Cross-Network Social User Embedding with Hybrid Differential Privacy Guarantees

Jiaqian Ren[1,2], Lei Jiang[1], Hao Peng[3], Lingjuan Lyu[4], Zhiwei Liu[5], Chaochao Chen[6], Jia Wu[7], Xu Bai[1], Philip S. Yu[8]

[1] Institute of Information Engineering, Chinese Academy of Sciences;
[2]School of Cyber Security, University of Chinese Academy of Sciences; [3]Beihang University;
[4]Sony AI; [5]Salesforce; [6]Zhejiang University; [7]Macquarie University; [8]University of Illinois Chicago.
* Corresponding authors (Jianglei@iie.ac.cn, penghao@buaa.edu.cn)

ABSTRACT

Integrating multiple online social networks (OSNs) has important implications for many downstream social mining tasks, such as user preference modelling, recommendation, and link prediction. However, it is unfortunately accompanied by growing privacy concerns about leaking sensitive user information. How to fully utilize the data from different online social networks while preserving user privacy remains largely unsolved. To this end, we propose a Cross-network Social User Embedding framework, namely DP-CroSUE, to learn the comprehensive representations of users in a privacy-preserving way. We jointly consider information from partially aligned social networks with differential privacy guarantees. In particular, for each heterogeneous social network, we first introduce a hybrid differential privacy notion to capture the variation of privacy expectations for heterogeneous data types. Next, to find user linkages across social networks, we make unsupervised user embedding-based alignment in which the user embeddings are achieved by the heterogeneous network embedding technology. To further enhance user embeddings, a novel cross-network GCN embedding model is designed to transfer knowledge across networks through those aligned users. Extensive experiments on three real-world datasets demonstrate that our approach makes a significant improvement on user interest prediction tasks as well as defending user attribute inference attacks from embedding.

CCS CONCEPTS

• **Computer systems organization → Embedded systems**; • **Security and privacy → Social network security and privacy**.

KEYWORDS

Network Integration, Differential Privacy, User Linkage, Representation Learning.

ACM Reference Format:
Jiaqian Ren[1,2], Lei Jiang[1], Hao Peng[3], Lingjuan Lyu[4], Zhiwei Liu[5], Chaochao Chen[6], Jia Wu[7], Xu Bai[1], Philip S. Yu[8] . 2022. Cross-Network Social User Embedding with Hybrid Differential Privacy Guarantees. In *Proceedings of the 31st ACM International Conference on Information and Knowledge Management (CIKM'22).* ACM, New York, NY, USA, 11 pages. https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Social media-based user embedding plays an important role in user representation, user analysis and many downstream applications. Nowadays, to incorporate more information and get enhanced user embeddings, new technologies [10, 26, 55, 56, 60, 61] which fuse and mine multiple social networks together show promising trends. However, this trend is now challenged by serious privacy concerns. Authors in [20] report that more than 80% US Internet users were worried about the usage of their personal data. Meanwhile, more rigorous regulations like EU's GDPR[1] are enacted to regulate the usage of personal information. For example, companies cannot share a user's data without his/her consent. Henceforth, raw social networks which encode individual's sensitive information (e.g., friendship, gender, occupation) should not be disclosed to others directly. This paper initiates the study on privacy-preserving social user embeddings across multiple online social networks (OSNs).

A toy example of the research problem is shown in Figure 1, which involves two different social networks (i.e., Foursquare and Twitter). Specifically, user Bob participates in both networks, and tends to post food on Foursquare and share his other daily life on Twitter. Only when these two networks are combined together can we capture his interests more comprehensively. Meanwhile, the privacy concerns lead to the user data protection by perturbing some of it to fake before data sharing. By taking apart the problem, we should 1) protect the privacy of social networks and 2) integrate social networks to conduct prediction.

Nevertheless, the real-world scenarios are complex because OSNs contain different types of information, which are formulated as heterogeneous social networks (HSNs). To protect privacy, existing works use 1) various anonymization techniques and 2) differential privacy (DP) mechanisms. Since anonymization techniques, including k-anonymity [48], l-diversity [34], t-closeness [22], etc, may be vulnerable to deanonymization attacks [37] and lacking a rigorous theoretical guarantee, we adopt DP in our work. As a strong and mathematically rigorous privacy-preserving framework, DP has been widely used to release social graphs [17, 19, 29, 31, 39, 41, 46, 51–53, 57]. Among them, the standard techniques [29, 31, 39, 46, 51, 53, 57] only consider to sanitize graph structure and

CIKM'22, October 17-22, 2022, Hybrid Conference, Hosted in Atlanta, Georgia, USA
© 2022 Association for Computing Machinery.
ACM ISBN 978-1-4503-XXXX-X/18/06.
https://doi.org/XXXXXXX.XXXXXXX

[1] https://gdpr-info.eu

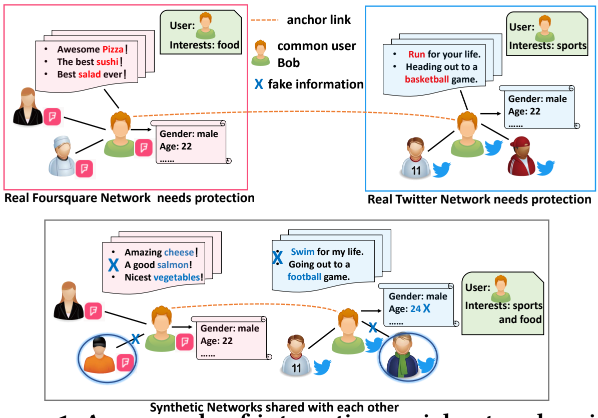


Figure 1: An example of integrating social networks with privacy protection.

ignore vertex attributes. The follow-up works [17, 19, 52] make up for this deficiency and study releasing attributed graphs. However, they are limited to homogeneous graphs. Considering real-world social networks contain multiple types of information, and different types of information have different privacy protection expectations, how to publish HSNs with proper differential privacy guarantees remains an open challenge. In this work, we propose a hybrid DP mechanism to handle each kind of data independently.

Suppose we have multiple protected heterogeneous social networks under differential privacy protection, another challenging problem is how to integrate them together to generate more comprehensive user representations. Obviously, two key issues in network fusion are user linkage and cross-network information transferring. Prior user linkage methods can be divided into 1) user attribute-based ones and 2) embedding-based ones. Attribute-based methods such as [2, 5, 8, 58] aim to link the same user across different OSNs through the comparison of real user information. Noted that the perturbed social networks do not hold accurate information anymore. Attribute-based methods cannot work. Embedding-based alignment methods [9, 21, 23, 27, 35] have gained lots of attention in recent years. It is worth noting that those perturbed networks which hold fake information still preserve important characteristics of the original ones. As embedding-based methods focus primarily on learning representations capturing essential characteristics, they can still learn useful knowledge and make the alignment. Thus, we adopt embedding-based methods in our framework to find user linkages. To further leverage multi-source data for improving cross-network analysis, a series of works [7, 26, 30, 54–56, 60, 61] have made great success in applications such as user profile modelling, social recommendation and so on. However, as mentioned before, most of them have ignored the privacy leakage problem. To sum up, how to support social network integration with proper user privacy protection is an important yet unsolved problem.

To achieve privacy-preserving social network integration for improving social user embeddings, we propose a novel framework, called DP-CroSUE. Particularly, to protect user privacy, we perturb each heterogeneous social network before data exchange. There are various sensitive data types in the social network, including graph topological data (user friendship), multi-dimensional numerical and categorical data (user attributes like age and gender), and text data (the posts), where each data has different privacy expectations. To address this issue, we introduce a hybrid differential privacy mechanism to add more proper perturbations to ensure data utility while preserving necessary privacy. To find user linkages across networks for further integration, we apply generative adversarial networks (GANs) to make unsupervised user alignment. Finally, a novel cross-network GCN embedding model including inter-graph propagation and hierarchy intra-graph propagation is proposed to transfer information both inside and across networks. We evaluate DP-CroSUE on three real-world social network platforms considering the embedding usefulness for user interest prediction tasks and its ability to resist two user attribute inference attacks: gender inference and occupation inference. Noted that the DP-CroSUE framework can also be generalized to other tasks like recommendation. We mainly focus on user interest prediction tasks for illustration purpose in this paper. Experimental results indicate that DP-CroSUE makes a good balance in both user feature utility and user privacy protection. The source code and data are available at GitHub[2].

