Appendix

Below are more details that are omitted in the main text due to space limitations.

1 Input Perturbation Algorithm

The input perturbation algorithm [3] is straightforward. It adds Laplace noise with privacy parameter ϵ to the weight value of every edge in graph G. The algorithm is described as follows:

Algorithm 1 Input Perturbation Algorithm

Input:

- Graph $G(V, E, w)$
- Privacy parameter ϵ

Output:

- Shortest-path distances $\widetilde{d}(u, v)$ for all node pairs $(u, v) \in V$

for each edge $e \in E$ **do**
  Sample noise $\eta \sim \text{Lap}(1/\epsilon)$
  $\widetilde{w}(e) \leftarrow w(e) + \eta$
end for
for all pairs of nodes $u, v \in V$ **do**
  Calculate the perturbed shortest-path distance $\widetilde{d}(u, v)$ on the graph with perturbed weights $\widetilde{w}$
end for
return All pairwise distances $\widetilde{d}(u, v)$

1.1 Privacy and Error Analysis

Theorem 1. *Given a graph $G = (V, E, w)$, the input perturbation algorithm satisfies ϵ-differential privacy.*

Proof. According to the edge-weight privacy model, the total weight on the edges of two neighboring graphs G and G' differs by at most 1. Consequently, the weight of each individual edge also differs by at most 1. For the weight function $w : E \to \mathbb{R}^+$, its ℓ_1-sensitivity satisfies $\Delta_1(w) \leq 1$. By the Laplace mechanism, adding noise sampled from $\mathrm{Lap}(1/\epsilon)$ to each edge weight ensures that the weight function satisfies ϵ-DP. According to the post-processing property of differential privacy, the shortest path algorithm acts as a post-processing function. Therefore, the entire input perturbation algorithm satisfies ϵ-DP. □

Theorem 2. *Given a graph $G = (V, E, w)$ with n nodes, the error upper bound of the input perturbation algorithm does not exceed $O\left(\frac{n \log(n/\beta)}{\epsilon}\right)$ with probability at least $1 - \beta$.*

Proof. For graph G, Laplace noise $\eta \sim \mathrm{Lap}(1/\epsilon)$ is added to each edge. A property of the Laplace distribution is that $\Pr[|\eta| \geq t/\epsilon] = e^{-t}$. For all $|E|$ edges of graph G, define the event $A_e = \{|\eta_e| \geq t/\epsilon\}$. By the union bound, the probability that there exists at least one edge for which the absolute value of the added noise exceeds t/ϵ is

$$\Pr\left(\bigcup_{e \in E} A_e\right) \leq \sum_{e \in E} \Pr(A_e) = |E| \cdot e^{-t}. \tag{1}$$

Setting $t = \log(|E|/\beta)$, it holds with probability at least $1 - \beta$ that the noise added to every edge in G does not exceed $\log(|E|/\beta)/\epsilon$. For any pair of nodes (u, v), the number of edges on the shortest path does not exceed $n - 1$. Therefore, the error of the shortest path distance satisfies

$$\begin{aligned} \left|\mathrm{dist}(u, v) - \widetilde{\mathrm{dist}}(u, v)\right| &\leq O\left(n \cdot \frac{\log(|E|/\beta)}{\epsilon}\right) \\ &= O\left(\frac{n \log(n/\beta)}{\epsilon}\right). \end{aligned} \tag{2}$$

Thus, the error upper bound of the input perturbation algorithm is $O\left(\frac{n \log(n/\beta)}{\epsilon}\right)$. □

2 Output Perturbation Algorithm

For a given graph G, the output perturbation algorithm [3] first computes the shortest paths between all pairs of nodes and then adds Laplace noise to these distances to achieve (ϵ, δ)-differential privacy. The details of the algorithm are as follows.

