Fault-Tolerant Secret Key Generation

Himanshu Tyagi

University of Maryland, College Park

Joint work with:
Navin Kashyap†  Yogesh Sankarasubramaniam∗  Kapali Viswanathan∗

†Indian Institute of Sciences, Bangalore   * HP Labs, Bangalore
Set of nodes: $\mathcal{M} = \{1, \ldots, m\}$

- Observations of the $i$th node: $X_i^n = (X_{i1}, \ldots, X_{in})$
- Denote by $X_{\mathcal{M}t}$ the correlated rvs $(X_{1t}, \ldots, X_{mt})$
- $X_{\mathcal{M}1}, \ldots, X_{\mathcal{M}n}$ are finite, discrete valued, i.i.d. rvs with known probability distribution.
$r$-Rounds Adaptive Protocol

Available Nodes: $A_0 = \mathcal{M}$
**Formulation**

**An Upper Bound**

**Symmetric Observations**

**Exchangeability**

**PIN Model**

---

**$r$-Rounds Adaptive Protocol**

Nodes Remaining: $A_1 = \{1, 2, 3, 4, 5, 6, 7\}$

Communication in round $j$ depends on:

- local observations
- the communication in the previous rounds
**$r$-Rounds Adaptive Protocol**

Nodes Remaining: $A_2 = \{2, 3, 4, 6, 7\}$

Communication in round $j$ depends on:
- local observations
- the communication in the previous rounds.
$r$-Rounds Adaptive Protocol

Nodes Remaining: $A_{r-1} = \{2, 4, 6\}$

Communication in round $j$ depends on:
- local observations
- the communication in the previous rounds.
Nodes Remaining: \( A_{r-1} = \{2, 4, 6\} = A_r \)

Communication in round \( j \) depends on:
- local observations and the communication in the previous rounds.

Assumption: \( A_r = A_{r-1} \)
Communication in round $j$ depends on:
local observations and the communication in the previous rounds.

Assumption: $A_r = A_{r-1}$

The overall communication depends on $A_r = A_{r-1} \subseteq ... \subseteq A_1$
- $F$ denotes the overall communication.
$K$ constitutes a *secret key* if:

1. **Recoverability:** $\Pr(K_i = K, i \in A_r) \approx 1$

2. **Security:** $I(K \wedge F) \approx 0$

The rate of the SK: $\frac{1}{n} H(K)$
**Definition (Achievable \((r, t)\)-fault-tolerant SK rate)**

\( R \geq 0 \) is an achievable \((r, t)\)-fault-tolerant SK rate if there is an \(r\)-rounds adaptive protocol that generates an SK of rate greater than \( R \) whenever not more than \( t \) nodes drop out.
$K$ constitutes a **perfect secret key** if:

1. **Perfect Recoverability**: $\Pr(K_i = K, i \in A_r) = 1$

2. **Perfect Security**: $I(K \land F) = 0$

The rate of the SK: $\frac{1}{n} H(K)$
Definition (Achievable \((r, t)\)-fault-tolerant perfect SK rate)

\( R \geq 0 \) is an achievable \((r, t)\)-fault-tolerant perfect SK rate if there is an \(r\)-rounds adaptive protocol that generates a perfect SK of rate greater than \( R \) whenever not more than \( t \) nodes drop out.
Fault-Tolerant Secret Key Capacity

$(r, t)$-fault-tolerant SK capacity $C^{r,t}(\mathcal{M})$:
Supremum of all achievable $(r, t)$-fault-tolerant rates.

$(r, t)$-fault-tolerant perfect SK capacity $C^{r,t}_0(\mathcal{M})$:
Supremum of all achievable $(r, t)$-fault-tolerant perfect SK rates.

Lemma

For $r \geq 1$,

$$C^{1,t}_0(\mathcal{M}) \leq C^{r,t}(\mathcal{M}) \leq C^{r+1,t}(\mathcal{M}).$$
An Upper Bound on Fault-Tolerant SK Capacity

**Theorem (Csiszár-Narayan 2004)**

The secret key capacity (for $t=0$) is given by

$$C(\mathcal{M}) = H(X_{\mathcal{M}}) - \min (R_1 + R_2 + ... + R_m),$$

where the $\min$ is taken over $(R_1, ..., R_m)$ that satisfy:

$$\sum_{i \in B} R_i \geq H(X_B \mid X_{\mathcal{M}\setminus B}), \quad B \subset \mathcal{M}.$$

The $\min$ value above is the minimum rate of communication for omniscience.

