APPENDIX A
PROOF OF THEOREM 3

Proof. Considering an arbitrary feature i with noise scale σ_i and score θ_i. To satisfy ε_i-LDP, for any pair of input value $t/k, t'/k \in \{\frac{1}{k}, \dots, 1\}$, we bound the following proportion:

$$\frac{Pr(u/k|t/k)}{Pr(u/k|t'/k)} \leq \arg\max_{u,t,t'} \frac{Pr(u/k|t/k)}{Pr(u/k|t'/k)} = \frac{Pr(1/k|1/k)}{Pr(1/k|k/k)}$$
$$= \exp[\frac{|1/k - 1/k| - |k/k - 1/k|}{\sigma_i}] \leq \exp\left(\frac{k-1}{k\sigma_i}\right) \leq e^{\varepsilon_i}$$

Therefore, we have:

$$\sigma_i \geq \frac{(k-1)}{k\varepsilon_i} \tag{17}$$

By the strong composition theorem, the total privacy budget for the whole feature vector is calculated as follows:

$$\sum_{i=1}^{d} \varepsilon_i = \varepsilon_f \sum_{i=1}^{d} \theta_i = \varepsilon_f \tag{18}$$

which completes our proof. □

APPENDIX B
PROOF OF LEMMA 4

Proof. Considering two neighbor graphs $G(\mathcal{V}, \mathcal{E}_{pub} \cup \mathcal{E}_{pri})$, $G(\mathcal{V}, \mathcal{E}_{pub} \cup \mathcal{E}'_{pri})$ which are different at one private edge. Without loss of generality, we consider $|\mathcal{E}_{pri}| = |\mathcal{E}'_{pri}| - 1$. First of all, we fix $\bar{N}_r$ and let $f(\bar{e}) = \chi(\bar{p}_r) - \chi(\bar{p}'_r)$ and $\Delta_e = \max|f(\bar{e})|$. The second order derivative of $h(\bar{p}_r)$ is given by: $\chi''(\bar{p}_r) = -\frac{1}{1-\bar{p}} - \frac{1}{\bar{p}} < 0, \forall \bar{p}$. Therefore, $\chi'(\bar{p}_r)$ is a monotonically decreasing function which implies that $f(\bar{e})$ is also a monotonically decreasing function ($\bar{p}_r > \bar{p}'_r$). With the monotonic property of $f(\bar{e})$ we can derive value of Δ_e when $\bar{e} = 1$ or $\bar{e} = \bar{N}_{max}$. Fixing $\bar{e}_r = 1$ and vary $\bar{N}_r$, therefore, $\Delta_e = \max_{\bar{N}_r} |f(\bar{N}_r)|$, where

$$f(\bar{N}_r) = 1 \log \frac{1}{\bar{N}_r} + (N-1) \log\left(1 - \frac{1}{\bar{N}_r}\right) \tag{19}$$

The first order derivative of $f(\bar{N}_r) = \log\left(1 - \frac{1}{\bar{N}_r}\right) < 0, \forall \bar{N}_r$, so that $f(\bar{N}_r)$ is monotonically decreasing by $\bar{N}_r$. Since $f(\bar{N}_r) < 0, \forall \bar{N}_r \in [1, +\infty]$, we conclude that $\Delta_e = -\min(f(\bar{N}_r)) = -f(\bar{N}_{max})$. Therefore,

$$\begin{aligned}
\Delta_e &= \log(\bar{N}_{max}) - (\bar{N}_{max} - 1) \log\left(\frac{\bar{N}_{max} - 1}{\bar{N}_{max}}\right) \\
&= \log(\bar{N}_{max}) - \log\left(\frac{\bar{N}_{max} - 1}{\bar{N}_{max}}\right)^{(\bar{N}_{max}-1)} \\
&= \log(\bar{N}_{max}) - \log\left(1 - \frac{1}{\bar{N}_{max}}\right)^{(\bar{N}_{max}-1)}
\end{aligned} \tag{20}$$

According to Cauchy's inequality, $N_{max} \leq \frac{(\bar{L}_r + \bar{R}_r)^2}{4}$ and the equality happens when $\bar{L}_r = \bar{R}_r = \frac{|\bar{V}|}{2} \rightarrow N_{max} = \frac{|\bar{V}|^2}{4}$ when $|\bar{V}|$ is even and $N_{max} = \frac{|\bar{V}|^2 - 1}{4}$ when $|\bar{V}|$ is odd. □

APPENDIX C
PRIVACY AND DATA UTILITY ANALYSIS

In this section, we focus on understanding the trade-off between privacy and data utility, regarding expected error bounds and DP-preserving graph structures, compared with baseline approaches.