Algorithm 2 Output Perturbation Algorithm

Input:

- Graph $G(V, E, w)$
- Privacy parameters ϵ, δ

Output:

- Shortest-path distances $\widetilde{d}(u, v)$ for all node pairs $(u, v) \in V$

for all pairs of nodes $u, v \in V$ **do**
  $d \leftarrow \text{dist}(u, v)$ {Compute the exact shortest-path distance}
  Sample noise $\eta \sim \text{Lap}\left(O\left(n\sqrt{\ln(1/\delta)}\right)/\epsilon\right)$
  $\widetilde{d}(u, v) \leftarrow d + \eta$
end for
return All pairwise distances $\widetilde{d}(u, v)$

2.1 Privacy and Error Analysis

Theorem 3. *Given a graph $G = (V, E, w)$, the output perturbation algorithm satisfies (ϵ, δ)-differential privacy.*

Proof. Since the total weight on the edges of two neighboring graphs G and G' differs by at most 1, the ℓ_1-sensitivity of the shortest path algorithm f satisfies $\Delta_1(f) \leq 1$ for any single pair query. Therefore, each shortest path query can be answered by adding noise from $\text{Lap}(1/\epsilon')$, satisfying $(\epsilon', 0)$-DP for that query. With n^2 such queries, according to the advanced composition theorem of differential privacy, the entire output perturbation algorithm satisfies $(n\sqrt{2\ln(1/\delta)}\epsilon' + n^2\epsilon'(e^{\epsilon'} - 1), \delta)$-DP for $\delta > 0$. By setting $\epsilon' = \epsilon/O(n\sqrt{2\ln(1/\delta)})$, the output perturbation algorithm achieves (ϵ, δ)-DP. □

Theorem 4. *Given a graph $G = (V, E, w)$ with n nodes, the error upper bound of the output perturbation algorithm does not exceed*

$$O\left(\frac{n\sqrt{\ln(1/\delta)}\log(n^2/\beta)}{\epsilon}\right)$$

with probability at least $1 - \beta$.

Proof. For a single shortest path distance, the magnitude of the added Laplace noise η satisfies $\Pr\left[|\eta| > n\sqrt{\ln(1/\delta)}\log(1/\beta)/\epsilon\right] = \beta$ due to the properties of the Laplace distribution. Applying the union bound over all n^2 pairwise queries, the probability that the error exceeds $n\sqrt{\ln(1/\delta)}\log(n^2/\beta)/\epsilon$ for *any* pair is at most β. Therefore, with probability at least $1 - \beta$, the error for all node pairs is bounded by $O\left(n\sqrt{\ln(1/\delta)}\log(n^2/\beta)/\epsilon\right)$. □

3 Proof of Theorem 3.2

Theorem 3.2 *Given a graph $G = (V, E, w)$ with n nodes, the error upper bound of the SynGraph algorithm does not exceed $O\left(\sqrt{n}\log^2(n/\beta)/\epsilon\right)$ with probability at least $1 - 2\beta$.*

Proof. For any pair of nodes (u, v) in G, consider the t-hop shortest path distance $\mathrm{dist}_t(u, v)$. After adding noise via the output perturbation mechanism, we have $\widetilde{\mathrm{dist}}_t(u, v)$. According to theorem4, with probability at least $1 - \beta$:

$$\left|\widetilde{\mathrm{dist}}_t(u, v) - \mathrm{dist}_t(u, v)\right| \leq O\left(t\log(n/\beta)/\epsilon\right) \leq O\left(\sqrt{n}\log^2(n/\beta)/\epsilon\right). \quad (3)$$

For any pair (u, v) within the sampled subset S, after adding output perturbation noise, with probability at least $1 - \beta$:

$$\left|\widetilde{\mathrm{dist}}_S(u, v) - \mathrm{dist}_S(u, v)\right| \leq O\left(\sqrt{n\ln(1/\delta)}\log(n/\beta)/\epsilon\right). \quad (4)$$

Therefore, for any pair (u, v) in G, with probability at least $1 - 2\beta$:

$$\begin{aligned}
\widetilde{\mathrm{dist}}(u,v) &= \min\left\{\widetilde{\mathrm{dist}}_t(u,v),\ \min_{m,n\in S}\left\{\widetilde{\mathrm{dist}}_t(u,m)+\widetilde{\mathrm{dist}}_S(m,n)+\widetilde{\mathrm{dist}}_t(n,v)\right\}\right\} \\
&\geq \min\left\{\mathrm{dist}_t(u,v),\ \min_{m,n\in S}\left\{\mathrm{dist}_t(u,m)+\mathrm{dist}_S(m,n)+\mathrm{dist}_t(n,v)\right\}\right\} \\
&\quad - O\left(\sqrt{n}\log^2(n/\beta)/\epsilon\right) - O\left(\sqrt{n\ln(1/\delta)}\log(n/\beta)/\epsilon\right) \\
&\geq \mathrm{dist}(u,v) - O\left(\sqrt{n}\log^2(n/\beta)/\epsilon\right) - O\left(\sqrt{n\ln(1/\delta)}\log(n/\beta)/\epsilon\right) \\
&\geq \mathrm{dist}(u,v) - O\left(\sqrt{n}\log^2(n/\beta)/\epsilon\right).
\end{aligned}$$

