NOVEL PROOFS AND ALGORITHMS FOR RANGE SEARCHABLE ENCRYPTION

Members:
Richard Ong Jun Quan
(NUS High School of Mathematics and Science)
Guan Keer (Hwa Chong Institution)
Claire-Leia Ng Shean Ee (Raffles Girls' School)

Mentors:
Ruth Ng Ii-Yung, John Khoo Teng Fong
(DSO National Laboratories)

Range Searchable Encryption (RSE)

Database encryption scheme to search for ranges (eg. Time, Salary) on an untrusted server privately

Index	Documents
\$1000	A, B, C
\$1500	D, E
\$2000	F
\$2500	G, H

Client Encrypt

Uploads to Server

Index	Documents
apoj	^, @, #
dskp	\$, !
mkjr	%
lmop	&, ?

Query: \$1500 - \$2500 → Token → **Documents:** \$, !, %, &, ? → Decrypt → **Documents**: D, E, F, G, H

Motivation

Cloud Storage - Increased accessibility and efficiency

Range Searching - Makes searching by range more convenient

User Privacy - Server has no knowledge of data being stored / returned

Current State

- Limited number of RSE schemes in literature
- Proof for MMEπ security is non-extensible
- Lack of flexible, efficient and secure c-cover algorithms

Novel Game Playing Provable Security Proofs for our RSEΩ scheme and the MMEπ scheme

Key Components

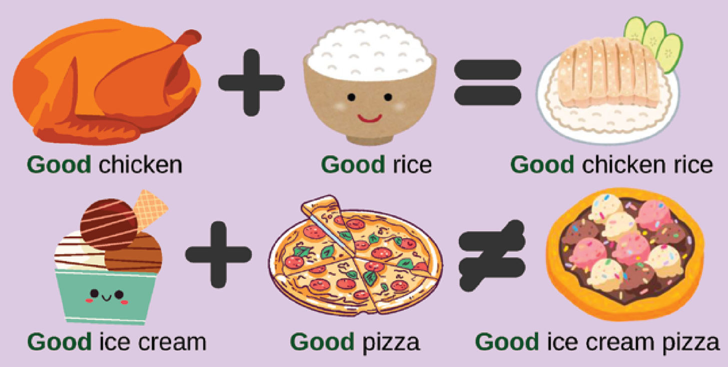
Game — *Algorithm designed to measure a certain security property by comparing real schemes to their ideal version*

Adversary — *Algorithm interacts with the game*
***Goal**: Differentiate between real and ideal schemes successfully*

Adversary's Advantage =
Pr [Correct Guess] - Pr [Wrong Guess]

Low advantage
↓
Low probability of correct guesses
↓
Adversary cannot differentiate between real and ideal scheme
↓
Real scheme achieves ideal properties
↓
Secure scheme

Purpose: Show security of a scheme follows from security of its components

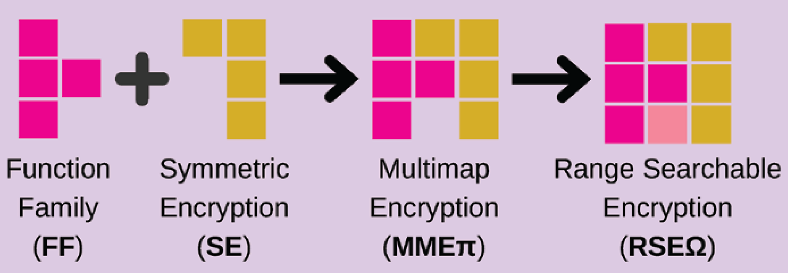


Good components don't always mean good products

Benefits

- Advantage allows us to concretely compare security
- Only have to focus on the security of underlying parts
- Mathematically prove security

Research Goal: Show the security of RSEΩ follows from FF and SE



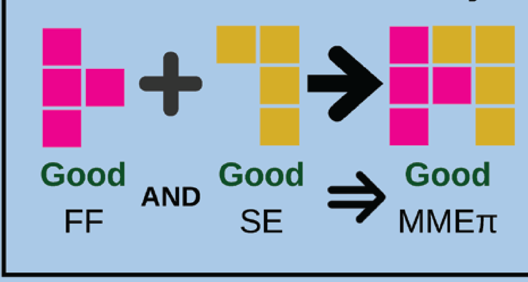
Our Contributions

MMEπ is an existing encrypted keyword search scheme used in many other schemes and implemented in the real world (eg. MongoDB)
RSEΩ is a novel scheme we designed built upon any MME scheme