a) Expected Error Bounds (EEB): To analyze data utility in FEATURERR mechanism, we study the EEB, quantified by the expected change of an embedding feature z_i after applying FEATURERR to achieve LDP-preserving $\bar{z}_i$. We consider the EEB as follows: $Err(z_i, z'_i) = \mathbb{E}\|z_i - z'_i\|_1$

We compare the EEB of FEATURERR with the state-of-the-art baselines: Duchi mechanism [29], Piecewise mechanism [30], Hybrid mechanism [30], Three-output mechanism [31], Sub-optimal mechanism (PM-SUB) [31], Laplacian and Gaussian mechanism [16], LPGNN [9]. Figure 10 illustrates the EEB of each algorithm as a function of the privacy budget ε_f.

Given a very tight privacy budget $\varepsilon_f < 0.5$, FEATURERR achieves competitive EEB with a marginal difference compared with the best baselines, i.e., the corrected LATENT and the corrected OME. When the privacy budget is higher $\varepsilon_f \in [0.5, 9]$, i.e., a wide range, our mechanism significantly outperforms other mechanisms by achieving the lowest EEB. Therefore, our mechanism can retain high data utility compared with baselines.

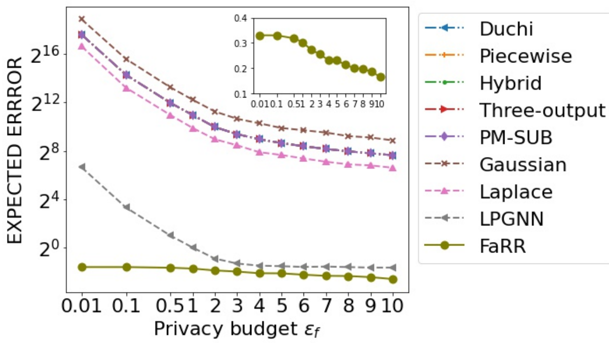

Fig. 10. Expected error bound.

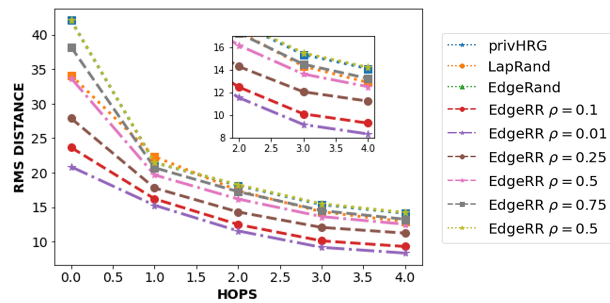

Fig. 11. RMS distance between G and $\tilde{G}$ with $\varepsilon_e = 1$.

b) Graph Structure Utility: We compare EDGERR with state-of-the-art edge-level DP-preserving approaches by measuring the Root Mean Square (RMS) distance between the legitimate graph G and the edge-level DP-preserving graph $\tilde{G}$ under a wide range of the privacy budget ε_e.

We conduct our analysis using Cora [34], a real-world dataset. Since the global sensitivity Δ_e is proportional to $\log|\bar{\mathcal{V}}|$, the utility of EDGERR depends on the proportion of the private edges in the legitimate graph G and the privacy budget ε_e. Let us denote ρ as the proportion of the private edges, i.e $\rho = \frac{|\mathcal{E}_{pri}|}{|\mathcal{E}|}$, we randomly pick a set of private edges with $\rho \in [0.01, 1.0]$. When $\rho = 1.0$, our mechanism is the same as the *privHRG* mechanism in [7], since all the edges are private. We consider the baseline: privHRG mechanism [7], LapGraph mechanism [5], and EdgeRand mechanism [5]. In practice, GNNs models usually aggregate the information from 2-hops to 3-hops of neighbors; thus, we calculate the distance of the random walk matrix up to 4-hops. We apply EDGERR on the Cora graph with different proportions of private edges; meanwhile, we apply each baseline on the original Cora graph G.

Illustrated in Figure 11, EDGERR achieves higher graph structure utility compared with the baselines. The RMS distance between the edge-level DP-preserving graph $\tilde{G}$ and the original graph G with different values of ρ is lower than the RMS distances produced by the baselines. This is reasonable since only the private edges are randomized while the public edges remain the same in EDGERR resulting in better graph structure utility compared to the baselines. Moreover, since EDGERR is optimized by leveraging the information from the public edges, the randomization process is not biased to the private subgraph G_{pri}. Also, the lower ρ is, the lower the RMS distance will be given fewer private edges to be randomized.