That is,

$$\widetilde{\mathrm{dist}}(u,v) - \mathrm{dist}(u,v) \geq -O\left(\sqrt{n}\log^2(n/\beta)/\epsilon\right). \tag{5}$$

If the true shortest path between u and v has more than t hops:

$$\begin{aligned}
\widetilde{\mathrm{dist}}(u,v) &\leq \widetilde{\mathrm{dist}}_t(u,v) \\
&\leq \mathrm{dist}_t(u,v) + O\left(\sqrt{n}\log^2(n/\beta)/\epsilon\right) \\
&\leq \mathrm{dist}(u,v) + O\left(\sqrt{n}\log^2(n/\beta)/\epsilon\right).
\end{aligned} \tag{6}$$

If the true shortest path has fewer than t hops:

$$\begin{aligned}
\widetilde{\mathrm{dist}}(u,v) &\leq \min_{m,n\in S}\left\{\widetilde{\mathrm{dist}}_t(u,m)+\widetilde{\mathrm{dist}}_S(m,n)+\widetilde{\mathrm{dist}}_t(n,v)\right\} \\
&\leq \min_{m,n\in S}\left\{\mathrm{dist}_t(u,m)+\mathrm{dist}_S(m,n)+\mathrm{dist}_t(n,v)\right\} \\
&\quad + O\left(\sqrt{n}\log^2(n/\beta)/\epsilon\right) + O\left(\sqrt{n\ln(1/\delta)}\log(n/\beta)/\epsilon\right) \\
&= \mathrm{dist}(u,v) + O\left(\sqrt{n}\log^2(n/\beta)/\epsilon\right) \\
&\quad + O\left(\sqrt{n\ln(1/\delta)}\log(n/\beta)/\epsilon\right) \\
&= \mathrm{dist}(u,v) + O\left(\sqrt{n}\log^2(n/\beta)/\epsilon\right).
\end{aligned} \tag{7}$$

In summary, with probability at least $1 - 2\beta$, the error upper bound is $O\left(\sqrt{n}\log^2(n/\beta)/\epsilon\right)$. □

4 ALDP Mechanism Specification

Given a weight vector $\mathbf{w}_i$ from client C_i with elements bounded in $[w_{\min}, w_{\max}]$, and privacy budget ϵ, the ALDP mechanism [2] operates as follows:

1. **Input:** Client's weight vector $\mathbf{w}_i$, value range $[w_{\min}, w_{\max}]$, privacy budget ϵ

2. **Midpoint calculation:**

$$\text{mid} = \frac{w_{\max} + w_{\min}}{2} \tag{8}$$

3. **Element-wise perturbation:** For each element w in $\mathbf{w}_i$:

   (a) Compute centered value: $\mu = w - \text{mid}$

   (b) Sample Bernoulli variable u with probability:

$$\mathbb{P}[u = 1] = \frac{e^\epsilon}{e^\epsilon + 1} \tag{9}$$

   (c) Generate perturbed value:

$$\widetilde{w} = \text{mid} + \mu \cdot \begin{cases} \frac{e^\epsilon + 1}{e^\epsilon - 1} & \text{if } u = 1 \\ -\frac{e^\epsilon + 1}{e^\epsilon - 1} & \text{otherwise} \end{cases} \tag{10}$$

4. **Output:** Perturbed weight vector $\widetilde{\mathbf{w}}_i$

Theorem 5 (Unbiased Estimation)**.** *Given a weight vector $\mathbf{w}_i$ from client C_i, for each weight w, the ALDP perturbation mechanism provides an unbiased estimate.*

