transmission over insecure channels via advanced information theory techniques

Csiszár-Körner 1979: extended Wyner's model

Bennett-Brassard-Robert 1988: showed benefits of public discussion for generating secrecy

Maurer 1993, Ahlswede-Csiszár 1993: basic source and channel models of generating a secret key

Maurer 1994: improved definition of security

Csiszár-Narayan 2004, 2008: multi-party source and channel models

Preliminaries

A random variable (RV) K is ε-secret from another RV Z if

$$I(K \wedge Z) \triangleq H(K) - H(K|Z) \leq \varepsilon.$$

Proposition (Shannon 1949). Let the RVs M and K have common range $\mathbf{K}$ on which a group operation $+$ is defined. If M and K are conditionally independent conditioned on another RV Z, then for $C = K + M$

$$I(M \wedge C|Z) \leq \log|\mathbf{K}| - H(K|Z).$$

Proof: $I(M \wedge C|Z) = H(C|Z) - H(C|MZ) \leq$
$\log|\mathbf{K}| - H(K|MZ) = \log|\mathbf{K}| - H(K|Z)$.
Convenient notation (Csiszár-Narayan 2004):

$$S(K|Z) \triangleq \log|\mathbf{K}| - H(K|Z) \text{ security index.}$$

Small security index $\Rightarrow$ good secret key (SK)

Definition. A RV U is ε-recoverable from a RV V if $\Pr\{U = \varphi(V)\} \geq \varepsilon$ for some function φ.

Let the RVs $V_1, \ldots, V_m$ represent the information available to m parties, thus V_i is the view of party i.

Definition. A RV K is ε-common randomness (ε-CR) for the m parties if K is ε-recoverable from each V_i. The RV K is an ε-secret key (ε-SK) relative to an adversary with view V^* if K is an ε-CR and $S(K|V^*) \leq \varepsilon$.

Mathematical models of generating CR and SK specify permissible protocols the parties may perform to arrive at views $V_1, \ldots, V_m$ from which they can obtain CR/SK as above.

Secure transmission over an insecure channel is a special case of generating SK.

Basic models for $m = 2$ users

Maurer 1993, Ahlswede-Csiszár 1993

Source model: Given i.i.d. repetitions of a triple of (correlated) RVs (X, Y, Z), suppose Alice sees $X^n = X_1 \dots X_n$, Bob sees $Y^n = Y_1 \dots Y_n$, Eve sees $Z^n = Z_1 \dots Z_n$.

Permissible protocols: Alice and Bob generate RVs Q_A, Q_B, independent of each other and of (X^n, Y^n, Z^n), then they communicate over a noiseless public channel, alternatingly sending messages $F_1, F_2, \dots, F_{2r-1}, F_{2r}$, each a function of the current view of Alice resp. Bob: for i odd resp. even, F_i is a function of (X^n, Q_A) resp. (Y^n, Q_B) and of the previous messages $F_1, \dots, F_{i-1}$.

The RVs Q_A, Q_B serve to model randomized choice of the messages F_i.

In models considered here, Eve has full access to the public communication $F = F_1 \dots F_{2r}$ but is unable to tamper with it.

When communication has been completed, Alice's view is $V_A = (X^n, Q_A, F)$, Bob's is $V_B = (Y^n, Q_B, F)$ and Eve's is $V_E = (Z^n, F)$.

Definition. An achievable SK rate is an $R > 0$ such that for each $\varepsilon > 0$, $\delta > 0$ and sufficiently large n, some admissible protocol enables Alice and Bob to generate ε-SK of rate $\frac{1}{n} \log |\mathbf{K}| > R - \delta$. The largest achievable SK rate is the secret key capacity C_{SK}.

Prior to Maurer (1994), instead of security index $< \varepsilon$ only $< \varepsilon n$ was required. The value of SK capacity remains the same (at least for a large class of models), even if one requires $\varepsilon = \varepsilon_n \to 0$ exponentially.

Channel model. Given a discrete memoryless channel (DMC) with one input and two outputs, with matrix

$$W = \{W(y,z|x) : \ x \in \mathbf{X}, y \in \mathbf{Y}, z \in \mathbf{Z}\},$$

Alice selects the DMC inputs $X_1, \ldots, X_n$, Bob and Eve see the outputs $Y_1, \ldots, Y_n$ resp. $Z_1, \ldots, Z_n$.