Our contributions can be summarized as follows: 1) We propose DP-CroSUE, the first attempt to integrate multiple HSNs for comprehensive social user embeddings with a hybrid differential privacy guarantee. Our approach demonstrates a competitive tradeoff between user feature utility and user privacy protection. 2) We introduce a hybrid differential privacy mechanism capturing the variation of privacy expectations for heterogeneous social graphs. 3) We propose a novel cross-network GCN embedding model including inter-graph propagation and hierarchy intra-graph propagation to transfer information both inside and across social networks to make complete information integration.

2 RELATED WORK

Differential Privacy. DP [13] has been widely used for privacy-preserving statistical analysis. The intuition behind it is to randomise the output to ensure that the presence of any individual in the input has a negligible impact on the probability of any particular output. Earlier DP mechanisms such as Laplace mechanism and Exponential mechanism are proposed to protect single numerical [11, 13] and categorical [14, 50] data. Whereas recently, mechanisms have been developed for different data types and domains. For example, authors in [49] extend original methods to handle multi-dimensional data. In the NLP domain, a few works directly inject high-dimensional DP noise into text representations [15, 16, 32, 33]. However, due to "the curse of dimensionality", they fail to strike a nice privacy-utility balance. Authors in [59] solve this problem by sampling a close word substitute to ensure utility. In the social network domain, a series of works [31, 46, 51, 53, 57] have been proposed to protect graph structure. They focus on edge-DP and protect graph topologies only. The follow-up works [17, 19] cover this shortage by taking users' attributes into account. However, they only consider the homogeneous networks. How to perturb heterogeneous social networks is not covered in the literature.

[2] https://github.com/RingBDStack/DP-CroSUE

User Linkage and Social Network Integration. User linkage [38], also known as social network alignment, has been an important issue to make further utilization of social network data. Existing works [2, 5, 8] make alignments by carefully comparing the user attributes. These solutions are now challenged by privacy concerns about the disclosure of sensitive user attributes. Recently, authors in [58] make the first attempt to study privacy-preserving user linkage across multiple OSNs. However, it still needs some user "volunteers" whose linkages are known in advance. In recent years, embedding-based alignment methods [9, 21, 23, 27, 35] have achieved great success in finding anchor users across two or more social networks, which gives opportunities to make social network integration. Meanwhile, the cross-network results have been proved to enhance various social network applications. For example, some works [6, 18, 25, 44] fuse different social networks to provide a better understanding of users' interests and behaviours.

3 TERMINOLOGY DEFINITION AND PROBLEM FORMULATION

3.1 Definitions

Differential Privacy [13] has emerged as a strong privacy definition for statistical data release with the intuition that a randomized algorithm behaves similarly on neighbouring datasets.

Definition 3.1. (ϵ-DP [12]) *For any $\epsilon > 0$, a randomized mechanism M satisfies ϵ-DP if for any two inputs x, x' in the domain of M, and for any output y of M, we have:*

$$\frac{\Pr\{M(x) = y\}}{\Pr\{M(x') = y\}} \leq \exp(\epsilon), \tag{1}$$

where $Pr\{\cdot\}$ represents probability, ϵ corresponds to privacy budget. Smaller ϵ asserts a better privacy protection but lower data utility.

Noted that early DP is oriented toward structured data. For the protection of graph data, the notion of ϵ-edge-DP is proposed.

Definition 3.2. (ϵ-Edge-DP [3]) *For any $\epsilon > 0$, a randomized mechanism M satisfies ϵ-Edge-DP if for any two neighbouring graphs $G_1, G_2 \in \mathcal{G}$, which differ by at most one edge, and for any output S of range(M), we have:*

$$\frac{\Pr\{M(G_1) = S\}}{\Pr\{M(G_2) = S\}} \leq \exp(\epsilon). \tag{2}$$

Being a very strong privacy notion, ϵ-DP, however, is unsuitable to protect text privacy with utility [59]. Because under ϵ-DP protection, a word will be transformed to any other words with equal probabilities, no matter how unrelated they are. Thus a sanitized token may not capture the semantics. The relaxed notion of Metric DP (MDP) can address this problem.

Definition 3.3. (MDP [1]) *Given $\epsilon > 0$ and a distance metric d, a randomized mechanism M satisfies MDP if for two inputs x, x' in the domain of M, and for any output y of M, we have:*

$$\frac{Pr\{M(x) = y\}}{Pr\{M(x') = y\}} \leq \exp(\epsilon \cdot d(x, x')). \tag{3}$$

For MDP, the indistinguishability of output distributions is further controlled by the corresponding distance between the inputs, and the metric d needs to be defined according to the specific application. Considering different characteristics and privacy expectations of heterogeneous social networks, we introduce a novel hybrid-DP notion which preserves various graph properties through carefully designing the injected noise.

Definition 3.4. (Hybrid-DP) *Given $\epsilon_a > 0$ for attribute feature, $\epsilon_g > 0$ for graph edge, $\epsilon_t > 0$ for textual data, and a distance metric d, suppose there are two neighbouring heterogeneous social graphs G and G' which differ in one user node's attribute vector, the presence of a single edge and one post, a randomized mechanism M satisfies hybrid-DP if for any output O of M, we have:*

$$\frac{Pr\{M(G) = O\}}{Pr\{M(G') = O\}} \leq exp(\epsilon_a + \epsilon_g + \epsilon_t \cdot d(G_t, G'_t)), \tag{4}$$

where G_t and G'_t represent the textual data (posts) in the graphs.

3.2 Problem Formulation

Generally, each online social network can be represented as a heterogeneous graph. In particular, as shown in Figure 2, we map every single network to a heterogeneous network $G = (V, E, R)$ containing two types of nodes: (i) users v_u; (ii) textual posts v_t written by users, and two types of edge relationships: (i) friendship r_f (user-user) and (ii) writing r_w (user-post). The user-type nodes have their respective multi-dimensional user attribute features $\mathbf{X}$.

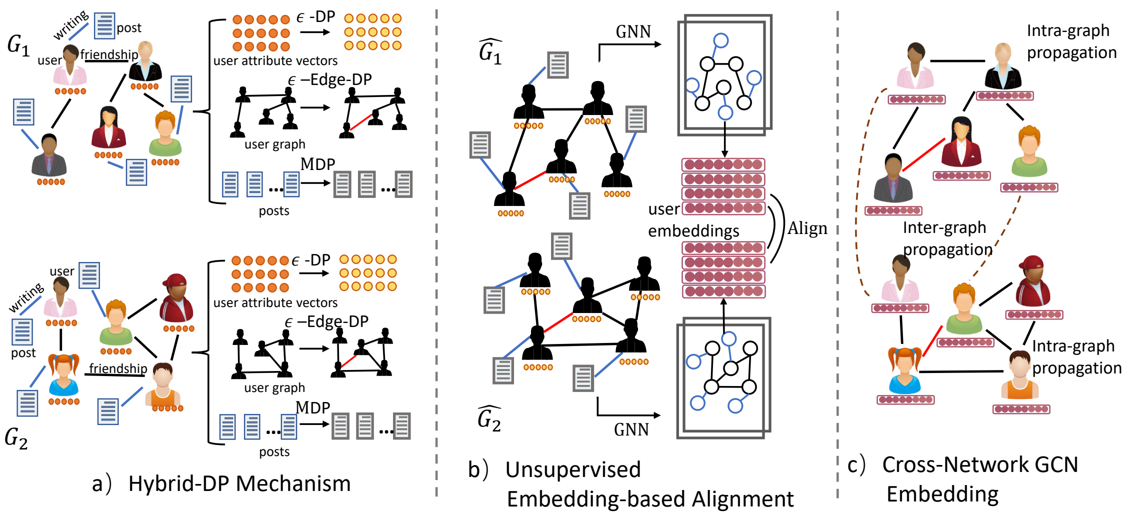
The problem of DP-CroSUE is to obtain more comprehensive representations of social users by combining multiple HSNs together without raw data leakage. Formally, this problem can be divided as follows: **1) Heterogeneous social graph protection with hybrid differential privacy guarantees**: Given a social network company's data which can be represented as a heterogeneous network G, we need to design a hybrid randomized mechanism M_h: $M_h(G) \longrightarrow \hat{G}$ by fully considering the different characteristics and privacy expectations of different data types. That is to say, for each data type, we should adopt the proper DP notion and set a proper privacy budget. **2) Cross-network information transferring**: Given two protected social networks $\hat{G}_1$ and $\hat{G}_2$, to transfer knowledge across these two networks, a set of anchor users need to be found. Next, with those anchor pairs working as a "bridge", we need to design a transferring architecture which propagates information across and within these two networks effectively.

