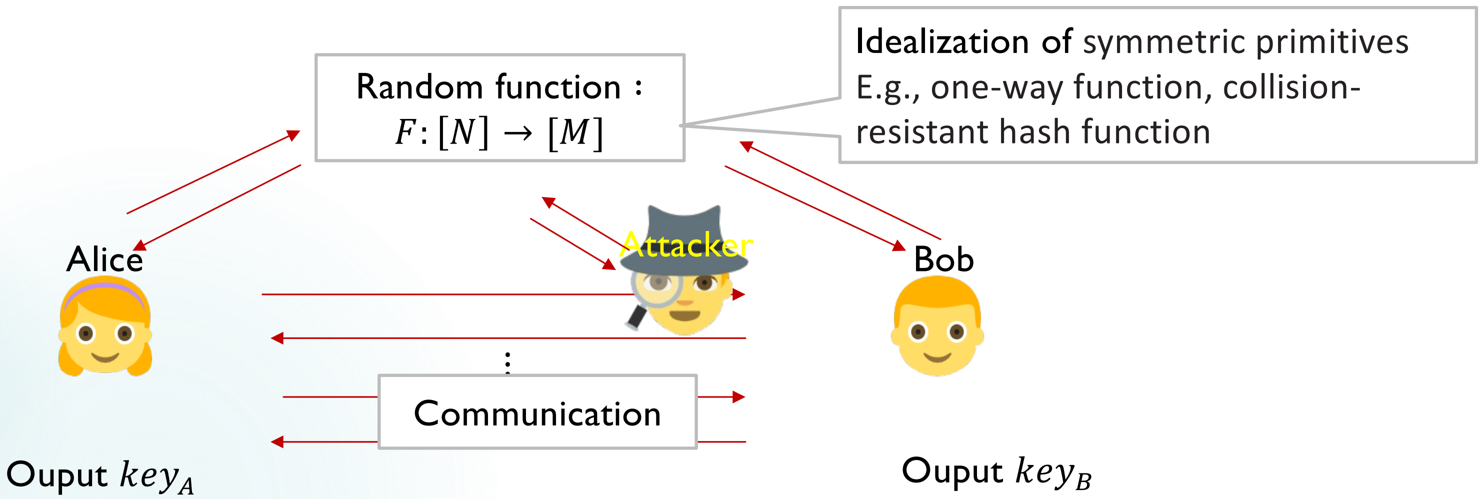
Communication Lower Bounds of Key-Agreement Protocols via Density Increment Arguments

Mi-Ying (Miryam) Huang Xinyu Mao **Guangxu Yang** Jiapeng Zhang

USC University of Southern California

Key-Agreement Protocols in the ROM



Correctness: $key_A = key_B$ (w.h.p.)

Security: any attacker sees the transcript and makes a few queries cannot guess key_A.

Upper Bounds: Merkle Puzzle [Merkle 78]

Random function
$F: [N] \to [M]$

Sample $x_1, \dots, x_\ell \leftarrow [N]$
Query the oracle to get
$F(x_1), \dots, F(x_\ell) \in [M]$

Sample $y_1, \dots, y_\ell \leftarrow [N]$
Query the oracle to get
$F(y_1), \dots, F(y_\ell) \in [M]$

Send $F(x_1), \dots, F(x_\ell)$ —— $F(x_1), \dots, F(x_\ell)$ ——→ Send $F(y_j)$ If $\exists\, i \in [\ell]\ s.t.\, F(y_j) = F(x_i)$

Ouput $key_A := x_i$ ←—— $F(y_j)$ —— Output $key_B := y_j$

- Correctness:
  - Set $N := 10\ell^2, |\{x_1, \dots, x_\ell\} \cap \{y_1, \dots, y_\ell\}| = 1$ w.h.p. by birthday paradox.
  - If M is large enough, $key_A = key_B$ w.h.p.
- Security: the shared key x^* is uniformly distributed → The attacker should makes at least $\Omega(\ell^2)$ queries.

Merkle puzzle only provides a quadratic gap between the efficiency of the honest parties and the attacker.

Can we do better ?

[Noam23] proposed a variant of the Merkle Puzzle with perfect completeness and the same security.

Previous Lower bounds:

Impagliazzo and Rudich [IR89]

Any key agreement protocol where Alice and Bob each make ℓ queries can be broken by the attacker with $O(\ell^6)$ queries.

Intersection queries

Barak and Mahmoody [BM09]

Any key agreement protocol where Alice and Bob each make ℓ queries can be broken by the attacker with $O(\ell^2)$ queries.

Heavy queries

$\Pr[q \in Q(V)] \geq \varepsilon$.

Merkel Puzzle is optimal w.r.t. query complexity of the attacker!

The heavy query techniques have found wide applications in the context of black-box separations and the power of random oracles in secure two-party computation [KSY11, BKSY11, MP12, DSLMM11, MMP14, HOZ13].

