Wiretap Channels with Random States

Paul Cuff (Princeton University), Ziv Golfeld, Haim Permuter
The Setting: 1975

- Wyner publishes five papers
- Two of interest in this talk
  - Wiretap Channel
  - Common Information
Wiretap Channel

Foundation of physical-layer security
Wiretap Channel

Secrecy Capacity:
- Reliable communication
- $Z^n$ contains no information about $M$
Solution

\[ C_s = \max_{P_{X,U}} I(U; Y) - I(U; Z) \]
Solutions

1975: Wyner introduced the problem and gave solution for degraded channels (\(U=X\) is sufficient)

1978: Csiszár and Körner gave solution for all channels
Encoding

- Random Codebook
- Pad with random garbage bits

Message: 01001011010111100100
Padding: 011001010

Transmitted together in one block
Given $M$, can decode $M'$. 

$$R' \approx I(X;Z)$$
Wyner’s security argument

\[
I(M, M'; Z^n) = I(M; Z^n) + I(M'; Z^n | M)
\]

\[
I(X^n; Z^n) \approx nI(X; Z)
\]

\[
H(M') = nR'
\]

Decodable if \( R' < I(X; Z) \)
Encoding Concept

Given $M$, distribution $\approx \prod P_{Y,Z|X}$

$R' \approx I(X;Z) \implies \text{Given } M, \text{ distribution } \approx \prod P_z$
Soft Covering

- Theorem 6.3 of Wyner's C.I. paper:

  \[ M (nR \text{ bits}) \rightarrow_{\text{Code}} u^n(M) \rightarrow_{\text{Memoryless \ QV|U}} V^n \]

Randomly select a codeword
Pass through a memoryless channel
Does induced output distribution match desired?
Output Distribution

Desired output distribution:

\[ Q_V(v) = \sum_u Q_{V|U}(v|u) Q_U(u) \]

Induced output distribution:

\[ P_{V^n|C} = 2^{-nR} \sum_{u^n(m) \in C} Q_{V^n|U^n=u^n(m)} \]

\[ Q_{V^n} = \prod Q_V \]

\[ Q_{U^n} = \prod Q_U \]

\[ Q_{V^n|U^n} = \prod Q_{V|U} \]
Output Distribution

\( P_{V|C} \) \( Q_V \)

u(1) u(5) u(4) u(2) u(3)
Gaussian Example
Gaussian Example
Soft Covering Lemma

- **Codebook size:** If $R > I(U;V)$

- **Codebook generation:** $U^n(m) \sim_{U \text{i.i.d.}}$

- **Success:** $P_{V^n|c} \approx Q_{V^n}$
Covering and Packing

Covering (compression)

\[ \mathcal{U}_n \rightarrow \mathcal{V}_n \]

Packing (transmission)

\[ \mathcal{U}_n \rightarrow \mathcal{V}_n \]
Covering

Hard covering:

\[ \bigcup_{u^n(m)} T_e(u^n(m)) \approx V^n \quad \text{in probability} \]

Soft covering:

\[ 2^{-nR} \sum_{u^n(m)} Q V^n | U^n = u^n(m) \approx Q V^n \]
Encoding Concept

\[ M \text{ (nR bits)} \xrightarrow{\text{Enc}} X^n \xrightarrow{\text{Dec}} M' \text{ (nR' bits)} \]

Given \( M \), distribution \( \approx \prod P_{Y,Z|X} \)

\[ R' \approx I(X;Z) \Rightarrow \text{Given } M, \quad \text{distribution } \approx \Pi P_z \]
Gelfand-Pinsker (state known to encoder)

- Reliable communication

Capacity:

M (nR bits) → Enc → X^n → Memoryless P_{Y|X,S} → Y^n → Dec

S^n i.i.d. from P_s
Solution (1980)

\[
C = \max_{P_{X,U \mid S}} I(U; Y) - I(U; S)
\]
Encoding

- Random Codebook
- Pad with \textit{skillfully chosen} bits

Transmitted together in one block
Similarities

- Virtually the same
- Same encoding
- Same converse (except, iid $S^n$ allows a skipped step)
- Same problem statement:
  - Wiretap: $M$ independent of $Z^n$
  - Gelfand-Pinsker: $M$ independent of $S^n$
Wiretap Channel with State

- Reliable communication
- $Z^n$ contains no information about $M$

Secrecy Capacity:
Same Encoding

\[ C_s \geq \max_{P_{X,U|S}} \ I(U;Y) - \max \left\{ I(U;Z), \ I(U;S) \right\} \]

Message

010010110101111100100

Padding

011001010

Transmitted together in one block

Chen and Han Vinck, 2006
Extract Key

Assume $S$ is known to the intended receiver as well:

$$C_s \geq \max_{P_{X,U|S}} \min \left\{ I(U;Y|S), \quad \begin{cases} I(U;Y|S), \\ H(S|Z,U) \end{cases} \right\}$$

Better in some cases!

