YDSP Young Defence Scientists Programme
DSTA Defence Science & Technology Agency
DSO National Laboratories

NEW FRAMEWORKS FOR ANALYSING SECURITY-EFFICIENCY TRADEOFFS IN RANGE SEARCHABLE ENCRYPTION

Members:
Soh Jun Heng, Ang Wei Sheng Wilson, Elton Ng Yew Tien (Catholic High School)

Mentors:
Ruth Ng Ii Yung, Phang Yan Feng Benito (DSO National Laboratories)

*What is Range Searchable Encryption (**RSE**)?*

(Sensitive) Employee Records (e.g. Salary, Bonuses) → Cloud Server

User (e.g. Company HR)

RSE WILL:	USER CAN:
Store files on **external** server	Access files **anywhere**
Index files	**Search** for files quickly
Encrypt files	Store files **securely**

Cloud Server

SERVER CANNOT:
See contents of files
Modify files

Methodology

- Complexity theory to describe space/time efficiency + security
  - Theoretical models of scheme performance **as input size increases**
  - **High-level view** of scheme performance

Advantages:

- Universal analysis, system independent
- More objective and applicable comparisons

Existing Literature	Problem	Our Contributions
Introduced different RSE schemes[1] [2]	**NO standardisation** in comparison ▼ Unable to directly compare schemes ▼ **Confusion** in picking scheme	**Standardise** comparison metrics
Qualitative security analysis of schemes		**Novel quantifiable measure** of information leakage of schemes
Comparison Metrics: (1) Time Efficiency, (2) Security, (3) Space Efficiency		

Impacts of Our Work

- Developers can pick **efficient RSE scheme**
- Save resources! (time, storage, **$$$**)

(1) Time Complexity Analysis

1. **(Setup)** Time to **set up** encrypted database on server
2. **(Query)** Time to **search for** + **retrieve** files from server

→ Runtime Analysis of algorithms in schemes

(2) Novel Volumetric Analysis Framework & Security Analysis

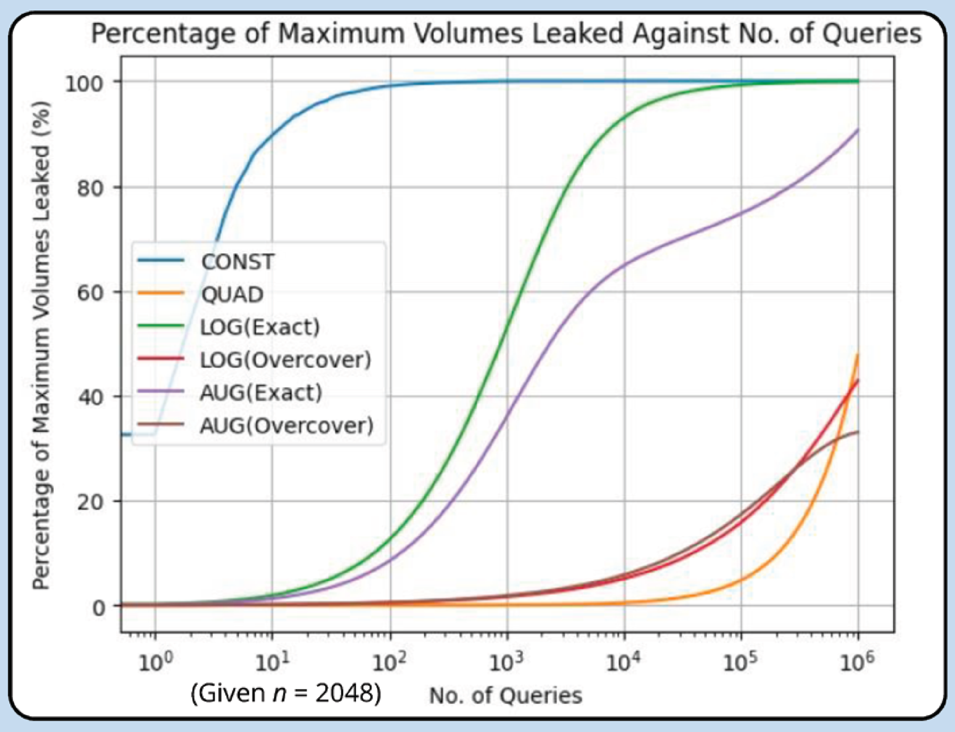
Different kinds of info leaked to attacker for each entry accessed in server

