Asymmetric Structure-preserving Subgraph Queries for Large Graphs

Zhe FAN, Byron CHOI, Jianliang XU
Hong Kong Baptist University

Sourva S. Bhowmick
Nanyang Technological University

Research Problem

- **Client**
  - Issue subgraph query $Q$

- **Service Provider (SP)**
  - Own the data graph $G$
  - Evaluate the subgraph query
  - Return Yes or No

- **Attack Model**
  - Honest-but-Curious
  - Chosen Plaintext Attack (CPA)

- **Privacy Target**
  - Structure of query graph $Q$ (adjacency matrix $M_Q$)

- **Problem Formulation**
  - Evaluate subgraph query with preserving the privacy target under the attack model

Challenges and Contributions

- **Challenges**
  - How to reduce the candidate mappings with preserving the query structure?
  - How to verify the candidate mappings without leaking the query structure?

- **Contributions**
  - A new candidate subgraph exploration
  - A novel neighborhood containment class
  - An efficient subgraph cache to prune redundant matching
  - A robust encoding scheme and verification method
  - Privacy analysis and extensive experiments

Structure-preserving Subgraph Isomorphism

- **Pre-processing at the client**
  - Retrieve optimization parameters
  - Encode entries of $M_Q$
  - Encrypt the encoded $M_Q$ (CGBE)

- **Verification at the client**
  - Decrypt and obtain the encoding

- **Reducing Candidate Mappings**
  - Generate Candidate Subgraphs
  - Minimize Candidate Subgraphs
  - Generate Candidate Matchings
  - Prune Candidate Matchings
  - Generate Candidate Mappings

- **Verifying Candidate Mappings**
  - SPVerify: Aggregate constant mappings and negligible positive
  - SPVerify*: Queries with bounded size and negligible positive

Privacy analysis

- Privacy of the encryption scheme
  - CGBE is secure under CPA
  - $M_Q$ is protected under CGBE

- Privacy of SPsubiso
  - Query structure is not used in reducing candidate mappings
  - Fixed number of mathematical operations regardless of the query structure in verifying candidate mappings

Experimental Evaluation

- Evaluate with real-world datasets on commodity machine
  - DBLP, Amazon, Youtube and LiveJournal
  - Intel i7 3.4GHz, 16GB

SUGAR: SecUre GrAph queRy services http://www.comp.hkbu.edu.hk/~bchoi/sugar.html