Chia and El Gamal, 2012

Note: They consider causal state information. This region is adapted to take advantage of non-causal state information.
Combined

Assume $S$ is known to the intended receiver as well:

$$C_s \geq \max_{P_{X,U|S}} \min \left\{ \begin{array}{c} I(U;Y|S), \\ H(S|Z,U) + [I(U;Y,S) - I(U;Z)]_+ \end{array} \right\}$$

Chia and El Gamal, 2012

Note: They consider causal state information.
This region is adapted to take advantage of non-causal state information.
Our Scheme

Superposition code

$U^n$ index is padding only
$V^n$ index is message and padding

All secrecy comes from $V$
$U^n$ is decoded by the eavesdropper
Our Scheme

\[
C_s \geq \max_{P_{X,U,V|S} : I(U;Y) \geq I(U;S)} \min \left\{ I(U,V;Y) - I(U,V;S), \right. \\
\left. I(V;Y|U) - I(V;Z|U) \right\}
\]

Can mimic Chia and El Gamal’s key extraction by setting \( V=S \)

Beats previous regions
Other Related Work

- Prabhakaran, Eswaran, and Ramchandran, 2012:
  - Same superposition code but require $U-V-(S,X)$ and $U \perp S$.

- Bassi, Bunin, Piantanida, and Shamai, 2016 (several papers):
  - Key generation and secure communication
  - Sources independent of channel
  - Generalized feedback
Simple Special case

- Unlimited public noise-free channel
- "Key Capacity with one-way communication"

\[ C_s = \max_{P_{U,V|S_x}} I(V; S_y|U) - I(V; S_z|U) \]

Achieved by our scheme
Apparently not by any others
Analysis Trick

- **Likelihood encoder to choose padding**

Soft Covering

**Distribution 1:**
- Choose $U^n$ codeword uniformly at random
- Generate $S^n$ memorylessly from $U^n$

**Distribution 2:**
- Use the above conditional distribution of $U^n$ given $V^n$
  (this is the likelihood encoder)
- Let $S^n$ be iid
Observation

- Lower layer of superposition code intended to be decoded by eavesdropper (i.e. "decoy")
- Villard-Piantanida secure source coding
- Key agreement with one way communication
Differential Privacy as a Mutual Information Constraint

Paul Cuff and Lanqing Yu
Database Privacy

- Let $X_1, X_2, \ldots, X_n$ be entries in a database.
  - E.g. $X_i$ is personal information about person $i$.
- Let $Y$ be the response to a query.
  - We can denote $Y_q$ to indicate that it depends on the query.
- The job of the information provider is to answer queries and protect individual privacy.
Differential Privacy

\( \varepsilon \)-DP:

- Let \( x \) and \( x' \) differ in only one entry (i.e., \( x_i = x'_i \) for all but one \( i \))

- \( p(y|x) \leq e^{\varepsilon} p(y|x') \)

Why \( x \) and \( x' \) differ in only one spot?

- Convince someone to put their data in your database

Why multiplicative constraint?

- Posterior update is small

Dwork, 2006
A Common Technique

Add Laplacean noise
Weaker DP

$(\varepsilon, \delta)$-DP:

$$P(Y \in A | x) \leq e^\varepsilon P(Y \in A | x') + \delta$$

Additive Gaussian noise often provides privacy

Dwork, 2006
Mutual Information
Differential Privacy

$\varepsilon$-MI-DP:

$$\max_{i, P_{X^n}} I(X_i; Y | X^{i-1}, X_{i+1}^n) < \varepsilon$$
Claim

$$\varepsilon$$-DP > MI-DP > $$(\varepsilon, \delta)$$-DP

Furthermore, if input or output alphabet is finite,

$$\text{MI-DP} = (\varepsilon, \delta)$$-DP

similar to semantic security proof, Bellare, Tessaro, Vardy, 2012
Privacy Ordering

\[ \alpha\text{-DP} > \beta\text{-DP} \text{ if for all } \beta > 0 \text{ there exists } \alpha \text{ such that } \alpha\text{-DP} \implies \beta\text{-DP}. \]
Subadditivity of DP

Multiple queries:

If $k$ queries $Y_{q1}, Y_{q2}, ..., Y_{qk}$ each have differential privacy $\varepsilon$ and are conditionally independent, the combined they have $k\varepsilon$ privacy.

