Appendix A

Lemma 1 *The dissimilarity between the distribution of original sensitive features and the distribution of k-hop neighbor sensitive features exhibits a pronounced correlation with the eigenvector corresponding to the largest magnitude eigenvalue of the adjacency matrix. Concurrently, the correlation with other eigenvectors diminishes exponentially.*

Proof. Assume $\mathbf{A} \in \mathbb{R}^{n \times n}$ is a adjacency matrix with real-valued entries. The k-hop neighbor feature matrix is $\mathbf{A}^k\mathbf{H}$, and the k-hop neighbor sensitive features are $\mathbf{A}^k\mathbf{H}[:, s]$. $|\lambda_1| > |\lambda_2| \geq ...... \geq |\lambda_n|$ are n real eigenvalues, and $\mathbf{p}_i$ ($i \in \{1, 2, ..., n\}$) are corresponding eigenvectors. The eigendecomposition of $\mathbf{A}$ can be written as $\mathbf{A} = \mathbf{P}\Lambda\mathbf{P}^\top$ with $\mathbf{P} = (\mathbf{p}_1, \mathbf{p}_2, ..., \mathbf{p}_n)$, $\|\mathbf{p}_i\| = 1$ ($i \in \{1, 2, ..., n\}$) and $\Lambda = diag(\lambda_1, \lambda_2, ..., \lambda_n)$: We assume $\alpha_i = \mathbf{H}[:, s]^\top \mathbf{p}_i$.

We use cosine similarity to measure the dissimilarity.

$$
\begin{aligned}
& cos(<\mathbf{p}_i, \mathbf{H}[:, s]>) \\
=& \frac{\mathbf{H}[:, s]^\top \mathbf{p}_i}{\|\mathbf{H}[:, s]\| \|\mathbf{p}_i\|} \\
=& \frac{\mathbf{H}[:, s]^\top \mathbf{p}_i}{\|\mathbf{H}[:, s]\|} \\
=& \frac{\mathbf{H}[:, s]^\top \mathbf{p}_i}{\sqrt{\mathbf{H}[:, s]^\top \mathbf{H}[:, s]}} \\
=& \frac{\mathbf{H}[:, s]^\top \mathbf{p}_i}{\sqrt{(\mathbf{P}^\top \mathbf{H}[:, s])^\top \mathbf{P}^\top \mathbf{H}[:, s]}} \\
=& \frac{\mathbf{H}[:, s]^\top \mathbf{p}_i}{\sqrt{\sum_{j=1}^{n} (\mathbf{H}[:, s]^\top \mathbf{p}_j)^2}} \\
=& \frac{\alpha_i}{\sqrt{\sum_{j=1}^{n} \alpha_j^2}}.
\end{aligned}
$$

Then,

$$
\begin{aligned}
& cos(<\mathbf{A}^k\mathbf{H}[:, s], \mathbf{H}[:, s]>) \\
=& \frac{(\mathbf{A}^k\mathbf{H}[:, s])^\top \mathbf{H}[:, s]}{\|\mathbf{A}^k\mathbf{H}[:, s]\| \|\mathbf{H}[:, s]\|} \\
=& \frac{(\mathbf{A}^k\mathbf{H}[:, s])^\top \mathbf{H}[:, s]}{\sqrt{(\mathbf{A}^k\mathbf{H}[:, s])^\top \mathbf{A}^k\mathbf{H}[:, s]} \sqrt{\mathbf{H}[:, s]^\top \mathbf{H}[:, s]}} \\
=& \frac{(\mathbf{P}\Lambda^k\mathbf{P}^\top\mathbf{H}[:, s])^\top \mathbf{H}[:, s]}{\sqrt{(\mathbf{P}\Lambda^k\mathbf{P}^\top\mathbf{H}[:, s])^\top (\mathbf{P}\Lambda^k\mathbf{P}^\top\mathbf{H}[:, s])} \sqrt{\mathbf{H}[:, s]^\top \mathbf{H}[:, s]}} \\
=& \frac{(\mathbf{P}^\top\mathbf{H}[:, s])^\top \Lambda^k (\mathbf{P}^\top\mathbf{H}[:, s])}{\sqrt{(\mathbf{P}^\top\mathbf{H}[:, s])^\top \Lambda^{2k} (\mathbf{P}^\top\mathbf{H}[:, s])} \sqrt{\mathbf{H}[:, s]^\top \mathbf{H}[:, s]}} \\
=& \frac{\sum_{i=1}^{n} \alpha_i^2 \lambda_i^k}{\sqrt{\sum_{i=1}^{n} \alpha_i^2 \lambda_i^{2k}} \sqrt{\sum_{i=1}^{n} \alpha_i^2}} \\
=& \frac{\alpha_1^2 + \sum_{i=2}^{n} \alpha_i^2 (\frac{\lambda_i}{\lambda_1})^k}{\sqrt{\alpha_1^2 + \sum_{i=2}^{n} \alpha_i^2 (\frac{\lambda_i}{\lambda_1})^{2k}} \sqrt{\sum_{i=1}^{n} \alpha_i^2}}.
\end{aligned}
$$

Thus, the correlation between $\mathbf{p}_i$ and $cos(<\mathbf{A}^k\mathbf{H}[:, s], \mathbf{H}[:, s]>)$ is proportional to $(\frac{\lambda_i}{\lambda_1})^k$. Because of $|\lambda_1| > |\lambda_i|$, the correlation decays exponentially.