Permissible protocols similar as before: Alice and Bob generate RVs Q_A, Q_B, then alternatingly send messages over a public channel, depending on their current views: any number of public messages (perhaps zero) may be exchanged between any two instances when Alice sends DMC inputs X_i (which are functions of Q_A and the public messages previously received from Bob).

SK capacity for channel model: same definition as for source model.

Variants of the classic models: the permissible public communication may be constrained, in rate or in the number of rounds, or both. One model whose SK capacity is known admits just one public message, sent by Alice, perhaps of constrained rate. Variants of the basic models not allowing randomization are also of interest.

If no public communication is allowed: channel model $\Rightarrow$ wiretap channel (see later)
source model becomes degenerate: no CR, let alone SK can be generated, except in trivial cases (Gács-Körner 1973)

General approach

For achievability results: Agree first on CR not caring for security (standard information theory techniques may be used, simple or more complex as superposition coding). Then, take a function of this CR whose value is already secure (privacy amplification).

Suitable tool: Secrecy Lemma, consequence of the Extractor Lemma (next slides).

Converses: Use bounds in terms of information measures, manipulate them judiciously, using known (or new) identities/inequalities. One identity has proved remarkably useful.

Extractor lemma

An ε-extractor for a family $\mathcal{P}$ of distributions on a set $\mathbf{U}$ is a mapping $\kappa : \mathbf{U} \to \{1, \ldots, k\}$ such that for each RV U whose distribution belongs to $\mathcal{P}$, the distribution of $\kappa(U)$ is ε-uniform:

$$\sum_{i=1}^{k} \left| P(\kappa^{-1}(i)) - \frac{1}{k} \right| \leq \varepsilon, \quad P \in \mathcal{P}.$$

Lemma (Ahlswede-Csiszár 1998)

If $P(\{u : P(u) \leq 1/d\}) \geq 1 - \eta$ for each $P \in \mathcal{P}$, then for any $\varepsilon > 0$, a randomly selected mapping $\kappa : \mathbf{U} \to \{1, \ldots, k\}$ is an $(\varepsilon + 2\eta)$-extractor for $\mathcal{P}$ with probability $\geq 1 - 2k|\mathcal{P}|e^{-\varepsilon^2(1-\eta)d/2k(1+\varepsilon)}$.

In applications, $|\mathbf{U}|, |\mathcal{P}|$ and d grow exponentially as $n \to \infty$. Then, ε and η may be exponentially small, and $\log k$, „the number of extracted random bits," may grow effectively as rapidly as $\log d$.

Secrecy lemma

Lemma

If for RVs U, V with values in $\mathbf{U}, \mathbf{V}$

$$P_{UV}(\{(u,v):\ P_{U|V}(u|v) \le 1/d\}) \ge 1-\eta^2, \quad \eta \le 1/3,$$

then in case $k \ln(2k|\mathbf{V}|) < \alpha^2 d$, $\alpha \le 1/6$, a randomly selected mapping $\kappa : \mathbf{U} \to \{1,\dots,k\}$ satisfies

$$S(\kappa(U)|V) \le (\alpha+2\eta)\log k + h(\alpha+\eta)$$

with probability $\ge 1 - 2k|\mathbf{V}|e^{-\alpha^2 d/k} > 0$.

Corollary

Let X^n, Z^n be i.i.d. repetitions of a pair of RVs (X, Z), and $V^{(n)}$ any RV with at most e^{nr} possible values. To any $\delta > 0$ there exists $\xi > 0$ such that for each n and $k \le \exp\{n(H(X|Z) - r - \delta)\}$ a randomly selected $\kappa : \mathbf{X}^n \to \{1,\dots,k\}$ gives $S(\kappa(X^n)|Z^n, V^{(n)}) < \exp\{-\xi n\}$, except with doubly exponentially small probability.

Theorem (Maurer 1993, Ahlswede-Csiszár 1993)

For the basic source model

$$I(X \wedge Y) - \min[I(X \wedge Z), I(Y \wedge Z)] \leq C_{SK} \leq I(X \wedge Y|Z).$$

For the basic channel model

$$\max\{I(X \wedge Y) - \min[I(X \wedge Z), I(Y \wedge Z)]\} \leq C_{SK} \leq \max\{I(X \wedge Y|Z)\},$$

with maximum over RVs satisfying $P_{YZ|X} = W$.