Communication Lower bounds

The amount of communication bits between Alice and Bob is also Important in practice!
For example, in Merkle's Puzzles, Alice and Bob need to exchange $\Omega(\ell)$ bits.

Conjecture [HMOYR18]

Any ℓ-query and c bits communication KA non-adaptive protocols could be broken by the attacker with $O(c\ell)$-queries.

Non-adaptive: Alice and Bob decide their queries before protocol execution, i.e., their queries are fully determined by their internal randomness.

Theorem [HMOYR18]

Any ℓ-query and c bits communication KA non-adaptive two rounds protocols could be broken by the attacker with $O(c\ell)$-queries.

Heavy queries and analyze the communication cost via ad hoc techniques

Our Contribution

Main Theorem

Any ℓ-query and c bits communication KA non-adaptive and prefect completeness protocols could be broken by the attacker with $O(c\ell)$-queries.

Perfect Completeness: $\Pr[Key_A = Key_B] = 1$

The protocol in [Noam23] is optimal.

Technical contribution:

1、Correlated queries: the queries are not only heavy queries but also highly related to communication transcripts.

2、Analyze the communication cost via density increment arguments.

Correlated Queries

Correlated Query

Let τ be a transcript and L be the current queries of the attacker .We say S $\subseteq [N]$ is ϵ-correlated w.r.t. attacker's view (τ, L) if

$$\mathbf{H}(\boldsymbol{F}(S)|\boldsymbol{R_A}, \boldsymbol{R_B}, L) - \mathbf{H}(\boldsymbol{F}(S)|\, \boldsymbol{R_A}, \boldsymbol{R_B}, L, \tau) \geq \epsilon$$

Algorithm of the attacker:

Initialize $i = 0$ and $L = \emptyset$.
While exists S $\subseteq [N]$ is ϵ-correlated w.r.t. the attack's view (τ, L) with $|S| \leq \ell$:
  Query F on S and receive $F(S)$.
  Update $L = L \cup \big(S, F(S)\big)$ and $i = i + 1$.

How to bound the expected number of iterations?

Output $b = \max_{i \in \{0,1\}} \Pr_{v \leftarrow (R_A, R_B, F)|_{\tau, L}} [Key_A\,(v)\ = i]$.

$(R_A, R_B, F)|_{\tau,L}$ is the distribution of all possible execution condition on communication transcript τ and queries L.

Density Increment Argument

Density Function

Let τ be a transcript and L be the queries of the attacker, the density function $\Phi(\tau, L)$ is defined as follows:

$$\Phi(\tau, L) = \mathbf{H}(\boldsymbol{F} \mid \boldsymbol{R_A}, \boldsymbol{R_B}, L) - \mathbf{H}(\boldsymbol{F} \mid \boldsymbol{R_A}, \boldsymbol{R_B}, L, \tau)$$

Lemma 1: The expected number of iterations of the algorithm is $O(\mathrm{CC}(\Pi)/\epsilon)$.

$$\Phi(\tau, \emptyset) \Longrightarrow \Phi(\tau, L_1) \Longrightarrow \Phi(\tau, L_1 \cup L_2) \Longrightarrow \ldots\ldots. \Longrightarrow \Phi(\tau, L_1 \cup \cdots \cup L_c)$$

By Chain Rule,
the density function Φ decreases at least ϵ in expectation after ϵ-correlated queries in each iteration.

Notice that the density function Φ is always non-negative since $\boldsymbol{F}$ is a uniform distribution condition on $(\boldsymbol{R_A}, \boldsymbol{R_B}, L)$.

Thus, the expected number of iterations given τ is $\frac{\Phi(\tau,\emptyset)}{\epsilon}$ and the expected number of iterations given protocol Π is

$$\mathrm{E}_{\tau \leftarrow \Pi}\left[\frac{\Phi(\tau, \emptyset)}{\epsilon}\right] \leq \frac{H(F|R_A, R_B, L) - H(F \mid R_A, R_B, L, \Pi)}{\epsilon} \leq \frac{H(\Pi)}{\epsilon} \leq \frac{CC(\Pi)}{\epsilon}$$

Summary and Proof Outline

Main Theorem

Any ℓ-query and c bits communication KA non-adaptive and prefect completeness protocols could be broken by the attacker with $O(c\ell)$-queries.

The proof outline is as follows:

Algorithm: The attacker queries the ϵ-correlated queries in each iteration and outputs the majority of the possible output based on it's view (τ, L).

Lemma 1: The expected number of iterations of the algorithm is $O(\mathrm{CC}(\Pi)/\epsilon)$.

Proved by density increment arguments.

Lemma 2: The success probability of the algorithm is at least $1 - \sqrt{\epsilon}$.

Proved by the rectangle view in communication complexity. We omitted the proof in this talk

Open Problems

Main Theorem

Any ℓ-query and c bits communication KA non-adaptive and prefect completeness protocols could be broken by the attacker with $O(c\ell)$-queries.

Imperfect completeness?

Adaptive protocols?

Other applications via our density increment argument or correlated queries?

Thank you for listening ☺