Appendix D
Algorithms

Algorithm 1 Feature level protection (FEATURERR)

Input: Privacy budget ε_f, number of bins k, number of features d, embedding features z, importance score vector α, sensitivity score vector β, hyper-parameter γ.
Output: Perturbed embedding features $\tilde{z}$
Normalize the values of z to $[0, 1]$
$\theta_i \leftarrow \gamma\alpha_i + (1-\gamma)\big[\beta_{min} + (\beta_{max} - \beta_i)\big], \forall i \in [d]$
$\theta_i \leftarrow \frac{\theta_i}{\sum_{j=1}^{d} \theta_j}, \forall i \in [d]$
$\varepsilon_i \leftarrow \theta_i \varepsilon_f$ and $\sigma_i = \frac{k-1}{k\varepsilon_i}, \forall i \in [d]$
for $i \in \{1, \ldots, d\}$ **do**
    for $t \in \{1, \ldots, k\}$ **do**
        if $\frac{t-1}{k} \leq z_i \leq \frac{t}{k}$: $z_i \leftarrow \frac{t}{k}$
    end for
end for
Initialize empty vector $\tilde{z}$ same shape with z.
for $i \in \{1, \ldots, d\}$ **do**
    $$\tilde{z}_i = \begin{cases} \frac{1}{k}, & \textit{with probability } Pr(\frac{1}{k}|z_i) \\ \ldots & \\ \frac{k}{k}, & \textit{with probability } Pr(\frac{k}{k}|z_i) \end{cases}$$
end for
return $\tilde{z}$

Algorithm 2 Edge level protection EDGERR

Input: Input graph $G(\mathcal{V}, \mathcal{E}_{pub} \cup \mathcal{E}_{pri})$, privacy budget ε_{e1}.
Output: randomized graph $\tilde{G}(\mathcal{V}, \mathcal{E}_{pub} \cup \tilde{\mathcal{E}}_{pri})$
Initialize the MCMC by a random dendrogram D_o;
$A \leftarrow$ the adjacency matrix of G
for each step t of MCMC **do**
    for each public step **do**
        Randomly pick an internal node r in D_{t-1}
        Sample $\mathcal{D}'$ from a possible structure of substree at r
        Accept $\mathcal{D}_t = D'$ with propability $\min\left(1, \frac{\exp \mathbb{L}_{pub}(\mathcal{D}')}{\exp \mathbb{L}_{pub}(\mathcal{D}_{t-1})}\right)$
    end for
    for each private step **do**
        Randomly pick an internal node r in D_{t-1}
        Sample $\mathcal{D}'$ from a possible structure of substree at r
        Accept $\mathcal{D}_t = D'$ with propability $\min\left(1, \frac{\exp\left(\frac{\varepsilon_{e1}}{\Delta_e}\mathbb{L}_{pri}\big(\mathcal{D}'\big)\right)}{\exp\left(\frac{\varepsilon_{e1}}{\Delta_e}\mathbb{L}_{pri}(\mathcal{D}_{t'-1})\right)}\right)$
    end for
    if converged **then**
        $\mathcal{D}* = \mathcal{D}_t$
    end if
end for
Applying **CalculateNoisyProb(**$G_{pri}(\mathcal{V}_{pri}, \mathcal{E}_{pri}), \mathcal{D}*, \varepsilon_{e2}, r_{root}$**)**
Initialize $\tilde{A}$ be zeros matrix
for each pair of $u, v \in \mathcal{V}$ **do**
    if $u \in \mathcal{V}_{pri}$ and $v \in \mathcal{V}_{pri}$ **then**
        Find the lowest common ancestor r of u and v
        $\tilde{A}_{u,v} \sim Ber(1, \tilde{p}_r)$ (Bernoulli's distribution); $\tilde{A}_{v,u} = \tilde{A}_{u,v}$
    else
        $\tilde{A}_{u,v} = A_{u,v}$; $\tilde{A}_{v,u} = \tilde{A}_{u,v}$
    end if
end for
Construct $\tilde{G}$ from $\tilde{A}$ and return $\tilde{G}$

Algorithm 3 `CalculateNoisyProb`$(G_{pri}, \mathcal{D}*, \varepsilon_{e2}, r)$ [7]