Proof. The expectation derivation is shown below:

$$\begin{aligned} \mathbb{E}[\widetilde{w}] &= \frac{e^\epsilon}{e^\epsilon + 1}\left(\text{mid} + \mu\frac{e^\epsilon + 1}{e^\epsilon - 1}\right) + \frac{1}{e^\epsilon + 1}\left(\text{mid} - \mu\frac{e^\epsilon + 1}{e^\epsilon - 1}\right) \\ &= \text{mid} + \mu\left[\frac{e^\epsilon}{e^\epsilon - 1} - \frac{1}{e^\epsilon - 1}\right] \\ &= \text{mid} + \mu\left(\frac{e^\epsilon - 1}{e^\epsilon - 1}\right) \\ &= \text{mid} + \mu = w \end{aligned} \tag{11}$$

where $\mu = w - \text{mid}$ and $\text{mid} = \frac{w_{\max} + w_{\min}}{2}$. $\square$

Algorithm 3 ALDP Perturbation Mechanism

Input:

- Weight vector $\mathbf{w}_i$ from client C_i
- Minimum weight $w_{\min}$, maximum weight $w_{\max}$
- Privacy budget ϵ

Output:

- Perturbed weight vector $\widetilde{\mathbf{w}}_i$

mid $\leftarrow (w_{\max} + w_{\min})/2$
for each $w \in \mathbf{w}_i$ **do**
    $\mu \leftarrow w - \text{mid}$
    Sample Bernoulli variable u with $\mathbb{P}[u = 1] = \frac{e^\epsilon}{e^\epsilon+1}$
    if $u = 1$ **then**
        $\widetilde{w} \leftarrow \text{mid} + \mu \cdot \frac{e^\epsilon+1}{e^\epsilon-1}$
    else
        $\widetilde{w} \leftarrow \text{mid} - \mu \cdot \frac{e^\epsilon+1}{e^\epsilon-1}$
    end if
    Store $\widetilde{w}$ in $\widetilde{\mathbf{w}}_i$
end for
return $\widetilde{\mathbf{w}}_i$

Theorem 6 (ϵ-LDP Guarantee)**.** *Given a weight vector $\mathbf{w}_i$ from client C_i, for each weight w, the ALDP mechanism satisfies ϵ-local differential privacy.*

Proof. For adjacent weights w and w', the privacy ratio is bounded as:

$$\frac{\mathbb{P}[\mathcal{M}(w) = \widetilde{w}]}{\mathbb{P}[\mathcal{M}(w') = \widetilde{w}]} \leq \frac{\max_w \mathbb{P}[\mathcal{M}(w) = \widetilde{w}]}{\min_{w'} \mathbb{P}[\mathcal{M}(w') = \widetilde{w}]} = e^\epsilon \tag{12}$$

This holds for both possible outputs of the mechanism. □

Theorem 7 (Error Bound)**.** *Given a weight vector $\mathbf{w}_i$ from client C_i, for each weight w, with probability at least $1 - \beta$, the error of the ALDP mechanism is bounded by $O\left((w_{\max} - w_{\min}) \log(2/\beta)/\epsilon\right)$.*

Proof. First, the variance is bounded by:

$$\mathrm{Var}[\widetilde{w}] \leq \left(\frac{w_{\max} - w_{\min}}{2}\right)^2 \left(\frac{e^\epsilon + 1}{e^\epsilon - 1}\right)^2 \tag{13}$$

Applying Bernstein's inequality:

$$\mathbb{P}\left[|\widetilde{w} - w| \geq t\right] \leq 2\exp\left(-\frac{t^2}{2\mathrm{Var}[\widetilde{w}] + \frac{2}{3}Mt}\right) \tag{14}$$

where $M = \frac{(w_{\max} - w_{\min})(e^\epsilon + 1)}{2(e^\epsilon - 1)}$. Setting $t = O\left((w_{\max} - w_{\min})\log(2/\beta)/\epsilon\right)$ gives the result. □

5 Evaluation of the SynGraph Algorithm on Real-World Datasets

5.1 Experiment Setup

In this section, we further evaluate the performance of the SynGraph algorithm on six real-world road network datasets (NY, BAY, CAL, COL, LKS, and FLA) from the 9th DIMACS Implementation Challenge[1]. The number of nodes and edges for each dataset is summarized in Table 1.