▼

Consolidate into **single measure**
1 'Volume' = 1 unit of info leakage for each entry in server

▼

Higher % of Max Volumes Leaked
→ **More Info Leaked** → **Less Secure!**



Overall Security-Efficiency Tradeoffs

Scheme	Space Efficiency			Time Efficiency		Security
	Storage	Bandwidth (Upload)	Bandwidth (Download)	Runtime (Setup)	Runtime (Query)	Volume Leakage
CONST	$\theta(n)$	$\theta(R)$	$\theta(R)$	$\theta(n)$	$\theta(R)$	$O(R)$
QUAD	$\theta(n^3)$	$\theta(1)$	$\theta(R)$	$\theta(n^3)$	$\theta(R)$	$O(1)$
LOG (Overcover)	$\theta(n \log_2 n)$	$\theta(1)$	$O(n)$	$\theta(n \log_2 n)$	$O(n)$	$O(1)$
LOG (Exact)	$\theta(n \log_2 n)$	$O(\log_2 R)$	$\theta(R)$	$\theta(n \log_2 n)$	$\theta(R)$	$O(\log_2 R)$
AUG (Overcover)	$\theta(n \log_2 n)$	$\theta(1)$	$O(n)$	$\theta(n \log_2 n)$	$O(n)$	$O(1)$
AUG (Exact)	$\theta(n \log_2 n)$	$O(\log_2 R)$	$\theta(R)$	$\theta(n \log_2 n)$	$\theta(R)$	$O(\log_2 R)$

Legend

Green	Excellent	Yellow	Fair
Lime	Good	Red	Terrible

R: Size of user query
n: No. of unique files in scheme
$n \geq R$ (n is the max user can query for)

O: Upper bound (scheme's worst case)
θ: Tight bound (scheme's average case)

(3) Space Complexity Analysis

(3.1) Storage

Given ***n*** unique files (input), some schemes store **duplicate files** on the server (due to scheme design).

More Files Stored → Less Space Efficient → Higher Server Cost

Scheme	CONST	LOG	AUG	QUAD
Total no. of files	n	$n(\log_2 n + 1)$	$2n(\log_2 n) - n + 2$	$\frac{n^3 + 3n^2 + 2n}{6}$

MOST space efficient → **LEAST** space efficient
n: number of unique files in scheme

(3.2) Bandwidth

Bandwidth (size of user's upload/download)

Upload (to server: query)

Download (from server: encrypted files)

→

- Simulations on average bandwidth
- Formulated "Worst Case"

Final Recommendations

(1) DO NOT IMPLEMENT CONST or QUAD
- Terrible security and storage respectively

(2) USE LOG or AUG INSTEAD
- Balanced across all metrics
- Suitable for most use cases

(3) Pros of Overcover + Exact Cover
- Overcover: more secure
- Exact Cover: better for frequent querying

Future Work

1. Devise other metrics
2. Evaluate more schemes

→ More options + information, Developers make better choices

References:
[1] Demertzis, I., Papadopoulos, S., Papapetrou, O., Deligiannakis, A., & Garofalakis, M. (2016). Practical private range search revisited. SIGMOD '16: Proceedings of the 2016 International Conference on Management of Data, 185–198. https://doi.org/10.1145/2882903.2882911
[2] Faber, S., Jarecki, S., Krawczyk, H., Nguyen, Q., Rosu, M., & Steiner, M. (2015). Rich Queries on Encrypted Data: Beyond Exact Matches. In Lecture notes in computer science (pp. 123–145). https://doi.org/10.1007/978-3-319-24177-7_7