Specifically, when $k \to \infty$, we have:

$$
\begin{aligned}
& \lim_{k \to \infty} cos(<\mathbf{A}^k\mathbf{H}[:, s], \mathbf{H}[:, s]>) \\
=& \lim_{k \to \infty} \frac{\alpha_1^2 + \sum_{i=2}^{n} \alpha_i^2 (\frac{\lambda_i}{\lambda_1})^k}{\sqrt{\alpha_1^2 + \sum_{i=2}^{n} \alpha_i^2 (\frac{\lambda_i}{\lambda_1})^{2k}} \sqrt{\sum_{i=1}^{n} \alpha_i^2}} \\
=& \frac{\alpha_1}{\sqrt{\sum_{i=1}^{n} \alpha_i^2}} \\
=& cos(<\mathbf{p}_1, \mathbf{H}[:, s]>).
\end{aligned}
$$

□

Appendix B

Lemma 2 *The sensitive feature distribution of $\mathbf{H}^{(k)}$ is the same as $\mathbf{H}$:*

$$\mathbf{H}^{(k)}[i, s] = q^k \mathbf{H}[i, s], \tag{1}$$

where $i \in \{1, 2, ..., n\}$, q denotes the number of nodes whose sensitive feature is 1.

Proof. The based case ($k = 1$): $\mathbf{H}^{(1)} = \mathbf{A}_s\mathbf{H}$ and $\mathbf{H}[i, s] \in \{0, 1\}$.

Because

$$\mathbf{H}^{(1)}[i, s] = \mathbf{A}_s[i, :]\mathbf{H}[:, s].$$

$$\mathbf{A}_s[i, j] = \begin{cases} 1, \mathbf{H}[i, s] = \mathbf{H}[j, s], \\ 0, \mathbf{H}[i, s] \neq \mathbf{H}[j, s]. \end{cases}$$

$$\mathbf{A}_s[i, j]\mathbf{H}[j, s] = \begin{cases} 0, \mathbf{H}[i, s] = 0, \\ 0, \mathbf{H}[i, s] = 1, \mathbf{H}[j, s] = 0, \\ 1, \mathbf{H}[i, s] = \mathbf{H}[j, s] = 1. \end{cases}$$

Thus,

$$\mathbf{H}^{(1)}[i, s] = \begin{cases} 0, \mathbf{H}[i, s] = 0, \\ p, \mathbf{H}[i, s] = 1. \end{cases}$$

The conclusion is valid.

Inductive hypothesis ($k = r$): Assume that the conclusion holds when $k = r$, thus $\mathbf{H}^{(r)}[i, s] = q^r \mathbf{H}[i, s]$. Thus,

$$\mathbf{H}^{(1)}[i, s] = \begin{cases} 0, \mathbf{H}^{(r-1)}[i, s] = 0, \\ q^r, \mathbf{H}^{(r-1)}[i, s] = q^{r-1}. \end{cases}$$

Inductive Step ($k = r + 1$):

Because $\mathbf{H}^{(r+1)}[i, s] = \mathbf{A}_s[i, :]\mathbf{H}^{(r)}[:, s]$.

$$\mathbf{A}_s[i, j]\mathbf{H}^{(r+1)}[j, s] = \begin{cases} 0, & \mathbf{H}^{(q)}[i, s] = 0, \\ 0, & \mathbf{H}^{(q)}[i, s] = q^r, \\ & \mathbf{H}^{(q)}[j, s] = 0, \\ q^r, & \mathbf{H}^{(q)}[i, s] = q^r, \\ & \mathbf{H}^{(q)}[j, s] = q^r. \end{cases}$$

Thus,

$$\mathbf{H}^{(r+1)}[i, s] = \begin{cases} 0, \mathbf{H}^{(r)}[i, s] = 0, \\ q^{(r+1)}, \mathbf{H}^{(r)}[i, s] = q^r. \end{cases}$$

The conclusion is confirmed in the inductive step.

Combining the foundational step, the inductive hypothesis, and the inductive step, we show that this mathematical statement holds for all positive integer k. □

Appendix C

More details of experimental results.

Methods	NBA		Bail		German		Credit		Income	
	F1 ↑	AUC ↑	F1 ↑	AUC ↑	F1 ↑	AUC ↑	F1 ↑	AUC ↑	F1 ↑	AUC ↑
GCN	74.42	71.10	79.15	87.43	81.61	62.81	78.83	69.45	51.02	**77.33**
GCNII	76.09	77.63	89.43	94.13	82.97	53.16	84.46	61.22	41.71	67.94
GAT	75.00	73.97	90.65	96.37	82.74	57.55	76.40	64.86	35.77	61.21
FairGNN	69.48	77.84	77.50	87.36	82.01	67.35	77.79	**69.73**	50.93	77.24
NIFTY	71.91	68.85	76.40	78.20	82.75	68.78	84.08	68.30	41.27	70.21
BIND	67.41	62.82	85.62	93.93	81.93	71.88	78.04	69.54	50.55	76.89
GTrans	76.92	74.81	91.56	97.32	84.24	75.54	80.75	68.13	38.90	63.60
SAN	73.42	75.51	93.07	97.84	83.37	76.79	80.46	63.30	32.09	63.29
NAG	76.09	78.91	90.67	97.46	84.42	71.44	86.91	63.30	40.20	63.60
FairGT	**78.16**	**81.22**	**94.46**	**99.07**	**84.54**	**78.32**	**87.19**	69.38	**51.46**	76.01

Table 1: Comparison of performance (F1 and AUC) in percentage (%). ↑ denotes the larger, the better. The best results are bold-faced.