Table 1: Statistics of Real-World Graph Datasets

Dataset	**Nodes**	**Edges**
NY	11,317	57,786
BAY	10,746	51,852
CAL	9,577	46,121
COL	7,566	34,670
LKS	9,317	43,165
FLA	9,763	47,608

The experiments in this section employ the following two evaluation metrics to assess the performance of the algorithm .

[1] `https://www.diag.uniroma1.it//challenge9/download.shtml`

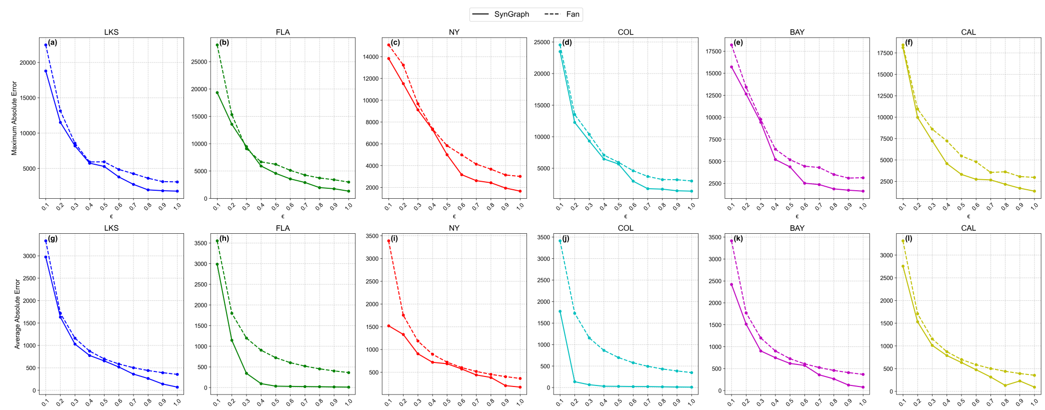

Figure 1: Error comparison between the SynGraph and Fan algorithms on six real-world graph datasets under $\delta = 10^{-5}$. (a)–(f): Maximum Absolute Error (MAE) for LKS, FLA, NY, COL, BAY, and CAL, respectively; (g)–(l): Average Absolute Error (AAE) for the same datasets.

- **Maximum Absolute Error (MAE):** Measures the maximum absolute deviation between the true and the estimated shortest path distances across all node pairs:

$$\text{MAE} = \max_{u,v \in V} \left| \text{dist}(u,v) - \widetilde{\text{dist}}(u,v) \right|. \tag{15}$$

- **Average Absolute Error (AAE):** Calculates the mean of the absolute errors over all node pairs:

$$\text{AAE} = \frac{1}{|V|^2} \sum_{u,v \in V} \left| \text{dist}(u,v) - \widetilde{\text{dist}}(u,v) \right|. \tag{16}$$

5.2 Evaluation on Real-World Graphs

To comprehensively assess the performance of the proposed algorithm, we conduct an empirical evaluation on real-world datasets, with the privacy parameter fixed at $\delta = 10^{-5}$. The state-of-the-art algorithm proposed by Fan et al. [1] is adopted as the baseline for comparison.

The second experiment evaluates the algorithms on six real-world graph datasets, with ϵ ranging from 0.1 to 1.0. Both MAE and AAE are reported, and each experiment is repeated 20 times to ensure reliability. The results are shown in Figure. 1. For both algorithms, errors decrease as ϵ increases, indicating improved accuracy with larger privacy budgets. Notably, SynGraph consistently outperforms the baseline across all datasets and privacy settings. In particular, when $\epsilon \geq 0.8$, SynGraph achieves an effective balance between privacy protection and query accuracy, highlighting its robustness and practical effectiveness.

References

[1] Chenglin Fan, Ping Li, and Xiaoyun Li. Breaking the linear error barrier in differentially private graph distance release. *arXiv preprint arXiv:2204.14247*, 2022.

[2] Jin Guo. Research on privacy-preserving federated learning schemes based on adaptive local differential privacy. Master's thesis, Xidian University, Xi'an,China, 2024.

[3] Adam Sealfon. Shortest paths and distances with differential privacy. In *Proceedings of the 35th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems*, pages 29–41, 2016.