4 MODEL

In this section, we describe DP-CroSUE in detail. The whole framework is shown in Figure 2.

4.1 Hybrid-DP Mechanism

As mentioned in Section 3.2, we consider privacy protection of three types of sensitive information in heterogeneous social networks - user attribute features, user friendship relationships and posts written by users. They correspond to three data formats respectively: multidimensional numerical and categorical data, graph edge data and textual data. Given the different characteristics of different data formats and users' heterogeneous privacy expectations, treating all the data types as equally sensitive will add too much unneeded noise and sacrifice utility. For example, the strict $\epsilon - DP$ notion applied in text data will totally change the original semantics and result in low utility. Besides, the contributions of a user's attribute



a) Hybrid-DP Mechanism

b) Unsupervised Embedding-based Alignment

c) Cross-Network GCN Embedding

Figure 2: An overview of our proposed DP-CroSUE framework, which contains three important steps. First, to share OSNs with each other, each raw heterogeneous social network is perturbed by the hybrid differential privacy mechanism (Section 4.1). Second, we utilize heterogeneous network embedding technology to get initial user embeddings and make unsupervised network alignment to find anchor users (Section 4.2). Finally, a novel cross-network GCN embedding model is proposed to get enhanced social user representations by integrating knowledge across networks (Section 4.3).

information, the numerous posts he/she has posted and the friendship relations to the final prediction tasks and information leakage differ a lot. Thus, a hybrid differential privacy mechanism which carefully designs the injected noise is highly needed. Specifically, we adopt ϵ-DP, ϵ-Edge-DP and MDP for attribute data, graph edge data and textual data respectively. Meanwhile, we also give insights into allocating proper privacy budgets.

We first introduce the three strategies for different data formats. To protect user attributes, we directly inject noise to the user attribute vector containing both numerical and categorical features. The extended PM algorithm proposed in [62] is adopted. According to the proof in [62], it satisfies ϵ-DP. For the protection of user friendship relations, we extract the homogeneous user graph and enforce edge-DP to it. Specifically, we adopt TmF [40] algorithm, which first computes a new, noisy number of graph edges, then utilizes a filter to decide whether the original edge should be preserved or not. As proved in [40], TmF satisfies ϵ-Edge-DP.

To protect text with utility, we develop a sanitation mechanism with a MDP guarantee. First, we inject each word in the corpus C to an embedding. We denote the injection function ϕ, which can be any of the well-known word embedding algorithms (e.g., Word2Vec [36], GloVe [42], or FastText [4]). Here we select Word2Vec[36]. For any two words x and x', we define their distance $d(x,x') = d_{euc}(\phi(x),\phi(x'))$, where d_{euc} represents the Euclidean distance. To achieve MDP, for each word x, we run the randomized mechanism M to sample a sanitized y with probability:

$$Pr\{M(x)=y\} = \frac{exp(-\frac{1}{2}\epsilon \cdot d_{euc}(\phi(x),\phi(y)))}{\sum_{y'\in C} exp(-\frac{1}{2}\epsilon \cdot d_{euc}(\phi(x),\phi(y')))}. \quad (5)$$

Theorem 4.1. *Given $\epsilon > 0$ and a distance metric d_{euc}, the randomized mechanism M depicted in Eq. 5 satisfies MDP.*

Proof of Theorem 4.1. Consider a sentence D only having one word <x>, another sentence D' which is <x'> ($x' \neq x$), and a possible output y. We set $C_x = (\sum_{y'\in C} exp(-\frac{1}{2}\epsilon \cdot d_{euc}(\phi(x),\phi(y'))))^{-1}$.

$$\begin{aligned}\frac{Pr\{M(x)=y\}}{Pr\{M(x')=y\}} &= \frac{C_x \cdot exp(-\frac{1}{2}\epsilon \cdot d_{euc}(\phi(x),\phi(y)))}{C_{x'} \cdot exp(-\frac{1}{2}\epsilon \cdot d_{euc}(\phi(x'),\phi(y)))} \\ &= \frac{C_x}{C_{x'}} \cdot exp(\frac{1}{2}\epsilon \cdot [d(x',y) - d(x,y)]) \\ &\leq \frac{C_x}{C_{x'}} \cdot exp(\frac{1}{2}\epsilon \cdot d(x,x')) \quad (6) \\ \leq \frac{\sum_{y'\in C} exp(-\frac{1}{2}\epsilon \cdot d_{euc}(\phi(x),\phi(y')))}{\sum_{y'\in C} exp(-\frac{1}{2}\epsilon \cdot d_{euc}(\phi(x'),\phi(y')))} &\cdot exp(\frac{1}{2}\epsilon \cdot d(x,x')) \\ \leq exp(\frac{1}{2}\epsilon \cdot d(x,x')) \cdot exp(\frac{1}{2}\epsilon \cdot d(x,x')) &= exp(\epsilon \cdot d(x,x')).\end{aligned}$$

We now analyze the overall differential privacy guarantee by combining all the above three information perturbations. We have the following theorem:

Theorem 4.2. *Assume that we independently adopt the three perturbation algorithms described above and the attribute feature, graph edge, and textual data satisfy ϵ_a-DP, ϵ_g-Edge-DP, and ϵ_t-MDP, respectively. This hybrid perturbation mechanism satisfies our hybrid-DP notion defined in Definition 3.4.*

Proof of Theorem 4.2. We assume the three kinds of data in the heterogeneous graph G are independent and denote the three perturbation mechanisms for user attributes, graph structure and user posts as M_a, M_g, M_t, respectively. Meanwhile, we utilize G_a, G_g and G_t to represent the attribute features, graph structure and posts alone. G'_a, G'_g and G'_t are the corresponding neighbouring datasets for each data type. Noted that the hybrid-DP mechanism M_h is the combination of the three randomized mechanisms mentioned above and G' denotes the whole perturbed heterogeneous social network. Since the attribute part, graph edge part and textual part satisfies ϵ-DP, ϵ-Edge-DP and MDP, respectively, we have:

$$\begin{aligned}\frac{\Pr\{M_h(G)=y\}}{\Pr\{M_h(G')=y\}} &= \frac{\Pr\{M_a(G_a)=y_a, M_g(G_g)=y_g, M_t(G_t)=y_t\}}{\Pr\{M_a(G'_a)=y_a, M_g(G'_g)=y_g, M_t(G'_t)=y_t\}} \\ &= \frac{\Pr\{M_a(G_a)=y_a\}}{\Pr\{M_a(G'_a)=y_a\}} \cdot \frac{\Pr\{M_g(G_g)=y_g\}}{\Pr\{M_g(G'_g)=y_g\}} \cdot \frac{\Pr\{M_t(G_t)=y_t\}}{\Pr\{M_t(G'_t)=y_t\}} \\ &\le \exp(\epsilon_a) \cdot \exp(\epsilon_g) \cdot \exp(\epsilon_t \cdot d(G_t, G'_t)) \\ &= \exp(\epsilon_a + \epsilon_g + \epsilon_t \cdot d(G_t, G'_t)).\end{aligned} \tag{7}$$

Note that these three perturbation strategies are designed according to the characteristics of different data formats. Additionally, how to set privacy budgets (ϵ_a, ϵ_g and ϵ_t) to achieve proper privacy levels as well as utility for heterogeneous graph properties is also challenging. We solve this by referring to the Task-relevance to Message-inference Ratio (TMR). The intuition behind is that less noise being injected into the extracted data type which is more relevant to the target task will bring out more utility. Meanwhile, more noise should be injected into those data types causing accurate personal message inference to satisfy privacy. In this work, we leverage the precision score obtained through a single piece of information in the interest prediction task to measure task relevance and use the sum of precision scores in inference attacks as the message-inference value. Since larger privacy budget means less injected noise, we set larger ϵ for data with high TMR values. More details can be seen in Section 5.2.

4.2 Embedding-based Social User Alignment

To find anchor user linkages, we leverage embedding-based alignment methods, which can be divided into two parts: heterogeneous social network embedding and unsupervised user linkage.