Limitations in Existing Security Proofs:

- Non standard assumptions about components
- Non-extensible (Cannot be used to prove security of schemes)

Proof of MMEπ security



Theorem 1 : Given MME_π, and adversary A there exists adversaries B, C and D such that

$$\mathrm{Adv}^{\mathrm{ss}}_{\mathsf{MME}_\pi}(A) \le \mathrm{Adv}^{\mathrm{ind}}_{\mathsf{SE}}(B) + \mathrm{Adv}^{\mathrm{prf}}_{\mathsf{F}}(C) + (m-x)\mathrm{Adv}^{\mathrm{prf}}_{\mathsf{F}}(D)$$

where m is the number of labels in $\mathbf{M}$ and x is the number of distinct queries

Proof of RSEΩ security

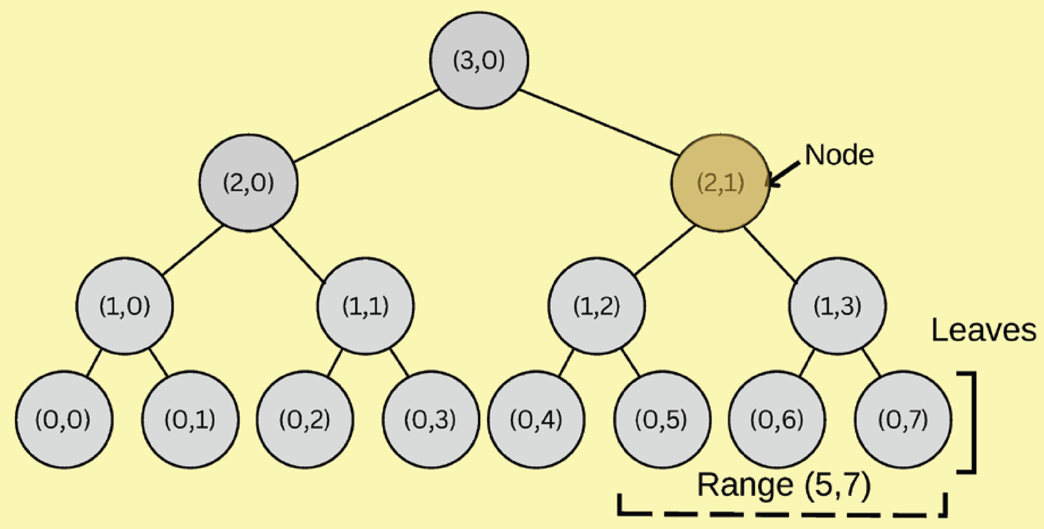
Theorem 2 : Given RSE_Ω, and adversary A there exists adversaries B, such that :

$$\mathrm{Adv}^{\mathrm{ss}}_{\mathsf{RSE}_\Omega}(A) \le \mathrm{Adv}^{\mathrm{ss}}_{\mathsf{MME}}(B)$$

- ✓ RSEΩ can be securely constructed and used
- ✓ Easy to analyse security of schemes using MMEπ + security of current implementations easily quantified

Our Novel Generic Overcover Algorithm with Greatest Efficiency and a Proof of Optimality

Binary Trees represent ranges in RSE



- **Documents** are stored under leaf nodes according to **index**
- **Overcover**: Set of nodes whose leaves covers exactly or beyond a range with extra nodes called **overhead (e)**
- **c-cover** - c nodes in cover (eg. 1-cover of (5,7): {(2,1)}, e=1)
- **Querying** for a cover will return documents in its range

- ✓ Improve efficiency by reducing network bandwidth
- ✓ Increase security by leaking less information

Our Contributions

Generic c-cover algorithm that is **optimal** (returns least c-cover with minimal overhead given a range (a,b) and integer c) and **efficient proof of optimality**

Relevance of Contributions:

- Overcover algorithms in literature are limited to certain c (1,3)
- Lack of proof of optimality
- Our previous c-cover algorithm was inefficient and hence impractical
- Our novel algorithm scales better
  $O(R^c) \rightarrow O(c^2)$ where R is the range size

eg. when R=200000 and c=5, our novel algorithm runs around 8×10^8 *times faster*

- ✓ Greater flexibility to adjust parameters
- ✓ Applicable in most RSEs
- ✓ Increased security and efficiency
- ✓ Reduced bandwidth

C-cover Algorithm Intuition

1. Find a **pair of nodes** that separate the range into two sides such that **overhead is minimised**
2. **Greedy algorithm** to find optimal cover