**Lemma (Upper Bound on $C^{r,t}(\mathcal{M})$)**

$$C^{0,t}_0(\mathcal{M}) \leq C^{r,t}(\mathcal{M}) \leq C^{r+1,t}(\mathcal{M}) \leq \min_{A \subseteq \mathcal{M}, |A| \geq m-t} C(A), \quad r \geq 1.$$

**Proof Idea:** Consider the sequence of sets $A_1 = ... = A_{r-1} = A_r = A$. 

Monotonicity of SK Capacity

**Theorem (Chan-Zheng 2010)**

\[
C(\mathcal{M}) = \min_{\mathcal{P} = \{C_1, \ldots, C_k\}} \frac{1}{k} D \left( X_M \parallel X_{C_1} X_{C_2} \ldots X_{C_k} \right),
\]

where the minimization is over all partitions \(\mathcal{P}\) of \(\mathcal{M}\).

**Lemma (Monotonicity of \(C(\mathcal{M})\))**

\[
C(\mathcal{M}) \geq \min_{A \subseteq \mathcal{M}, |A| = m - 1} C(A).
\]

**Lemma (Upper Bound on \(C^{r,t}(\mathcal{M})\))**

\[
C_0^{1,t}(\mathcal{M}) \leq C^{r,t}(\mathcal{M}) \leq C^{r+1,t}(\mathcal{M}) \leq \min_{A \subseteq \mathcal{M}, |A| = m - t} C(A), \quad r \geq 1.
\]
Is this Upper Bound Tight??

Lemma (Upper Bound on $C^{r,t}(\mathcal{M})$)

$$C_0^{1,t}(\mathcal{M}) \leq C^{r,t}(\mathcal{M}) \leq C^{r+1,t}(\mathcal{M}) \leq \min_{A \subseteq \mathcal{M} \mid |A| = m-t} C(A), \quad r \geq 1.$$
Lemma (Upper Bound on $C^{r,t}(\mathcal{M})$)

$$C_{0}^{1,t}(\mathcal{M}) \leq C^{r,t}(\mathcal{M}) \leq C^{r+1,t}(\mathcal{M}) \leq \min_{A \subseteq \mathcal{M}, |A| = m-t} C(A), \quad r \geq 1.$$ 

Yes.

When the observations of the nodes are symmetric
Exchangeable Random Variables

\[ P_{X_1, \ldots, X_m} = P_{X_{\sigma(1)}, \ldots, X_{\sigma(m)}}, \text{ for all permutations } \sigma \text{ of } \{1, \ldots, m\} \]

For disjoint sets \( B_1, B_2 \): \( H(X_{B_1} | X_{B_2}) \) depends only on \(|B_1|, |B_2|\)

Define: \( g(i | j) = H(X_1, \ldots, X_i | X_{i+1}, \ldots, X_{i+j}) \)

Lemma (Minimum Rate of Communication for Omniscience)

For \( \alpha_m = \frac{g(m-1|1)}{m-1} \),

\( (\alpha_m, \ldots, \alpha_m) \) is an optimal rate-vector for omniscience, i.e., \( R_{CO} = m\alpha_m \).

Lemma

\( \alpha_m \) is nonincreasing in \( m \).

Proof: Uses properties \( g(i | j) \) inherited from \( H(\cdot) \).
2-rounds adaptive protocol:

1. Each node communicates using random mapping of rate $\alpha_m$.
   
   $A_1$ = set of nodes that communicate in round 1, $|A_1| = k$

2. Nodes in $A_1$ send further communication of rate $\alpha_k - \alpha_m$
   - if $A_2 \neq A_1$ the protocol fails.

Observation: Two random mappings of rates $R_1$ and $R_2$ can serve as a single random mapping of rate $R_1 + R_2$ in (multiterminal) Slepian-Wolf coding.

