Channel Upgrading for Semantically-Secure Encryption on Wiretap Channels

Ido Tal
Technion

Alexander Vardy
UCSD
The wiretap channel

Alice, Bob, and Eve

- **Alice** sends a message $U$ of $k$ bits to Bob through the encoder.
- **Main Channel** $W_{Bob}$ transmits $X$ of $n$ bits.
- **Random bits** $r$ are sent to Eve through the wiretap channel $W_{Eve}$.
- **Bob** decodes the message to obtain $\hat{U}$.
- **Eve** observes $Z$.

Wiretap channel essentials

- **Reliability**: $\lim_{n \to \infty} \Pr\{\hat{U} \neq U\} = 0$
- **Security**: $\lim_{n \to \infty} \frac{I(U; Z)}{n} = 0$
- **Random bits**: In order to achieve the above, Alice sends and Bob receives $r$ random bits, $r/n = I(W_{Eve})$. 
Semantic security

Information theoretic security, revisited

- **Assumption**: input $U$ is uniform.
- **Assumption**: figure of merit is mutual information, $I(U; Z)/n$.

Semantic security

We achieve $\sigma$ bits of semantic security if:

- For all distributions on the message set of Alice
- For all functions $f$ of the message
- For all strategies Eve might employ
- The probability of Eve guessing the value of $f$ correctly increases by no more than $2^{-\sigma}$ between the case in which Eve does not have access to the output of $W$ and the case that she does.

That is, having access to $W$ hardly helps Eve, for sufficiently large $\sigma$. 
**Notation**

The channel model

- Denote $W = W_{Eve}$.
- Let $W : \mathcal{X} \rightarrow \mathcal{Y}$ be a memoryless channel.
- Finite input alphabet $\mathcal{X}$
- Finite output alphabet $\mathcal{Y}$
- The channel $W$ is symmetric:
  - The output alphabet $\mathcal{Y}$ can be partitioned into $\mathcal{Y}_1, \mathcal{Y}_2, \ldots, \mathcal{Y}_T$.
  - Let $A_t = [W(y|x)]_{x \in \mathcal{X}, y \in \mathcal{Y}_t}$.
  - Each row (column) of $A_t$ is a permutation of the first row (column).
The BT scheme

The function $\Psi$

$$\Psi(W) \overset{\text{def}}{=} \log_2 |\mathcal{Y}| + \sum_{y \in \mathcal{Y}} W(y|0) \log_2 W(y|0),$$

$$= \log_2 |\mathcal{Y}| - H(\mathcal{Y}|X).$$

Theorem (The BT scheme)

Let $W : \mathcal{X} \rightarrow \mathcal{Y}$ be the SDMC from Alice to Eve. Then, the BT scheme achieves at least $\sigma$ bits of semantic security with a codeword length of $n$ and $r$ random bits, provided that

$$r = 2(\sigma + 1) + \sqrt{n \log_2 (|\mathcal{Y}| + 3)} \sqrt{2(\sigma + 3) + n \cdot \Psi(W)}.$$

The function $\Psi$

**Asymptotics**

\[
r = 2(\sigma + 1) + \sqrt{n \log_2(|Y| + 3)} \sqrt{2(\sigma + 3)} + n \cdot \Psi(W).
\]

Thus, the asymptotic number of random bits we need to transmit is

\[
\lim_{n \to \infty} \frac{r}{n} = \Psi(W).
\]

**$\Psi$ versus $I$**

\[
\Psi(W) \overset{\text{def}}{=} \log_2 |Y| + \sum_{y \in Y} W(y|0) \log_2 W(y|0),
\]

\[
= \log_2 |Y| - H(Y|X) \geq H(Y) - H(Y|X) = I(W)
\]

How can we “make” $\Psi(W)$ close to $I(W)$?
Equivalent channels

Degraded channel

A DMC $W : \mathcal{X} \to \mathcal{Y}$ is (stochastically) degraded with respect to a DMC $Q : \mathcal{X} \to \mathcal{Z}$, denoted $W \preceq Q$, if there exists an intermediate channel $P : \mathcal{Z} \to \mathcal{Y}$ such that

$$W(y|x) = \sum_{z \in \mathcal{Z}} Q(z|x) \cdot P(y|z).$$

If $W \preceq Q$ and $Q \preceq W$, then $W$ and $Q$ are equivalent, $W \equiv Q$. 