Input: Private graph $G_{pri}(\mathcal{V}_{pri}, \mathcal{E}_{pri})$, sampled dendrogram $\mathcal{D}*$, privacy budget ε_{e2}, internal node r
Output: Perturbed set $\{\tilde{p}_r\}$
$\lambda_b \leftarrow \frac{1}{\varepsilon_{e2}\bar{L}_r\bar{R}_r}$; $\lambda_c \leftarrow \frac{1}{\varepsilon_{e2}(\bar{L}_r+\bar{R}_r)(\bar{L}_r+\bar{R}_r-1)}$
if $\lambda_b \geq \tau_1$ and $\lambda_c \geq \tau_e$ **then**
    $\tilde{e}_r \leftarrow$ number of edges in the sub-graph induced from all private leave nodes of sub-tree rooted at r;
    $\tilde{p}_r \leftarrow \min\left(1, \frac{\tilde{e}_r+Lap(\frac{1}{\varepsilon_{e2}})}{(\bar{L}_r+\bar{R}_r)(\bar{L}_r+\bar{R}_r-1)/2}\right)$
    for each child internal node r' of r **do**
        $\tilde{p}'_r \leftarrow \min\left(1, \frac{\tilde{e}_r+Lap(\frac{1}{\varepsilon_{e2}})}{(\bar{L}_r+\bar{R}_r)(\bar{L}_r+\bar{R}_r-1)/2}\right)$
    end for
else
    $\tilde{p}_r \leftarrow \min\left(1, \frac{\bar{e}_r+Lap(\frac{1}{\varepsilon_{e2}})}{\bar{L}_r\bar{R}_r}\right)$
    $r_L \leftarrow$ left child of r; $r_R \leftarrow$ right child of r;
    `CalculateNoisyProb`$(G_{pri}, \mathcal{D}*, \varepsilon_{e2}, r_L)$
    `CalculateNoisyProb`$(G_{pri}, \mathcal{D}*, \varepsilon_{e2}, r_R)$
end if
return $\{\tilde{p}_r\}$

Appendix E
Supplementary results

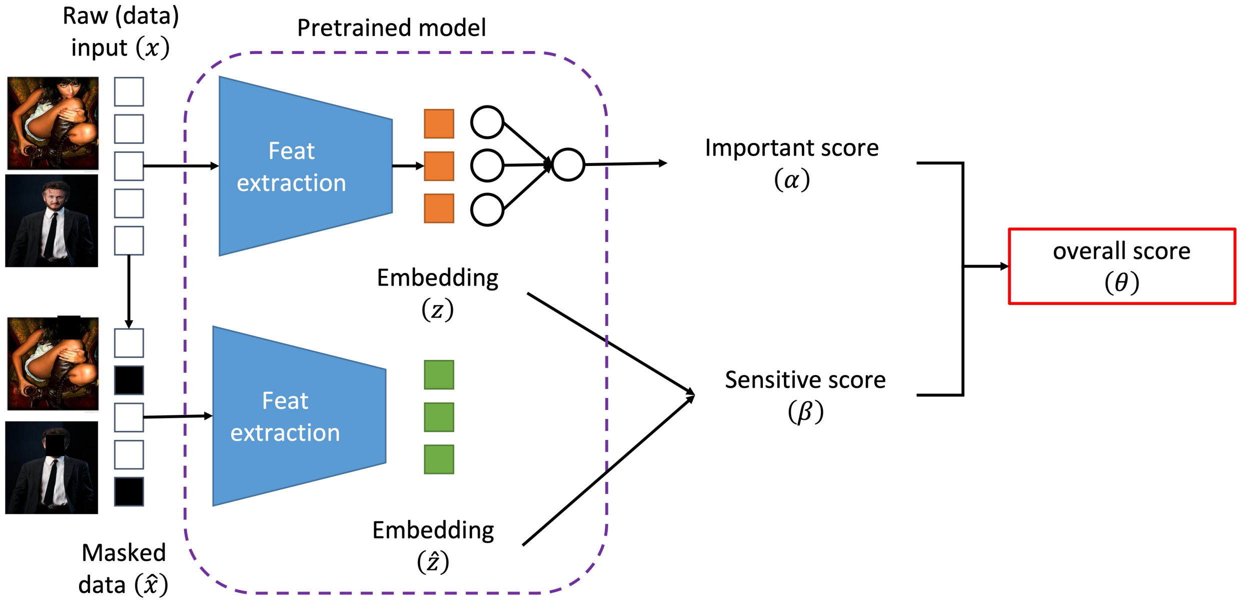

Fig. 12. An overview of feature's importance and sensitivity indication process.

Fig. 13. Additional results of model performance of the combination of feature-level and edge-level protection of the FLICKR-MIR dataset.

(a) PPI with $\rho = 0.05$ (F1-SCORE)

(b) PPI with $\rho = 0.1$ (F1-SCORE)

(c) PPI with $\rho = 0.2$ (F1-SCORE)

Fig. 14. Additional results of model performance of the combination of feature-level and edge-level protection for the PPI dataset.