CS848 Winter 2021 Assignment 3

[**TODO**]: [Student Name]

Due by March 29, 11pm

Please submit a pdf solution on Learn: "Submit" → "Dropbox folder" → "Assignment 3".

Q1. Design Choice I

Consider a database schema consists of 3 attributes Grade(StudentID, ClassID, Score). Below is a listed of intended queries on this database:

- Select * from Grade where StudentID = "xh23";
- Select ClassID, Count(*) from Grade Group By ClassID;
- Select Count(*) from Grade where Score $\in [60, 100]$;
- Select Max(Score) from Grade where classID ="cs848";

(a) Ciphertext Schema

[8 points] Consider the possible encryption schemes (e.g. slide 23, or the onion of encryption on slide 33). What ciphertext schema do you recommend for an encrypted database stored using CryptDB?
[**TODO**]: State the ciphertext schema, e.g. slide 31 or slide 34, and justify your choice in terms of (i) security guarantee; (ii) storage cost; (iii) performance.

(b) Non-sensitive/sensitive Attributes

[2 points] If ClassID is not a sensitive attribute that requires no encryption, will you make any changes to your ciphertext schema?
[**TODO**]: State Yes/No and justify your choice.

Q2. Design Choice II

Recall the edge table from Assignment 1: a table of edges (e.g. Table 1) for an undirected graph of n nodes. Each row in the table represents an edge (v_i, v_j) in the graph, the node with smaller index stores in the first column (source) and the node with bigger index stores in the second column (dest). Suppose this table is encrypted and stored on the cloud. The client is interested in querying the triangle count in the graph.

source	dest
v_1	v_2
v_1	v_3
v_2	v_3
$\vdots$	$\vdots$
v_{n-1}	v_n

Table 1: Edge Table

(a) Strong Encryption

[5 points] If all the node ids in the Edge Table are encrypted with RND encryption, please discuss how to answer a triangle counting query on this encrypted table stored on a untrusted cloud?
[TODO]: Please state the assumptions (trusted client/hardware), the encryption protocols for the columns, and how the query could be answered, the performance overhead for the client and the server.

(b) Weaker Encryption

[5 points] Please describe how to apply a weaker encryption such that the query can be processed more efficiently. Also, describe what is the security trade-off for the performance gain.
[TODO]: Please state the assumptions (trusted client/hardware), the encryption protocols for the columns, and how the query could be answered, the performance overhead for the client and the server, and state the possible information leakage/problems