To fully encode all kinds of information in the perturbed heterogeneous social network $\hat{G} = (V, \hat{E}, R)$, we adopt relation-specific transformations and the propagation function is:

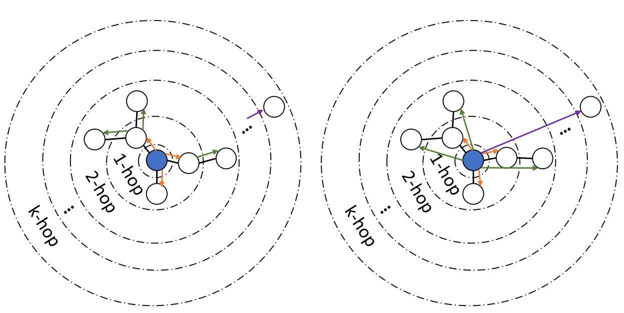
$$\mathbf{h}_i^{(l+1)} = \sigma\left(\sum_{r\in R}\sum_{j\in\mathcal{N}_i^r} \frac{1}{c_{i,r}} \mathbf{W}_r^{(l)}\mathbf{h}_j^{(l)} + \mathbf{W}_0^{(l)}\mathbf{h}_j^{(l)}\right), \tag{8}$$

where $\mathbf{h}_i^{l+1}$ is the updated representation of node v_i in the $(l+1)$-th layer. $\mathcal{N}_i^r$ denotes the set of neighbor indices of node i under relation $r \in R$ and $c_{(i,r)} = |\mathcal{N}_i^r|$. Suppose the final user embedding is $\mathbf{z}_{v_u}$ ($v_u \in \mathcal{V}_u$), the loss function can be expressed as:

$$L_{\hat{\mathcal{G}}} = -\log\left(\sigma(\mathbf{z}_{v_u}^\top \mathbf{z}_{v'_u})\right) - Q \cdot \mathbb{E}_{v_n \sim P_n(V_u)} \log\left(\sigma(\mathbf{z}_{v_u}^\top \mathbf{z}_{v_n})\right), \tag{9}$$

where v'_u is the one-hop neighbours of v_u. $P_n(V_u)$ defines negative sampling distribution of user nodes. Q is the number of negative samples.

Next, we leverage the obtained social user embeddings $\mathbf{Z}_{u1}$ and $\mathbf{Z}_{u2}$ to make the alignment. As the spaces of $\mathbf{Z}_{u1}$ and $\mathbf{Z}_{u2}$ are learnt independently, we need to learn the matrix $\mathbf{W}$ such that $\mathbf{W} = argmin\,||\mathbf{W}\mathbf{Z}_{u1} - \mathbf{Z}_{u2}||$ to reconcile them. Here the source network is $\hat{G}_1$ and the target one is $\hat{G}_2$. We get $\mathbf{W}$ by Generative Adversarial Networks (GANs), where the generator learns the transformation matrix $\mathbf{W}$, ensuring that the transformed $\mathbf{W}\mathbf{Z}_{u1}$ approximates $\mathbf{Z}_{u2}$ as closely as possible. The discriminator tries to classify whether the embeddings are real $\mathbf{Z}_{u2}$ or those transformed ones. When $\mathbf{W}$ is trained well, we utilize the similarity measure CSLS proposed in [21] to find equivalent users.



(a) Iterative Propagation (b) Hierarchy Propagation

Figure 3: The propagation processes of the iterative architecture in GCN and the hierarchy architecture in DP-CroSUE. The node in blue represents the anchor user.

4.3 Cross-network GCN Embedding Model

After finding anchor users, we further integrate the partially aligned homogeneous user networks $\hat{G}_{u1} = (V_{u1}, \hat{E}_{u1})$ and $\hat{G}_{u2} = (V_{u2}, \hat{E}_{u2})$. To integrate them together, we propose a novel cross-network GCN embedding model, which is composed by an inter-graph propagation layer and a hierarchy intra-graph propagation layer.

- **inter-graph propagation layer:** We first update those aligned user nodes in both networks by fusing their information together. Considering the existing network disparity of $\hat{G}_{u1}$ and $\hat{G}_{u2}$, we leverage transformation matrices $\mathbf{W}_{12}$ and $\mathbf{W}_{21}$ to project embeddings into the other network space. For example, $\mathbf{W}_{12}$ is used to project users of $\hat{G}_{u1}$ into the space of $\hat{G}_{u2}$. Suppose two social users v_{u1}^{anchor} and v_{u2}^{anchor} are aligned, the representations of them after inter-graph propagation are:

$$\begin{aligned}\mathbf{z}'^{\text{anchor}}_{u1} &= \sigma\left(\mathbf{z}_{u1}^{anchor} + \mathbf{W}_{21}\mathbf{z}_{u2}^{anchor}\right), \\ \mathbf{z}'^{anchor}_{u2} &= \sigma\left(\mathbf{z}_{u2}^{anchor} + \mathbf{W}_{12}\mathbf{z}_{u1}^{anchor}\right),\end{aligned} \tag{10}$$

where σ denotes the activation function.

- **hierarchy intra-graph propagation layer:** We set the two user embeddings Z'_{u1} and Z'_{u2} with those aligned users getting updated through inter-graph propagation, while others are still the initial representations learned by GNN model. Now those anchor users' representations contain knowledge from both networks. Considering the number of anchor users is limited, propagating the cross-graph information from those aligned users to the entire graph needs more hops. Considering the over-fitting and over-smoothing problems of GCN [28], we leverage a hierarchy intra-graph propagation layer which emphasizes the role of anchor users and directly transfers information to nodes within k-hop range. Suppose $\hat{\mathbf{A}}_{u1}$ and $\hat{\mathbf{A}}_{u2}$ are the normalized adjacency matrices of the two user networks, and $\mathbf{D}_{u1}^{anchor}$ and $\mathbf{D}_{u2}^{anchor}$ are the diagonal matrices with the positions of anchors being set to one while others

Table 1: Statistics of datasets used in the experiments.

Dataset	Network	Nodes			Relationships	
		#(Users)	#(Anchor Users)	#(Posts)	#(friendship)	#(write)
Foursquare-Twitter	Foursquare	5392	3388	48756	76972	48756
	Twitter	5223		615,515	164,920	615,515
Weibo	Sub-weibo1	8117	2969	158,823	12000	158,823
	Sub-weibo2	8539		153,741	12000	153,741

Table 2: The TMR values of attribute, friendship and posts data.

data	Foursquare	Twitter	Sub-Weibo1	Sub-Weibo2
attribute	0.287	0.274	0.118	0.144
friendship	0.743	0.665	0.225	0.249
posts	0.645	0.537	0.207	0.222

are zeroes, the propagation function is:

$$\begin{aligned} \mathbf{H}_{u1}^{l} &= (\alpha \cdot \mathbf{D}_{u1}^{anchor} + \hat{\mathbf{A}}_{u1}^{l}) \cdot \mathbf{Z}'_{u1}, (l = 1, 2, ..., k), \\ \mathbf{H}_{u2}^{l} &= (\alpha \cdot \mathbf{D}_{u2}^{anchor} + \hat{\mathbf{A}}_{u2}^{l}) \cdot \mathbf{Z}'_{u2}, (l = 1, 2, ..., k), \\ \mathbf{O}_{u1} &= \sigma\left(\mathbf{W}_{u1} \cdot \text{stack}(\mathbf{Z}'_{u1}, \mathbf{H}_{u1}^{1}, ..., \mathbf{H}_{u1}^{k})\right), \\ \mathbf{O}_{u2} &= \sigma\left(\mathbf{W}_{u2} \cdot \text{stack}(\mathbf{Z}'_{u2}, \mathbf{H}_{u2}^{1}, ..., \mathbf{H}_{u2}^{k})\right). \end{aligned} \tag{11}$$

To show the difference between the original iterative propagation layer in GCN and our hierarchy one clearly, we depict their propagation processes in Figure 3. As shown in Figure 3(a), for the iterative architecture, it needs k transformations to transfer the carrying knowledge in the anchor user to its k-hop neighbours. For example, the forward propagation process of iterative architecture in $\hat{G}_{u1}$ is formulated as $\mathbf{O}^{(u1)} = \sigma\left(\hat{\mathbf{A}}_{u1}\left(\dots\left(\sigma\left(\hat{\mathbf{A}}_{u1}\mathbf{Z}'_{u1}\mathbf{W}_{u1}^{(0)}\right)\dots\right)\mathbf{W}_{u1}^{(k-1)}\right)$. Obviously, when k becomes large, there will be too many transformation parameters $\mathbf{W}_{u1}^{(i)}$ which will make the network difficult to train. Meanwhile, according to [28], if we apply iterative architecture to a connected graph, when k goes to infinity, nodes become indistinguishable. However, as for the hierarchy architecture, each anchor user can be transferred to other users inside k-hops directly. It only needs one transformation thus the knowledge of anchor users will not be changed too much when they are transferred to distant users. Meanwhile, the emphasis of anchor users can further help transfer cross-network knowledge. By considering the information of multiple hops together, the hierarchy architecture has more capacities in capturing its local information compared to the iterative architecture, which helps make users distinguishable.

