Processing Analytical Queries over Encrypted Data

Štěpán Kozák

FI MUNI

14.10.2013

Motivation of search outsourcing

Industry needs to have a cheap storage with **search as a service** over the stored data.

Key advantages of clouds:

- **Low initial investments**
- **Low storage costs**
- **Very good scalability and flexibility**

Problems of the outsourcing:

- **Data can be sensitive!**
- High cost of data transfer via network

MONOMI

MONOMI is the privacy-preserving extension over Postgre database.

Supports 19 out of 22 TPC-H queries.

MONOMI - Cryptography it uses

Encryption scheme	**SQL operations**	**Leakage**
Randomized AES	None	None
Deterministic AES	GROUP BY, EQUI-JOIN	Duplicates
OPE	ORDER BY, $<$, $>$ predicates	Order
Paillier	a + b, SUM(a)	None

MONOMI - Example

Consider following Orders table with scheme:

OrderID (int)	Price (float)

MONOMI stores the columns in the following way

- **OrderID** encrypted with **deterministic AES**
- **Price** encrypted with **Paillier** encryption

MONOMI - Example query 1

Consider issuing following query:

```
SELECT OrderID
FROM Orders
WHERE Price > 100
```

How does it translate within MONOMI?

MONOMI - Example query 1

Original query:

```
SELECT OrderID
FROM Orders
WHERE Price > 100
```

Transformed query:

```
SELECT OrderID_Det_AES
FROM Orders
WHERE Price_OPE > Encrypt_OPE(100)
```

MONOMI - Example query 2

Consider issuing following query:

```
SELECT SUM(Price) AS Total
FROM Orders
GROUP BY OrderID
HAVING Total > 100
```

How does it translate within MONOMI?

MONOMI - Example query 2

Original query:

```
SELECT SUM(Price) AS Total
FROM Orders
GROUP BY OrderID
HAVING Total > 100
```

Transformed query:

```
SELECT PAILLIER_SUM(Price_Paillier) AS Total
FROM Orders
GROUP BY OrderID_Det_AES
```

MONOMI Optimizer

- **Per-row precomputations**
  MONOMI materializes some precalculated values so it's possible to operate with them in an encrypted manner
- **Space-efficient encryption**
  Especially Paillier cryptosystem uses very large integers (2048 bit) to store ciphertexts. MONOMI packs values from multiple columns in a single row and packs values from multiple rows, into a single Paillier plaintext.

MONOMI Designer

Tool which computes recommended physical design

- What encryption should be used for particular columns
- Which computed columns to precompute

Takes as the input

- Sample query set
- Sample data set to be used

MONOMI Performance

The authors say the over head is only 1.24x comparing to unencrypted DB (ranging from 1.03x to 2.33x).

However ...

They do not consider the communication cost.

Questions

?