Proof: Source model, lower bound: Alice can send a message $F = F(X^n)$ of rate arbitrarily close to $H(X|Y)$ such that X^n is ε-recoverable from (Y^n, F), thus X^n is an ε-CR (even with $\varepsilon = \varepsilon_n \to 0$ exponentially).
Last Corollary $\Rightarrow$ there exists $\kappa : \mathbf{X}^n \to \{1, \ldots, k\}$ with

$$k = \exp\{n(H(X|Z) - H(X|Y) - \delta)\} = \exp\{n[I(X \wedge Y) - I(X \wedge Z) - \delta]\}.$$

To complete the proof, exchange the roles of Alice and Bob.

Upper bound: Suppose K is an ε-CR thus ε-recoverable from both (X^n, Q_A, F) and (Y^n, Q_B, F), and satisfies weak security $S(K|Z^nF) < n\varepsilon$. Then

$$\log|\mathbf{K}| - n\varepsilon \le H(K|Z^nF) \le I(X^nQ_A \wedge Y^nQ_B|Z^nF) + 2(\varepsilon \log|\mathbf{K}| + 1)$$
$$\Rightarrow (1-2\varepsilon)(\log|\mathbf{K}|)/n - \varepsilon - 2/n \le I(X^nQ_A \wedge Y^nQ_B|Z^nF).$$

By induction on the number of public messages,
$I(X^nQ_A \wedge Y^nQ_B|Z^nF) \le$

$$I(X^nQ_A \wedge Y^nQ_B|Z^n) = I(X^n \wedge Y^n|Z^n) = nI(X \wedge Y|Z).$$

Channel model admits emulating the source model, sending i.i.d. $X^n \Rightarrow$ lower bound follows from that for the source model.

Upper bound: again, $I(X^nQ_A \wedge Y^nQ_B|Z^nF)$ has to be bounded, needs a little more effort.

Remark: A single-letter formula for C_{SK} is known in special cases only. Full solution available for restricted source model with one-way public communication (Ahlswede-Csiszár 1993, Csiszár-Narayan 2000).

Wiretap channel (Wyner 1975)

Two DMCs with common input, $W_1 : \mathbf{X} \to \mathbf{Y}$, $W_2 : \mathbf{X} \to \mathbf{Z}$.

Message RV M uniform on set $\mathbf{M}$.
$\to$ channel input X^n (randomized encoding allowed)
$\to$ channel outputs Y^n, Z^n

Requirements: M ε-recoverable from Y^n, and weakly secret from Z^n: $I(M \wedge Z^n) < n\varepsilon$.

Secrecy capacity C_S: largest $\frac{1}{n} \log |\mathbf{M}|$, in limit $n \to \infty$.

Theorem (Csiszár-Körner 1978)

$$C_S = \max[I(V \wedge Y) - I(V \wedge Z)]$$

for RVs $V \to X \to YZ$ with $P_{Y|X} = W_1$, $P_{Z|X} = W_2$; the range of the auxilliary RV V may be assumed not larger than $|\mathbf{X}|$.

Corollary. $C_S = 0$ iff W_2 is less noisy than W_1. (Wyner assumed W_2 was a degraded version of W_1; then $V = X$ suffices.)

Remarks

(i) The weak secrecy condition can be sharpened to $I(M \wedge Z^n) < \varepsilon$ or further, requiring $\varepsilon = \varepsilon_n \to 0$ exponentially; this does not affect secrecy capacity (Csiszár 1996).

(ii) The wiretap channel is equivalent to a channel model of generating SK not allowing public communication. While the SK there has to be only nearly uniform, its slight modification is suitable for M here. In absence of public communication, the full matrix $W(y, z|x)$ need not be known, only

$$W_1(y|x) \triangleq \sum_{z \in \mathbf{Z}} W(y, z|x), \quad W_2(z|x) \triangleq \sum_{y \in \mathbf{Y}} W(y, z|x).$$

(iii) A channel model with public communication may have positive SK capacity also if the corresponding wiretap channel has $C_S = 0$ (Maurer 1993).

Proof sketch

Achievability: that of $I(X \wedge Y) - I(X \wedge Z)$ suffices, since randomized encoding is allowed.

CR agreement: randomly select $N = \exp\{n(I(X \wedge Y) - \delta)\}$ sequences from the distribution P_X^n. Most outcomes of the random selection give a good codeword set for the DMC W_1: if Alice selects as channel input a RV U uniformly distributed on this set, denoted by $\mathbf{U}$, then U will be ε-CR for Alice and Bob.