• **Objective Function:** We jointly train these two networks together to integrate them. The loss includes three parts: the graph-based losses $\mathcal{L}_{\hat{G}_{u1}}$ and $\mathcal{L}_{\hat{G}_{u2}}$ computed as per Eq. 9, and the hard alignment regularization which measures the anchor nodes' representation distance in the two networks to regularize the output representation spaces. Overall, the total loss is:

$$\mathcal{L}_t = \mathcal{L}_{\hat{G}_{u1}} + \mathcal{L}_{\hat{G}_{u2}} + d(\mathbf{O}_{u1}^{anchor}, \mathbf{O}_{u2}^{anchor}). \tag{12}$$

5 EXPERIMENTS

In this section, we evaluate DP-CroSUE on both user interest prediction capacities and privacy protection strength. Specifically, we aim to answer the following questions: **Q1**: Compared with undisturbed single network knowledge, can the cross-network results of DP-CroSUE get improved in user interest prediction tasks? **Q2**: Can DP-CroSUE effectively protect sensitive user attribute information? **Q3**: How does each part of the perturbation paradigm of different information affect the accuracy of user interest prediction and the privacy-preserving property of DP-CroSUE? **Q4**: How does the novel cross-network GCN embedding model perform?

5.1 Datasets

We select three online social platforms: Foursquare, Twitter and Weibo, and make two partially-aligned dataset pairs. One is the Foursquare-Twitter pair from [61]. For the other, we crawled the Weibo platform and divided it into two sub-networks containing a part of sharing users. We manually labelled users from 10 possible interest categories according to their biographies and posts, including: (1) travel; (2) art; (3) health; (4) food; (5) technology; (6) sports; (7) business; (8) politics; (9) game; (10) fashion. Each user may be classified into multiple interests (multi-label classification). Table 1 shows the detailed statistical information about these datasets. The corresponding user attribute information extracted form the datasets are: **(1) user screen name** (We decompose the name into sub-string sets and convert them into numerical values using SimHash [45].), **(2) number of followers**, **(3) number of followees**, **(4) number of posts**, **(5) gender**, **(6) occupation**.

5.2 Baseline and Hyperparameter Setting

We compare DP-CroSUE with the following methods.

• **Single-view methods:** In general, single-view methods can be divided into text-based ones and structure-based ones. We select **Word2Vec** [36], **SANTEXT** [59], and **DeepWalk** [43] as single-view baselines. Among them, Word2Vec maps a sequence of social media posts into a vector representation. SANTEXT gets text representations under differential privacy. DeepWalk learns the latent representations of users in the homogeneous user network.

• **Multi-view methods:** For multi-view methods, we choose **SNE** [24], **UDMF** [63] and **R-GCN** [47] as multi-view baselines. SNE learns user representations by preserving both structural proximity and attribute proximity. UDMF is a novel hybrid DNN-based framework that fuses information across different modalities. R-GCN aggregates information on HSNs based on different relations.

• **Variants of DP-CroSUE:** We utilize **R-GCN** and **DP-R-GCN** to observe the performances of single-network user embeddings

Table 3: Overall performance for the user interest prediction tasks. Higher precision and F1 score represent better interest prediction performance. (Bold: best; Underline: runner-up.)

Methods	Foursquare		Twitter		Sub-Weibo1		Sub-Weibo2	
	Precision	Micro-f1	Precision	Micro-f1	Precision	Micro-f1	Precision	Micro-f1
Word2Vec [36]	0.463±0.013	0.455±0.009	0.463±0.002	0.450±0.003	0.226±0.001	0.238±0.002	0.231±0.011	0.236±0.009
SANTEXT [59]	0.457±0.022	0.446±0.018	0.446±0.020	0.442±0.015	0.207±0.012	0.193±0.020	0.229±0.019	0.233±0.016
DeepWalk [43]	0.453±0.012	0.447±0.008	0.413±0.010	0.409±0.007	0.215±0.003	0.176±0.004	0.247±0.004	0.243±0.004
SNE [24]	0.509±0.006	0.511±0.005	0.554±0.007	0.547±0.007	0.249±0.012	0.238±0.011	0.275±0.003	0.274±0.002
UDMF [63]	0.515±0.003	0.514±0.003	0.523±0.007	0.515±0.006	0.254±0.004	0.252±0.005	0.308±0.003	0.306±0.002
R-GCN [47]	0.530±0.002	0.522±0.001	0.546±0.001	0.542±0.001	0.283±0.002	0.279±0.002	0.317±0.007	0.312±0.005
DP-R-GCN	0.514±0.001	0.510±0.002	0.534±0.004	0.532±0.003	0.235±0.001	0.230±0.002	0.271±0.006	0.270±0.005
DP-CroSUE	0.549±0.006	0.541±0.005	0.561±0.001	0.558±0.001	0.303±0.003	0.296±0.002	0.332±0.003	0.334±0.002
CroSUE	**0.556±0.004**	**0.551±0.003**	**0.571±0.002**	**0.569±0.001**	**0.332±0.001**	**0.330±0.001**	**0.360±0.004**	**0.361±0.003**

obtained on the raw heterogeneous social networks and the perturbed ones, respectively. We also create a non-private model called **CroSUE**, which is a variant of DP-CroSUE deleting the Hybrid DP perturbation mechanism, to show the performance of the cross-network user embeddings without privacy protection.

In DP-CroSUE, to get the perturbed heterogeneous social graph, we set the privacy budget $\epsilon_a = 5$ for user attribute feature perturbation, $\epsilon_g = 10$ for user graph edge perturbation and $\epsilon_t = 7.5$ for textual data perturbation. Our guidance is the TMR values as shown in Table 2. The TMR values of attribute information and textual information are obtained through the corresponding attribute features and sentence embeddings (i.e., averaging word embeddings [36]), respectively, while the TMR of friendship information is obtained through the features learnt by DeepWalk [43]. To get initial user embeddings, we use a two-layer R-GCN. The first layer embedding dimension is 256 and the second embedding dimension is 128. For the cross-network GCN model, we set $k = 4$ and $\alpha = 2$ in the hierarchy intra-graph propagation layer. The final embedding dimension is 128.

5.3 Evaluation Metrics

The evaluation metrics for user interest prediction and attribute inference attack are depicted below:

User Interest Prediction. After getting the social user embeddings, we utilize the decision tree classifier to conduct multi-label user interest classification. We randomly pick 80% of data for training, while the rest 20% is for evaluation. To be more convincing, we report the mean and standard deviation of the results after repeating experiments for 5 times. The specific metrics we leverage to judge the quality are precision and micro-$F1$ score.

Attribute Inference Attack. To evaluate all models' robustness and capabilities in preserving sensitive user attribute information, we quantify the privacy leakage in social user embeddings through two inference attacks: gender inference and occupation inference. For gender inference, it is a binary classification problem. We use Logistic Regression as the attacker to make classification. For the occupation inference attack, we use the decision tree classifier to make classification. Similarly, we repeat the experiments for 5 times.

5.4 User Interest Prediction (Q1)

All models' performances on predicting user interests are summarized in Table 3. DP-CroSUE outperforms all baselines and is only next to CroSUE, which is the same method using the unsanitized network data. Generally, single-view based embedding methods - DeepWalk, SANTEXT and Word2vec get the worst performances. Compared to single-view methods, SNE which incorporates structure and attributes together works better. It achieves maximum relative improvements of 14.1% in precision on Twitter compared to DeepWalk. What's more, UDMF and R-GCN which incorporate more kinds of information perform even better than SNE on most datasets. For example, compared to SNE, R-GCN gets an improvement of 13.4% in precision on Sub-Weibo1. These observations demonstrate that incorporating more user information can help predict user interests. Compared with R-GCN, DP-R-GCN gets worse precision and Micro-$f1$ score. This is reasonable because DP-R-GCN extracts user information from the perturbed heterogeneous social network. These methods mentioned above are all single-network methods. Obviously, the best single-network method is R-GCN applied in real social networks. As we can see, DP-CroSUE consistently outperforms R-GCN on all the datasets. For example, the precision of DP-CroSUE is 2.0% higher than R-GCN on Sub-Weibo1 dataset. This indicates that DP-CroSUE, though operating on perturbed networks for the purpose of information protection, can achieve satisfactory results. In addition, CroSUE achieves the best performance. This further confirms the effectiveness of fusing networks together. However, since CroSUE works directly on real social networks, it faces user information leakage problems.