Equivalent channel
Letter Splitting

Splitting function

- Let an SDMC $W : \mathcal{X} \to \mathcal{Y}$ be given.
- Denote the corresponding partition as $\mathcal{Y}_1, \mathcal{Y}_2, \ldots, \mathcal{Y}_T$.
- A function $s : \mathcal{Y} \to \mathbb{N}$ is an output letter split of $W$ if
  - $s(y) = s(y')$ for all $1 \leq t \leq T$ and all $y, y' \in \mathcal{Y}_t$.
  - By abuse of notation, define $s(\mathcal{Y}_t)$.

Resulting channel

Applying $s$ to $W$ gives $Q : \mathcal{X} \to \mathcal{Z}$

- **Output alphabet:** $\mathcal{Z} = \bigcup_{y \in \mathcal{Y}} \{y_1, y_2, \ldots, y_s \mid s = s(y)\}$.
- **Transition probabilities:** $Q(y_i|x) = W(y|x)/s(y)$
- Namely, each letter $y$ is duplicated $s(y)$ times. The conditional probability of receiving each copy is simply $1/s(y)$ times the original probability in $W$. 
Properties of $Q$

- Since $W$ is symmetric, so is $Q$.
- $W \equiv Q$.

Lemma

For a positive integer $M \geq 1$, define

$$s(y) = \lceil M \cdot W(y) \rceil, \quad \text{where} \quad W(y) = \frac{1}{|\mathcal{X}|} \sum_{x \in \mathcal{X}} W(y|x).$$

Let $Q : \mathcal{X} \to \mathcal{Z}$ be the resulting channel. Then,

$$\Psi(Q) - I(W) = \Psi(Q) - I(Q) \leq \log_2 \left( 1 + \frac{|\mathcal{Y}|}{M} \right),$$

and $|\mathcal{Z}| \leq M + |\mathcal{Y}|$. 
The number of random bits needed to achieve semantic security is at most

\[
r = 2(\sigma + 1) + \sqrt{n} \log_2 (M + |\mathcal{Y}| + 3) \sqrt{2(\sigma + 3)} + n \cdot \left( I(W) + \log_2 \left( 1 + \frac{|\mathcal{Y}|}{M} \right) \right).
\]

Consequences

- Setting, say, \( M = n \) and taking \( n \to \infty \) gives us

\[
\lim_{n \to \infty} \frac{r}{n} = I(W).
\]

- What about the finite \( M \) and \( n \) case?
Algorithm A: Greedy algorithm to find optimal splitting function

**input** : Channel $W : \mathcal{X} \to \mathcal{Y}$, a partition $\mathcal{Y}_1, \mathcal{Y}_2, \ldots, \mathcal{Y}_T$ where each subset is of size $\mu$, a positive integer $M$ which is a multiple of $\mu$

**output**: A letter-splitting function $s$ such that $\sum_{y \in \mathcal{Y}} s(y) = M$ and $\Psi(Q)$ is minimal

// Initialization
$s(\mathcal{Y}_1) = s(\mathcal{Y}_2) = \cdots = s(\mathcal{Y}_T) = 1$ ;

// Main loop
for $i = 1, 2, \ldots, \frac{M}{\mu} - T$ do

\[
t = \arg \max_{1 \leq t \leq T} \sum_{y \in \mathcal{Y}_i} W(y) \log_2 \left( \frac{s(\mathcal{Y}_i) + 1}{s(\mathcal{Y}_i)} \right) ;
\]

$s(\mathcal{Y}_t) = s(\mathcal{Y}_t) + 1$;

return $s$;
Greedy algorithm

Theorem

Given a valid input to Algorithm A, the output is a valid letter-splitting function \( s \), such that \( \sum_{y \in Y} s(y) = M \) and the resulting channel \( Q \) is such that \( \Psi(Q) \) is minimized.