Simple MI-DP Proof:

$$I(X;Y_1,Y_2) = I(X;Y_1) + I(X;Y_2|Y_1)$$

$$\leq I(X;Y_1) + I(X;Y_2)$$

For clarity, conditioned database variables are omitted.
Common complaint

- Differentially privacy doesn’t not mean that you can’t learn about $X_i$.
- Consider a database with correlated entries.

Simple MI-DP Explanation:

$$I(X_i; Y) \neq I(X_i; Y | X^{i-1}, X^{i+1})$$
Precise Bounds
$(\epsilon, \delta)$-Closeness

$P \approx Q$

if

$P(A) \leq e^\epsilon Q(A) + \delta, \quad \forall A \in \mathcal{F},$

$Q(A) \leq e^\epsilon P(A) + \delta, \quad \forall A \in \mathcal{F}.$
Special Cases

\[ P^{(\varepsilon,0)} \approx Q \quad \iff \quad \left| \ln \left( \frac{dP}{dQ} (a) \right) \right| \leq \varepsilon \quad \forall a \in \Omega. \]

\[ P^{(0,\delta)} \approx Q \quad \iff \quad \|P - Q\|_{TV} \leq \delta. \]
Relation to Detection Theory

\[ P^{(\epsilon, \delta)} \approx Q \]

Optimal tradeoff within this region

\( \epsilon \) determines slope

\( \delta \) determines intercept

Kairouz, Oh, Viswanath, 2015
Tightest \((\varepsilon, \delta)\) Conversion

\[
P^{(\varepsilon, \delta)} \approx Q \iff P^{(\varepsilon', \delta')} \approx Q.
\]

for

\[
\varepsilon' \leq \varepsilon
\]

\[
\delta' = 1 - \frac{(e^{\varepsilon'} + 1)(1 - \delta)}{e^\varepsilon + 1}
\]
Simple Claim

\[ P^{(\epsilon, 0)} \approx Q \implies D(P\|Q) \leq \epsilon \text{ nats,} \]
\[ D(Q\|P) \leq \epsilon \text{ nats.} \]

Tightest Claim

\[ P^{(\epsilon, 0)} \approx Q \implies D(P\|Q) \leq \epsilon \frac{(e^\epsilon - 1)(1 - e^{-\epsilon})}{(e^\epsilon - 1) + (1 - e^{-\epsilon})} \text{ nats,} \]
\[ D(Q\|P) \leq \epsilon \frac{(e^\epsilon - 1)(1 - e^{-\epsilon})}{(e^\epsilon - 1) + (1 - e^{-\epsilon})} \text{ nats.} \]
Pinsker

\[ D(P \parallel Q) \leq \epsilon \text{ nats} \implies P \left(0, \sqrt{\frac{\epsilon}{2}}\right) \approx Q. \]
Relative entropy to Mutual Information

If

$$D \left( P_Y | X = x_1 \ || \ P_Y | X = x_2 \right) \leq \epsilon \quad \forall x_1, x_2 \in \mathcal{X}$$

then

$$I(X; Y) \leq \epsilon$$

Hint: Radius of information ball
Mutual Information to Total Variation

\[ I(X;Y) \leq \epsilon \implies \| P_{Y|X=x_1} - P_{Y|X=x_2} \|_{TV} \leq \delta' \]

\[ \delta' = 1 - 2h^{-1}(\ln 2 - \epsilon) \]

\[ \leq \sqrt{2\epsilon} \]

Tightest bound, achieved with binary channel
Finite Alphabet

\[ \| P_{Y|X=x_1} - P_{Y|X=x_2} \|_{TV} \leq \delta \quad \forall x_1, x_2 \in \mathcal{X} \quad \implies \quad I(X;Y) \leq \epsilon' \]

\[ \epsilon' = 2h(\delta) + 2\delta \ln \left( \min \left\{ |Y|, \max_i |X_i| + 1 \right\} \right) \]

Continuity of entropy

Continuity of conditional entropy

harder step

inspired by Alicki and Fannes, 2004
Observation

\[
\max_{P_{X^n}} I(X_i; Y|X_{i-1}^i, X_{i+1}^n) = \max_{\prod_{t=1}^n P_{X_t}} I(X_i; Y) \quad \forall i
\]

Either could be used for definition of MI-DP

Pointed out by Thomas Steinke