5.5 Attribute Inference Attacks (Q2)

Table 4 shows the results of the attribute inference attack models depicted in Section 5.3 on all the datasets. Note that lower scores show higher resistance to inference attacks. As shown in Table 4, all baselines operating on real social networks except for DeepWalk achieve high precision and Micro-$f1$ score in attribute inference attacks. For example, the precision scores of DeepWalk on all datasets are nearly 50.0%. That means DeepWalk has no ability to infer user's gender. **This is predictable as DeepWalk only extracts network structure information but does not involve user personal attribute features at all.** It is also worth noting that DP-CroSUE can achieve the second or third best results on almost all datasets. **Given that DeepWalk does not have a chance to learn user attribute features during training, the result that DP-CroSUE is only worse than DeepWalk highly validates the capability of DP-CroSUE in private attribute protection.** Meanwhile, the

Table 4: Overall performance for user gender and occupation inference attacks. Lower Precision and F1 scores represent better privacy protection. (Bold: best; Underline: runner-up.)

Attribute	Methods	Foursquare		Twitter		Sub-Weibo1		Sub-Weibo2	
		Precision	Micro-f1	Precision	Micro-f1	Precision	Micro-f1	Precision	Micro-f1
Gender	Word2Vec [36]	0.569±0.001	0.587±0.001	0.705±0.001	0.700±0.001	0.570±0.001	0.567±0.001	0.533±0.001	0.621±0.001
	SANTEXT [59]	0.559±0.001	0.578±0.001	0.644±0.001	0.639±0.001	0.560±0.001	0.555±0.001	0.517±0.001	0.614±0.001
	DeepWalk [43]	**0.508±0.001**	**0.510±0.001**	**0.501±0.001**	**0.509±0.002**	**0.506±0.001**	**0.517±0.001**	**0.498±0.001**	**0.583±0.001**
	SNE [24]	0.615±0.001	0.624±0.001	0.682±0.001	0.684±0.001	0.590±0.001	0.591±0.001	0.549±0.001	0.647±0.001
	UDMF [63]	0.995±0.001	0.995±0.001	0.999±0.001	0.999±0.001	0.999±0.001	0.999±0.001	0.998±0.001	0.998±0.001
	R-GCN [47]	0.791±0.001	0.794±0.001	0.738±0.001	0.739±0.001	0.727±0.001	0.727±0.001	0.684±0.001	0.727±0.001
	DP-R-GCN	0.568±0.002	0.579±0.001	0.624±0.001	0.627±0.001	0.538±0.001	0.537±0.001	0.518±0.001	0.617±0.001
	DP-CroSUE	0.554±0.001	0.575±0.001	0.640±0.001	0.644±0.001	0.536±0.001	0.535±0.001	0.506±0.001	0.601±0.001
	CroSUE	0.655±0.001	0.644±0.001	0.703±0.001	0.704±0.001	0.649±0.001	0.649±0.001	0.611±0.001	0.674±0.001
Attribute	Methods	Precision	Micro-f1	Precision	Micro-f1	Precision	Micro-f1	Precision	Micro-f1
Occupation	Word2Vec [36]	0.149±0.023	0.147±0.022	0.158±0.007	0.154±0.008	0.524±0.018	0.528±0.017	0.508±0.007	0.519±0.007
	SANTEXT [59]	0.139±0.012	0.140±0.018	0.143±0.010	0.139±0.007	0.510±0.013	0.513±0.012	0.497±0.004	0.490±0.014
	DeepWalk [43]	**0.102±0.016**	**0.098±0.015**	**0.120±0.003**	**0.118±0.005**	**0.450±0.001**	**0.462±0.001**	0.495±0.023	0.484±0.028
	SNE [24]	0.187±0.019	0.176±0.018	0.165±0.012	0.156±0.011	0.527±0.006	0.529±0.007	0.520±0.008	0.520±0.006
	UDMF [63]	0.702±0.016	0.707±0.017	0.938±0.001	0.954±0.001	0.963±0.002	0.974±0.002	0.912±0.008	0.934±0.011
	R-GCN [47]	0.160±0.015	0.158±0.014	0.184±0.005	0.185±0.004	0.525±0.002	0.536±0.003	0.553±0.004	0.540±0.007
	DP-R-GCN	0.134±0.005	0.135±0.006	0.160±0.005	0.157±0.007	0.474±0.006	0.497±0.012	0.411±0.012	0.417±0.009
	DP-CroSUE	0.132±0.015	0.136±0.017	0.134±0.005	0.128±0.005	0.466±0.016	0.474±0.017	**0.405±0.015**	**0.397±0.013**
	CroSUE	0.150±0.014	0.155±0.012	0.198±0.005	0.194±0.003	0.539±0.003	0.543±0.002	0.488±0.015	0.482±0.014

better performance of DP-R-GCN compared to R-GCN also indicates the effectiveness of our hybrid DP mechanism. For example, compared to R-GCN, the precision of DP-R-GCN drops significantly on Foursquare dataset in both gender inference and occupation inference (22.3% and 2.6% reduction). In addition, we noticed that DP-CroSUE, which incorporates information from two perturbed networks, performs even better in resisting attribute inference attacks than DP-R-GCN on most datasets. The reason is that if a user's attribute is disturbed into fake in one network, but remains true in another, linking these two nodes may result in both being inferred to false. This further shows the superiority of DP-CroSUE in privacy protection.

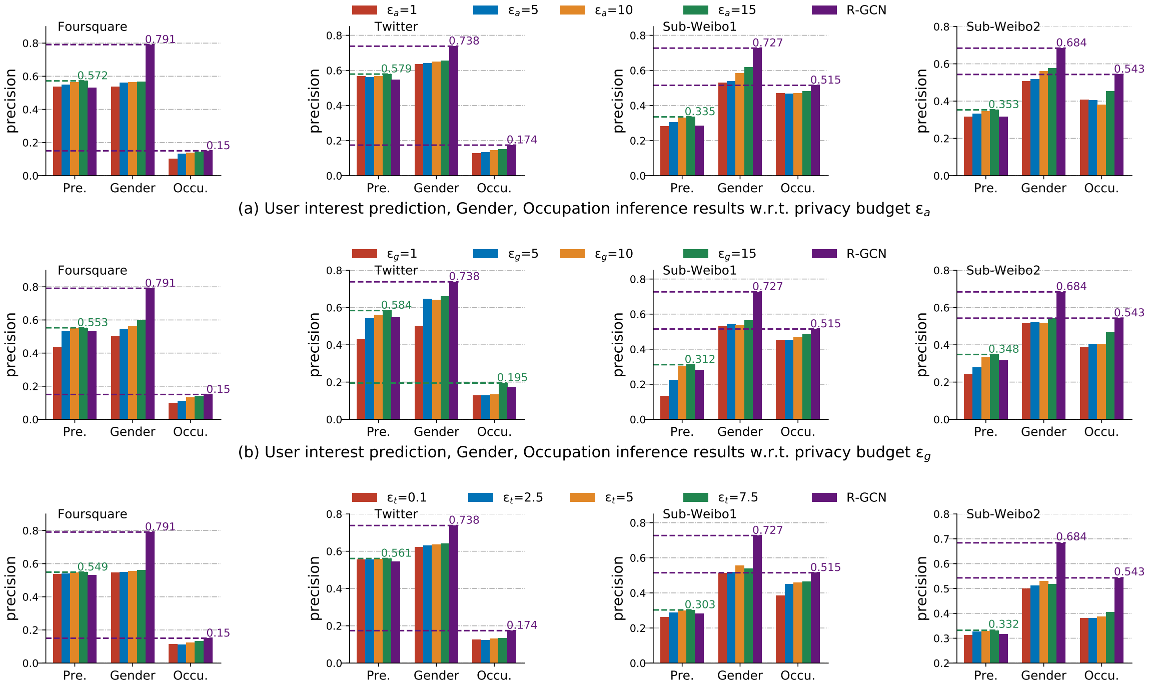
5.6 Accuracy and Privacy (Q3)

To investigate how does each part of the perturbation paradigm of DP-CroSUE affect its performance on user interest prediction and privacy protection, we independently vary the privacy budget ϵ_a, ϵ_g and ϵ_t, and report the new results in Figure 4.