Privacy amplification: One verifies via the Secrecy Lemma the existence of $\kappa : \mathbf{U} \to \{1, \ldots, k\}$ with

$$\frac{1}{n} \log k = I(X \wedge Y) - I(X \wedge Z) - 2\delta, \quad S(\kappa(U)|Z^n) < \varepsilon.$$

This argument gives that one can have even $\varepsilon = \varepsilon_n \to 0$ exponentially.

Converse: Suppose $M \to X^n \to Y^n Z^n$, $P_{Y^n|X^n} = W_1^n$, $P_{Z^n|X^n} = W_2^n$, M is ε-recoverable from Y^n and $S(M|Z^n) = \log|\mathbf{M}| - H(M|Z^n) < n\varepsilon$. The latter and $S(M|Y^n) \leq \varepsilon \log|\mathbf{M}| + 1$ (Fano's inequality) imply

$$(1-\varepsilon)\frac{1}{n}\log|\mathbf{M}| - \frac{1}{n} - \varepsilon$$
$$< \frac{1}{n}[H(M|Z^n) - H(M|Y^n)] = \frac{1}{n}[I(M \wedge Y^n) - I(M \wedge Z^n)].$$

Key identity: For arbitrary RVs U, V and sequences of RVs Y^n, Z^n

$$I(V \wedge Y^n|U) - I(V \wedge Z^n|U) = n[I(\tilde{V} \wedge Y_J|\tilde{U}) - I(\tilde{V} \wedge Z_J|\tilde{U})],$$

where $\tilde{U} = JUY_1 \dots Y_{J-1}Z_{J+1} \dots Z_n$, $\tilde{V} = \tilde{U}V$, and J is a RV uniform on $\{1, \dots, n\}$, independent of (U, V, Y^n, Z^n).

Apply this identity with $U = \text{const}$, $V = M \Rightarrow$ last bound equals

$$I(\tilde{V} \wedge Y_J|\tilde{U}) - I(\tilde{V} \wedge Z_J|\tilde{U}),$$

where $\tilde{U} = JUY_1 \dots Y_{J-1}Z_{J+1} \dots Z_n$, $\tilde{V} = \tilde{U}M$.

Renaming $\tilde{U}, \tilde{V}, X_J, Y_J, Z_J$ to U, V, X, Y, Z, these RVs satisfy the Markov conditions $U \to V \to X \to YZ$, and $P_{Y|X} = W_1$, $P_{Z|X} = W_2$.

It follows that if R is an achievable SK rate then

$$R \leq \sup[I(V \wedge Y|U) - I(V \wedge Z|U)]$$

for RVs as above. This supremum is equal to $\sup[I(V \wedge Y) - I(V \wedge Z)]$.

Standard arguments show the last supremum does not change if $|\mathbf{V}| \leq |\mathbf{X}|$ is required.

Broadcast channel with confidential messages

Two DCMs $W_1 : \mathbf{X} \to \mathbf{Y}$, $W_2 : \mathbf{X} \to \mathbf{Z}$.
Three messages M_0, M_1, M_s uniform on $\mathbf{M}_0, \mathbf{M}_1, \mathbf{M}_s$

$$(M_0, M_1, M_s) \to \text{ channel input } X^n \to \text{ output } Y^n, Z^n.$$

Requirements: (M_0, M_1, M_s) ε-recoverable from Y^n,
M_0 ε-recoverable from Z^n, $I(M_s \wedge Z^n) < \varepsilon$.

Goal: find all rate triples (R_0, R_1, R_s) achievable as $n \to \infty$.

Theorem (Csiszár-Körner 1978)

Necessary and sufficient is the existence of RVs
$UV \to X \to YZ$ *with* $P_{Y|X} = W_1$, $P_{Z|X} = W_2$ *such that*

$$R_0 \le \min[I(U \wedge Y), I(U \wedge Z)], \quad R_s \le I(V \wedge Y|U) - I(V \wedge Z|U),$$
$$R_0 + R_1 + R_s \le I(V \wedge Y|U) + \min[I(U \wedge Y), I(U \wedge Z)].$$

Here $|\mathbf{U}| \le |\mathbf{X}| + 3$, $|\mathbf{V}| \le |\mathbf{X}| + 1$ *may be assumed.*

Proof idea: In (asymmetric) broadcast channel without secrecy, random messages $M_0, \tilde{M}_1$ from sets $\mathbf{M}_0, \tilde{\mathbf{M}}_1$ are encoded into channel input X^n; both M_0 and $\tilde{M}_1$ have to be ε-recoverable from Y^n, and M_0 also from Z^n.