Performance of the protocol:

- Nodes in $A_2 = A_1$ recover $X_{A_1}^n$

- Rate of communication = $k\alpha_k$

- Nodes in $A_2$ generate SK of rate $C(A_2)$
For exchangeable rvs, for \( r \geq 2 \),

\[
C^{r,t}(\mathcal{M}) = \min_{A \subseteq \mathcal{M} \mid |A| = m-t} C(A) = g(m - t|0) - \frac{(m - t)g(m - t - 1|1)}{m - t - 1}.
\]
The Pairwise-Independent-Network Model

Graph \( G = (\mathcal{V}, \mathcal{E}) \)

\[
\begin{align*}
\text{Ye-Reznik 2007, Nitinawarat et.al. 2010}
\end{align*}
\]

\( B_{ij} \): unbiased bit corresponding to the edge \( e_{ij} \)

Random Variables \( \{ B_{ij} : i, j \in \mathcal{M} \} \) are mutually independent.

\( X_i = \{ B_{ij} \text{ corresponding to edges } e_{ij} \text{ incident on } i \} \)
The Pairwise-Independent-Network Model

Assumption: The graph $G$ is complete.

Symmetry: For $B_1 \cap B_2 = \emptyset$, $H(X_{B_1} | X_{B_2})$ depends only on $|B_1|, |B_2|$.

$$C^{1,t}_{0}(\mathcal{M}) \leq C^{2,t}_{0}(\mathcal{M}) = g(m-t|0) - \frac{(m-t)g(m-t-1|1)}{m-t-1} = \frac{m-t}{2}$$
Assume that $G$ is a $(t + 1)$-connected, spanning graph.

Noninteractive protocol to generate 1-bit of fault-tolerant SK:

\[ \{ B_{ij} \oplus B_{ij'} : e_{ij}, e_{ij'} \in \mathcal{E} \} \]

For $A \subseteq \mathcal{M}$ with $|A| \geq m - t$: let $e_A$ be an edge between nodes in $A$.

Claim: $H(B_{e_A} | (F_A, X_i)) = 0$ and $I(B_{e_A} \land F_A) = 0$, $i \in A$.

$B_{e_A}$ constitutes a 1-bit SK for $A$.
Assume that $G$ is a $(t + 1)$-connected, spanning graph.

Noninteractive protocol to generate 1-bit of fault-tolerant SK:

$$\{ B_{ij} \oplus B_{ij'} : e_{ij}, e_{ij'} \in \mathcal{E} \}$$

This noninteractive protocol generates 1-bit SK for each spanning tree.

Nitinawarat et.al. use the interactive protocol of Csiszár-Narayan.
Noninteractive protocol above gives 1-bit of SK for each spanning tree

- sufficiently many spanning trees must remain when nodes drop out

Consider \( n = 2 \): Any two nodes share 2 independent bits

- Can find a spanning tree packing such that:
  - any subset \( A \) contains \(|A|\) spanning trees

Thus, a subset of size \( \geq m - t \) can pack \( m - t \) spanning trees

Secret key rate attained: \( \frac{m-t}{2} \)
Optimal Fault-Tolerant SK Generation Protocol
Theorem

For the PIN model corresponding to a complete graph,

\[ C_{0,t}^r (\mathcal{M}) = C^{r,t} (\mathcal{M}) = \frac{m - t}{2}, \quad r \geq 2. \]
An Alternative Protocol

A protocol to generate $\left\lfloor \frac{m}{2} \right\rfloor - t$ bits of SK for $n = 1$:

First consider $m$ even.

Tree remains connected if a leaf node drops out.

- Fix a matching in $G$. 
An Alternative Protocol

A protocol to generate \( \left\lfloor \frac{m}{2} \right\rfloor - t \) bits of SK for \( n = 1 \):

First consider \( m \) even.

Tree remains connected if a leaf node drops out.

- Fix a matching in \( G \).
- There is a spanning tree corresponding to each edge in the matching.
Future Directions

- This work is a first step towards the larger goal of information-theoretic SK agreement for dynamic groups.

- Incorporate rejoining of terminals that drop out.

- What if the central switch has additional side information?