• **Impact of user attribute perturbation privacy budget ϵ_a.** We vary the privacy budget $\epsilon_a \in \{1, 5, 10, 15\}$. The results are shown in Figure 4(a). In general, for all the datasets, the results of user interest prediction get improvement with the increase of ϵ_a. Because larger ϵ_a means less privacy thus more accurate user attributes, which helps predict user interests. For example, if we know a user is a woman, she may be more interested in fashion compared to a man. Meanwhile, for all the ϵ_a we selected, DP-CroSUE consistently outperforms R-GCN in terms of user interest prediction. This further confirms the effectiveness of our model. As for the privacy protection capabilities, larger ϵ_a, negatively influences the privacy protection results. Because more accurate user attributes are encoded into the final user representations. However, the gender and occupation information still can be effectively protected even when ϵ_a is set to a large value. From Figure 4(a), even when ϵ_a is set to 15, the precision scores of attribute inference attacks are lower than R-GCN. Our analysis is that by aligning and aggregating two perturbed datasets together, user's gender and occupation information may be further masked. Since if a user's gender or occupation is perturbed into fake in one network with the user attribute perturbation algorithm, his/her attribute in the other network may also be influenced. All the above results validate the superiority of DP-CroSUE in privacy protection.

• **Impact of user edge perturbation privacy budget ϵ_g.** We choose the value of ϵ_g from $\{1, 5, 10, 15\}$ and plot the results in Figure 4(b). Similarly, larger ϵ_g results in better user interest prediction performance. However, the prediction precision degrades significantly when the privacy budget is small. For example, when $\epsilon_g = 1$, the precision is even lower than R-GCN. As we know, the nature of GCN is smoothing - making connected nodes similar. Real social networks have the homophily property. People tend to follow those who have similar interests with them. If we inject a large amount of noise into the network structure, the homophily level of the graph will be impacted. Thus user prediction results will decrease while the user information can be better protected. In this way, we should choose a proper ϵ_g to achieve a good trade-off between privacy protection and recommendation accuracy.

• **Impact of user text perturbation privacy budget ϵ_t.** We study the impact of privacy budget in text perturbation with $\epsilon_t \in \{0.1, 2.5, 5, 7.5\}$. As we can see from Figure 4(c), in general, larger ϵ_t brings higher user interest prediction precision. Meanwhile, we find that the improvement is relatively slow with the increase of ϵ_t. The reason behind is the high utility property of the relaxed MDP notion applied in the textual data perturbation algorithm. Even when privacy budget is comparably small, with the computed substitution words, the original semantics may still be kept. Thus the user interest prediction performance can be guaranteed. What's more, it is worth noting that the perturbation of user text effectively preserves user attribute information compared to R-GCN. Since users may declare their gender and occupation in the original posts,



(a) User interest prediction, Gender, Occupation inference results w.r.t. privacy budget ϵ_a

(b) User interest prediction, Gender, Occupation inference results w.r.t. privacy budget ϵ_g

(c) User interest prediction, Gender, Occupation inference results w.r.t. privacy budget ϵ_t

Figure 4: User interest prediction and privacy protection results w.r.t. privacy budget ϵ_a, ϵ_g, ϵ_t.

substituting the original words can prevent information leakage effectively. For example, the word "girl" may be substituted by the word "boy".

To sum up, to make a good trade-off of interest prediction tasks and privacy protection, we recommend a comparably small attribute privacy budget, and comparably large edge and text privacy budgets. A good budget setting can be: $\epsilon_a = 5, \epsilon_g = 10$ and $\epsilon_t = 7.5$. The total privacy budget is $\epsilon_a + \epsilon_g + \epsilon_t \cdot d_{euc} = 15 + 7.5 \cdot d_{euc}$.

5.7 Ablation Study (Q4)

To study the performance of the novel cross-network GCN embedding model in network integration and user interests prediction, we evaluate different degraded model versions and record the precision scores in Table 5. For convenience, we denote the original cross-network GCN model introduced in Section 4.3 as CroGCN. We create CroGCN-NH by replacing the hierarchy intra-graph propagation layer with iterative propagation layers like standard GCN model. Besides, we also remove the hard alignment regularization and denote the new version as CroGCN-NA. As shown in Table 5, on all the datasets, CroGCN gets highest user interest prediction scores. The better performance of CroGCN compared with CroGCN-NH demonstrates the superiority of the hierarchy intra-graph propagation layer in transferring knowledge to the whole network. Meanwhile, the worse performance of CroGCN-NA also proves the importance of the hard alignment regularization.

6 CONCLUSION

In this work, we propose DP-CroSUE, which obtains privacy-preserving cross-network social user embeddings. DP-CroSUE perturbs users' different information including attribute features, friendship relations and user posts through a hybrid-DP mechanism to allow further data sharing. Next, through embedding-based alignment, anchor nodes can be found and different social networks can be learnt jointly in a pairwise manner without knowing the real user information. Extensive experiments demonstrate that DP-CroSUE simultaneously guards users against personal attribute inference attacks and maintain great utility in interest prediction tasks. Noted that DP-CroSUE can be easily generalized to other tasks like recommendation, exploitation on other tasks is left to future works.

Table 5: Ablation results of CroGCN in user interest prediction tasks.

Methods	Foursquare	Twitter	Sub-Weibo1	Sub-Weibo2
CroGCN-NH	0.539±0.006	0.543±0.002	0.298±0.006	0.325±0.009
CroGCN-NA	0.538±0.009	0.545±0.004	0.291±0.005	0.322±0.008
CroGCN	0.549±0.006	0.561±0.001	0.303±0.003	0.332±0.003

ACKNOWLEDGMENTS

The authors of this paper were supported by the National Key R&D Program of China through grant 2021YFB1714800, NSFC through grant U20B2053, S&T Program of Hebei through grant 21340301D, Beijing Natural Science Foundation through grant 4222030, and the Fundamental Research Funds for the Central Universities. Philip S. Yu was supported by NSF under grants III-1763325, III-1909323, III-2106758, and SaTC-1930941. Thanks for computing infrastructure provided by Huawei MindSpore platform. For any correspondence, please refer to Hao Peng.

REFERENCES

[1] Mário Alvim, Konstantinos Chatzikokolakis, Catuscia Palamidessi, and Anna Pazii. 2018. Local differential privacy on metric spaces: optimizing the trade-off with utility. In *CSF*. IEEE, 262–267.

[2] Athanasios Andreou, Oana Goga, and Patrick Loiseau. 2017. Identity vs. attribute disclosure risks for users with multiple social profiles. In *ASONAM*. 163–170.

[3] Jeremiah Blocki, Avrim Blum, Anupam Datta, and Or Sheffet. 2012. The johnson-lindenstrauss transform itself preserves differential privacy. In *SFCS*. IEEE, 410–419.

[4] Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2017. Enriching word vectors with subword information. *TACL* 5 (2017), 135–146.

[5] Kseniya Buraya, Aleksandr Farseev, Andrey Filchenkov, and Tat-Seng Chua. 2017. Towards user personality profiling from multiple social networks. In *AAAI*. 4909–4910.

[6] Xuezhi Cao and Yong Yu. 2017. Joint User Modeling Across Aligned Heterogeneous Sites Using Neural Networks. In *ECML-PKDD*. Springer, 799–815.

[7] Sicong Che, Zhaoming Kong, Hao Peng, Lichao Sun, Alex Leow, Yong Chen, and Lifang He. 2022. Federated multi-view learning for private medical data integration and analysis. *ACM TIST* 13, 4 (2022), 1–23.

[8] Terence Chen, Mohamed Ali Kaafar, Arik Friedman, and Roksana Boreli. 2012. Is more always merrier? A deep dive into online social footprints. In *WOSN*. 67–72.

[9] Xiaokai Chu, Xinxin Fan, Di Yao, Zhihua Zhu, Jianhui Huang, and Jingping Bi. 2019. Cross-network embedding for multi-network alignment. In *WWW*. 273–284.