Proof

- Proooving \( \sum_{y \in Y} s(y) = M \):
  - After the initialization step, \( \sum_{y \in Y} s(y) = \mu \cdot T \).
  - Each iteration increments the sum by \( \mu \).
  - So, in the end, \( \sum_{y \in Y} s(y) = M \).
- Proooving optimality:
  - Since \( Q \equiv W \), we have \( I(Q) = I(W) \).
  - Minimizing \( \Psi(Q) \) is equivalent to maximizing
    \[
    I(Q) - \Psi(Q) = \sum_{y \in Y} -W(y) \log_2 \left( \frac{W(y)}{s(y)} \right) - \log_2 M.
    \]
Greedy algorithm

Proof, continued

- Clearing away constant terms, maximize

\[ \sum_{y \in Y} W(y) \log_2 s(y). \]

- We now recast the optimization problem. Define the set

\[ A = \bigcup_{y \in Y} \bigcup_{i=1}^{M/\mu - T} \left\{ \delta(y, i) = W(y) \log_2 \left( \frac{i+1}{i} \right) \right\}. \]

- Finding the optimal \( s(y) \) is equivalent to choosing \( M/\mu - T \) numbers from the set \( A \) such that
  - Their sum is maximal, and
  - if \( \delta(y, i) \) was picked and \( i > 1 \), then \( \delta(y, i - 1) \) must be picked as well.

- The last constraint is redundant. The proof follows.
What would we do if the output alphabet of $W$ is infinite?

To begin with, in this case, $\Psi$ is not even defined.

**Solution:** Replace $W$ by a channel $Q$ which is **upgraded** and has a finite output alphabet.

A channel $Q$ is **upgraded** with respect to $W$ if $W \preceq Q$.

A method to upgrade $W$ to $Q$ was previously presented by the authors in “How to Construct Polar Codes”.

The method we now show is **better**, with respect to $\Psi$. 
Notation

Assumptions

- Assume the input alphabet is binary, and denote $\mathcal{X} = \{1, -1\}$.
- Let the output alphabet be the reals, $\mathcal{Y} = \mathbb{R}$.
- Symmetry: $f(y|1) = f(-y| -1)$.
- Positive value more likely when $x = 1$
  \[ f(y|1) \geq f(y| -1), \quad y \geq 0. \]
- Likelihood increasing in $y$:
  \[ \frac{f(y_1|1)}{f(y_1| -1)} \leq \frac{f(y_2|1)}{f(y_2| -1)}, \quad -\infty < y_1 < y_2 < \infty. \]
The channel $Q$

Partitioning $\mathbb{R}$

- Let the channel $W$ and a positive integer $M$ be given.
- **Initialization**: Define $y_0 = 0$.
- **Recursively** define, for $1 \leq i < M$ the number $y_i$ as such that
  \[
  \int_{-y_i}^{-y_{i-1}} f(y|1) \, dy + \int_{y_{i-1}}^{y_i} f(y|1) \, dy = \frac{1}{M}.
  \]
- Lastly, “define” $y_M = \infty$.
- For $1 \leq i \leq M$, the regions
  \[
  A_i = \{ y : -y_i < y \leq -y_{i-1} \} \cup \{ y : y_{i-1} \leq y < y_i \}
  \]
  form a partition of $\mathbb{R}$, which is equiprobable with respect to $f(\cdot|1)$ and $f(\cdot| -1)$
  \[
  f(A_i|1) = f(A_i| -1) = \frac{1}{M}.
  \]
The channel $Q$