A random construction called superposition coding yields a good broadcast channel code. Then Secrecy Lemma implies existence of $\kappa : \tilde{\mathbf{M}}_1 \to \mathbf{M}_s$ with $\mathbf{M}_s$ of right size and $S(\kappa(\tilde{M}_1)|Z^n) < \varepsilon$.

The message $\tilde{M}_1$ splits, in effect, into (M_1, M_s) where $M_s = \kappa(\tilde{M}_1)$ is secret from Z^n.

Converse: Key identity plays crucial role.

Further generalizations of the model are known; single-letter solution is achievable for the Cognitive Interference Channel (Liang et al. 2009).

Multiparty models of generating SK

Source model. Given i.i.d. repetitions of an m-tuple of RVs $X_{\mathbf{M}} = \{X_i : i \in \mathbf{M}\}$, where $\mathbf{M} = \{1, \ldots, m\}$, the ith party observes the repetitions $X_i^n = X_{i1} \ldots X_{in}$ of X_i. There may be a set $\mathbf{D} \subset \mathbf{M}$ of compromised parties. All parties including the compromised ones cooperate, via communication over a noiseless public channel, to generate SK for a specified set of parties $\mathbf{A} \subset \overline{\mathbf{D}}$. The eavesdropper knows $X_{\mathbf{D}}^n = \{X_i : i \in \mathbf{D}\}$ and all communication, but (unlike previously) she does not have information unavailable to any of the m parties.

Permissible protocols: similar to the case of $m = 2$, but each compromised party $i \in \mathbf{D}$ immediately reveals X_i^n.

SK capacity: defined as in the case $m = 2$, is denoted by $C_{SK}(\mathbf{A}|\mathbf{D})$.

Step 1 of generating SK is agreement on CR; conveniently, let this CR be $X^n_{\mathcal{M}}$.

The total communication F of the non-compromised parties is communication for ε-ominiscience (for the set $\mathbf{A}$) if for each $i \in \mathbf{A}$ the whole $X^n_{\mathbf{M}}$ is ε-recoverable from X^n_i and $(F, X^n_{\mathbf{D}})$.

Omniscience rate $R_{OS}(\mathbf{A}|\mathbf{D})$: the smallest R such that for each $\varepsilon > 0$, $\delta > 0$ and sufficiently large n communication for ε-omniscience is possible with total rate $< R + \delta$.

Theorem (Csiszár-Narayan 2004)

$C_{SK}(\mathbf{A}|\mathbf{D}) = H(X_{\mathbf{M}}|X_{\mathbf{D}}) - R_{OS}(\mathbf{A}|\mathbf{D})$.

Achievability is immediate from the Corollary of the Secrecy Lemma.

Here, $R_{OS}(\mathbf{A}|\mathbf{D})$ equals the minimum of $\sum_{i\in\overline{\mathbf{D}}} R_i$ for vectors $\{R_i,\ i \in \overline{\mathbf{D}}\}$ satisfying the constraints

$$\sum_{i\in\mathbf{B}} R_i \geq H(X_{\mathbf{B}}|X_{\overline{\mathbf{B}}}), \quad \mathbf{B} \subset \overline{\mathbf{D}}, \mathbf{B} \not\supset \mathbf{A}.$$

(For $\mathbf{D} = \emptyset$, $\mathbf{A} = \mathbf{M}$, see Wyner-Wolf-Willems 2002.)

Channel model: Given a DMC $W : \mathbf{X}_1 \to \mathbf{X}_2 \times \cdots \times \mathbf{X}_m$, Party 1 controls the inputs, Party i observes output i ($i = 2, \ldots, m$). In other aspects, similar to the source model.

Theorem (Csiszár-Narayan 2008)

The SK capacity of a channel model is the maximum of the SK capacities of the source models emulated by taking i.i.d. channel inputs.

Achievability is obvious, converse is hard.

Additional results of Csiszár-Narayan 2008

If the set **D** of compromised parties contains $i = 1$ (DMC input), the SK capacity can be attained using a deterministic input sequence.

If the set **A** of SK-seeking parties contains $i = 1$ (DMC input), SK capacity can be attained by transmission: Party 1 generates the SK and transmits it over the DMC, not relying on any public messages from the outputs; in general, the receivers do need public communication to recover the transmitted SK.