A IoT-benchmark Data Generation Algorithm

Algorithm 1 describes the data generation strategy. For parameter details, please refer to Table 1.

Table 1: Parameters Description of Data Generator

Notation	Data Features
μ_v	Mean of values
μ_d	Mean of deltas
θ_d	Variance of deltas
γ	Repeat Rate
η	Increase rate
l	Series length

Algorithm 1: Numerical data generator [45]

Input: μ_v, μ_d, θ_d, γ, η, length n
Output: TS
$DS := empty\_list()$;
while $|DS| < n$ **do**
  $isRepeat := random\_index(\gamma)$;
  if $isRepeat$ **then**
  else
    $repeat\_len := random(8, T)$;
    $DS.append(0, repeat\_len)$;
  end
  $isPositive := random\_index(\eta)$;
  $delta := 0$;
  if $isPositive$ **then**
  else
    while $delta \leq 0$ **do**
      $delta := random\_guass(\mu_d, \theta_d)$;
    end
  end
  while $delta \geq 0$ **do**
    $delta := random\_guass(\mu_d, \theta_d)$;
  end
  $DS.append(delta)$;
end
$TS := prefix\_sum(DS)$;
$TS.zoom(\mu_v)$;
return TS;

B Proof of Proposition 4.11

The proof of Proposition 4.11 is as follows.

Proof. Let $Cost\left(I_u\left(c\right)\right)$ denote the cost of the traditional query process of a compressed database, i.e. decompressing first, then restore and query on uncompressed data, where

$$\begin{aligned}
&\mathrm{Cost}(I_u\left(c\right)) = Cost\left(\left(Q_u \circ R_u \circ U\right)\left(c\right)\right) = Cost\left(Q_u\left(R_u\left(U\left(c\right)\right)\right)\right)\\
&= \mathrm{Cost}\left(\left\{u: \begin{array}{c} u_1 \leftarrow op_1\left(u_0\right), u_2 \leftarrow op_2\left(u_1\right), \\ \ldots, u \leftarrow op_n\left(u_{n-1}\right) \end{array} \middle| \begin{array}{c} i \in \{1, \ldots, n\} \\ op_i \in \Pi \end{array} \right\}\right)\\
&\quad + \mathrm{Cost}(R_u\left(u\right)) + Cost\left(U\left(c\right)\right)\\
&= \mathrm{Cost}(U\left(c\right)) + Cost\left(R_u\left(u\right)\right) + Cost\left(\left\{op_1, op_2, \ldots, op_{n-1}\right\}\left(u_1\right)\right).
\end{aligned}$$

Let $Cost\left(I_c\left(c\right)\right)$ denote the cost of the partial homomorphic query process of a compressed database, where

$$\begin{aligned}
&\mathrm{Cost}(I_c\left(c\right)) = Cost\left(\left(I'_u \circ Q_c \circ R_c\right)\left(c\right)\right) = Cost\left(I'_u\left(Q_c\left(R_c\left(c\right)\right)\right)\right)\\
&= \mathrm{Cost}\left(I'_u\left(c_j\right)\right) + Cost\left(Q_c\left(c_0\right)\right) + Cost\left(R_c\left(c\right)\right)\\
&= \mathrm{Cost}\left(\left\{u: \begin{array}{c} u_j \leftarrow op_j\left(u_{j-1}\right), u_{j+1} \leftarrow op_{j+1}\left(u_j\right), \\ \ldots, u \leftarrow op_n\left(u_{n-1}\right) \end{array} \middle| \begin{array}{c} i \in \{j, \ldots, n\} \\ op_i \in \Pi \end{array} \right\}\right)\\
&\quad + \mathrm{Cost}\left(\left\{c_j: \begin{array}{c} c_1 \leftarrow op'_1\left(c_0\right), c_2 \leftarrow op'_2\left(c_1\right), \\ \ldots, c_j \leftarrow op'_j\left(c_{j-1}\right) \end{array} \middle| \begin{array}{c} i \in \{1, \ldots, j\} \\ op'_i \in \Theta \end{array} \right\}\right)\\
&\quad + \mathrm{Cost}\left(U\left(c_j\right)\right) + Cost\left(R_c\left(c\right)\right)\\
&= \mathrm{Cost}(R_c\left(c\right)) + Cost\left(\left\{op'_1, op'_2, ..op'_j\right\}\left(c_0\right)\right) + Cost\left(U\left(c_j\right)\right)\\
&\quad + \mathrm{Cost}\left(\left\{op_{j+1}, op_{j+2}, \ldots, op_{n-1}\right\}\left(u_{j-1}\right)\right),
\end{aligned}$$

with $\varphi\left(u_i\right) = c_i, j \in \{1...n-1\}$. Then we have
$\mathrm{Cost}(I_u\left(c\right)) - Cost\left(I_c\left(c\right)\right)$

$$\begin{aligned}
&= \mathrm{Cost}(U\left(c\right)) + Cost\left(R_u\left(u\right)\right) + Cost\left(\left\{op_1, op_2, \ldots, op_{n-1}\right\}\left(u_0\right)\right)\\
&\quad - \left(Cost\left(R_c\left(c\right)\right) + Cost\left(\left\{op'_1, op'_2, ..op'_j\right\}\left(c_0\right)\right) + Cost\left(U\left(c_j\right)\right)\right.\\
&\quad \left. + \mathrm{Cost}\left(\left\{op_{j+1}, op_{j+2}, \ldots, op_{n-1}\right\}\left(u_{j-1}\right)\right)\right)\\
&= \left(Cost\left(U\left(c\right)\right) - Cost\left(U\left(c_j\right)\right)\right) + \left(Cost\left(R_u\left(u\right)\right) - Cost\left(R_c\left(c\right)\right)\right)\\
&\quad + \left(Cost\left(\left\{op_1, op_2, \ldots, op_{n-1}\right\}\left(u_0\right)\right) - Cost\left(\left\{op'_1, op'_2, ..op'_j\right\}\left(c_0\right)\right)\right.\\
&\quad \left. - \mathrm{Cost}\left(\left\{op_{j+1}, op_{j+2}, \ldots, op_{n-1}\right\}\left(u_{j-1}\right)\right)\right)\\
&= \left(Cost\left(U\left(c\right)\right) - Cost\left(U\left(c_j\right)\right)\right) + \left(Cost\left(R_u\left(u\right)\right) - Cost\left(R_c\left(c\right)\right)\right)\\
&\quad + \left(Cost\left(\left\{op_1, op_2, \ldots, op_j\right\}\left(u_0\right)\right) - Cost\left(\left\{op'_1, op'_2, ..op'_j\right\}\left(c_0\right)\right)\right)
\end{aligned}$$

Note that when $j = n-1$, query Q is a fully homomorphic query.
By Lemma 4.10, $Size\left(c\right) \geq Size\left(c_j\right)$, thus, $Cost\left(U\left(c\right)\right) \geq Cost\left(U\left(c_j\right)\right)$. And by Definition 4.9, we have

$$Cost\left(\left\{op_1, op_2, \ldots, op_j\right\}\left(u_0\right)\right) \geq Cost\left(\left\{op'_1, op'_2, \ldots, op'_j\right\}\left(c_0\right)\right).$$

And by Definition 4.8, we have

$$Cost\left(R_u\left(u\right)\right) \geq Cost\left(R_c\left(c\right)\right)$$

Thus, we have $Cost\left(I_u\left(c\right)\right) - Cost\left(I_c\left(c\right)\right) \geq 0$, i.e.,

$$Cost\left(I_c\left(c\right)\right) \geq Cost\left(I_c\left(c\right)\right).$$

□