The likelihood ratios $\lambda_i$

- Recall the partition

$$A_i = \{y : -y_i < y \leq -y_{i-1}\} \cup \{y : y_{i-1} \leq y < y_i\},$$

which is equiprobable

$$f(A_i|1) = f(A_i|-1) = \frac{1}{M}.$$

- Define the likelihood ratios

$$\lambda_i = \frac{f(y_i|1)}{f(y_i|-1)}.$$

- By our previous assumptions,

$$1 \leq \lambda_{i-1} = \inf_{y \in B_i} \frac{f(y|1)}{f(y|-1)} \leq \sup_{y \in B_i} \frac{f(y|1)}{f(y|-1)} \leq \lambda_i.$$
The channel $Q$

- The channel $Q : \mathcal{X} \to \mathcal{Z}$ is defined as follows.
- **Input alphabet:** $\mathcal{X} = \{-1, 1\}$.
- **Output alphabet:** $\mathcal{Z} = \{z_1, \bar{z}_1, z_2, \bar{z}_2, \ldots, z_M, \bar{z}_M\}$.
- **Conditional probability:**

  $Q(z|1) = \begin{cases} 
  \frac{\lambda_i}{M(\lambda_i+1)} & \text{if } z = z_i \text{ and } \lambda_i \neq \infty, \\
  \frac{1}{M(\lambda_i+1)} & \text{if } z = \bar{z}_i \text{ and } \lambda_i \neq \infty, \\
  \frac{1}{M} & \text{if } z = z_i \text{ and } \lambda_i = \infty, \\
  0 & \text{if } z = \bar{z}_i \text{ and } \lambda_i = \infty, 
  \end{cases}$

and

$Q(z_i| -1) = Q(\bar{z}_i|1), \quad Q(\bar{z}_i| -1) = Q(z_i|1)$.

- For $1 \leq i \leq M$, the **likelihood ratio** of $z_i$ is $Q(z_i|1) / Q(z_i| -1) = \lambda_i$. 

Properties of $Q$

- **Finite output alphabet**: $|\mathcal{Z}| = 2M$.
- **Optimal $\Psi$**: $\Psi(Q) = I(Q)$, since $Q(z_i) = Q(\bar{z}_i) = \frac{1}{2M}$.
- $Q$ is upgraded with respect to $W$, $W \preceq Q$.
- **Key question**: What is $I(Q) - I(W)$?

The channel $Q'$

- Define $Q' : \mathcal{X} \to \mathcal{Z}$ as a “shifted version” of $Q$.

$$Q'(z|1) = \begin{cases} \frac{\lambda_{i-1}}{M(\lambda_{i-1}+1)} & \text{if } z = z_i, \\ \frac{1}{M(\lambda_{i-1}+1)} & \text{if } z = \bar{z}_i, \end{cases}$$

and

$$Q'(z_i|1) = Q'(\bar{z}_i|1), \quad Q'(\bar{z}_i|1) = Q'(z_i|1).$$

- $Q'$ is degraded with respect to $W$, $Q' \preceq W$.
- To sum up,

$$Q' \preceq W \preceq Q.$$
Theorem

Let $W : \mathcal{X} \to \mathcal{Y}$ be a continuous channel as defined above. For a given integer $M$, let $Q : \mathcal{X} \to \mathcal{Z}$ be the upgraded channel described previously. Then, $|\mathcal{Z}| = 2M$ and

$$\Psi(Q) - I(W) \leq \frac{1}{M}.$$

Proof.

We know that

$$\Psi(Q) = I(Q),$$

and that

$$I(Q') \leq I(W) \leq I(Q).$$

Thus, it suffices to prove that

$$I(Q') - I(Q) \leq \frac{1}{M}.$$

Because $Q'$ is a “shifted version” of $Q$, the above difference telescopes to $1/M$. 

□