[10] Jinming Cui, Chaochao Chen, Lingjuan Lyu, Carl Yang, and Wang Li. 2021. Exploiting Data Sparsity in Secure Cross-Platform Social Recommendation. *Advances in Neural Information Processing Systems* 34 (2021).

[11] Bolin Ding, Janardhan Kulkarni, and Sergey Yekhanin. 2017. Collecting telemetry data privately. *Advances in Neural Information Processing Systems* 30 (2017).

[12] John C Duchi, Michael I Jordan, and Martin J Wainwright. 2013. Local privacy and statistical minimax rates. In *FOCS*. IEEE, 429–438.

[13] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. 2006. Calibrating noise to sensitivity in private data analysis. In *TCC*. Springer, 265–284.

[14] Úlfar Erlingsson, Vasyl Pihur, and Aleksandra Korolova. 2014. Rappor: Randomized aggregatable privacy-preserving ordinal response. In *SIGSAC*. 1054–1067.

[15] Oluwaseyi Feyisetan, Borja Balle, Thomas Drake, and Tom Diethe. 2020. Privacy- and utility-preserving textual analysis via calibrated multivariate perturbations. In *WSDM*. 178–186.

[16] Oluwaseyi Feyisetan, Tom Diethe, and Thomas Drake. 2019. Leveraging hierarchical representations for preserving privacy and utility in text. In *ICDM*. IEEE, 210–219.

[17] Tianxi Ji, Changqing Luo, Yifan Guo, Jinlong Ji, Weixian Liao, and Pan Li. 2019. Differentially private community detection in attributed social networks. In *ACML*. PMLR, 16–31.

[18] Meng Jiang, Peng Cui, Nicholas Jing Yuan, Xing Xie, and Shiqiang Yang. 2016. Little Is Much: Bridging Cross-Platform Behaviors through Overlapped Crowds. In *AAAI*, Vol. 30.

[19] Zach Jorgensen, Ting Yu, and Graham Cormode. 2016. Publishing attributed social graphs with formal privacy guarantees. In *SIGMOD*. 107–122.

[20] Rimma Kats. 2018. Many Facebook Users Are Sharing Less Content. *Recuperado de https://www. emarketer. com/content/many-facebook-users-are-sharing-less-content-because-of-privacy-concerns* (2018).

[21] Guillaume Lample, Alexis Conneau, Marc'Aurelio Ranzato, Ludovic Denoyer, and Hervé Jégou. 2018. Word translation without parallel data. In *ICLR*.

[22] Ninghui Li, Tiancheng Li, and Suresh Venkatasubramanian. 2007. t-closeness: Privacy beyond k-anonymity and l-diversity. In *ICDE*. IEEE, 106–115.

[23] Zhehan Liang, Yu Rong, Chenxin Li, Yunlong Zhang, Yue Huang, Tingyang Xu, Xinghao Ding, and Junzhou Huang. 2021. Unsupervised Large-Scale Social Network Alignment via Cross Network Embedding. In *CIKM*. 1008–1017.

[24] Lizi Liao, Xiangnan He, Hanwang Zhang, and Tat-Seng Chua. 2018. Attributed social network embedding. *TKDE* 30, 12 (2018), 2257–2270.

[25] Bang Hui Lim, Dongyuan Lu, Tao Chen, and Min-Yen Kan. 2015. # mytweet via instagram: Exploring user behaviour across multiple social networks. In *ASONAM*. IEEE, 113–120.

[26] Tzu-Heng Lin, Chen Gao, and Yong Li. 2019. Cross: Cross-platform recommendation for social e-commerce. In *SIGIR*. 515–524.

[27] Li Liu, William K Cheung, Xin Li, and Lejian Liao. 2016. Aligning Users across Social Networks Using Network Embedding.. In *IJCAI*. 1774–1780.

[28] Meng Liu, Hongyang Gao, and Shuiwang Ji. 2020. Towards deeper graph neural networks. In *SIGKDD*. 338–348.

[29] Peng Liu, YuanXin Xu, Quan Jiang, Yuwei Tang, Yameng Guo, Li-e Wang, and Xianxian Li. 2020. Local differential privacy for social network publishing. *Neurocomputing* 391 (2020), 273–279.

[30] Zhiwei Liu, Liangwei Yang, Ziwei Fan, Hao Peng, and Philip S. Yu. 2021. Federated social recommendation with graph neural network. *ACM TIST* (2021).

[31] Wentian Lu and Gerome Miklau. 2014. Exponential random graph estimation under differential privacy. In *SIGKDD*. 921–930.

[32] Lingjuan Lyu, Xuanli He, and Yitong Li. 2020. Differentially Private Representation for NLP: Formal Guarantee and An Empirical Study on Privacy and Fairness. In *Findings of the Association for Computational Linguistics: EMNLP 2020*. 2355–2365.

[33] Lingjuan Lyu, Yitong Li, Xuanli He, and Tong Xiao. 2020. Towards differentially private text representations. In *SIGIR*. 1813–1816.

[34] Ashwin Machanavajjhala, Daniel Kifer, Johannes Gehrke, and Muthuramakrishnan Venkitasubramaniam. 2007. l-diversity: Privacy beyond k-anonymity. *TKDD* 1, 1 (2007), 3–es.

[35] Tong Man, Huawei Shen, Shenghua Liu, Xiaolong Jin, and Xueqi Cheng. 2016. Predict anchor links across social networks via an embedding approach.. In *IJCAI*, Vol. 16. 1823–1829.

[36] Tomas Mikolov, Kai Chen, Gregory S. Corrado, and Jeffrey Dean. 2013. Efficient Estimation of Word Representations in Vector Space. In *ICLR*.

[37] Arvind Narayanan and Vitaly Shmatikov. 2009. De-anonymizing social networks. In *SSP*. IEEE, 173–187.

[38] Arvind Narayanan and Vitaly Shmatikov. 2010. Myths and fallacies of" personally identifiable information". *CACM* 53, 6 (2010), 24–26.

[39] Hiep Nguyen, Abdessamad Imine, and Michaël Rusinowitch. 2016. Network structure release under differential privacy. *Transactions on Data Privacy* 9, 3 (2016), 26.

[40] Hiep H Nguyen, Abdessamad Imine, and Michaël Rusinowitch. 2015. Differentially private publication of social graphs at linear cost. In *ASONAM*. IEEE, 596–599.

[41] Hao Peng, Haoran Li, Yangqiu Song, Vincent Zheng, and Jianxin Li. 2021. Differentially private federated knowledge graphs embedding. In *CIKM*. 1416–1425.

[42] Jeffrey Pennington, Richard Socher, and Christopher D Manning. 2014. Glove: Global vectors for word representation. In *EMNLP*. 1532–1543.

[43] Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. 2014. Deepwalk: Online learning of social representations. In *SIGKDD*. 701–710.

[44] Fuxin Ren, Zhongbao Zhang, Jiawei Zhang, Sen Su, Li Sun, Guozhen Zhu, and Congying Guo. 2020. BANANA: when Behavior ANAlysis meets social Network Alignment. In *IJCAI*. 1438–1444.

[45] Caitlin Sadowski and Greg Levin. 2007. Simhash: Hash-based similarity detection. *Technical report, Google* (2007).

[46] Alessandra Sala, Xiaohan Zhao, Christo Wilson, Haitao Zheng, and Ben Y Zhao. 2011. Sharing graphs using differentially private graph models. In *SIGCOMM*. 81–98.

[47] Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne Van Den Berg, Ivan Titov, and Max Welling. 2018. Modeling relational data with graph convolutional networks. In *ESWC*. Springer, 593–607.

[48] Latanya Sweeney. 2002. k-anonymity: A model for protecting privacy. *IJUFKS* 10, 05 (2002), 557–570.

[49] Stacey Truex, Ling Liu, Ka-Ho Chow, Mehmet Emre Gursoy, and Wenqi Wei. 2020. LDP-Fed: Federated learning with local differential privacy. In *EdgeSys*. 61–66.

[50] Tianhao Wang, Jeremiah Blocki, Ninghui Li, and Somesh Jha. 2017. Locally differentially private protocols for frequency estimation. In *USENIX Security 17*. 729–745.

[51] Yue Wang and Xintao Wu. 2013. Preserving differential privacy in degree-correlation based graph generation. *Transactions on data privacy* 6, 2 (2013), 127.

[52] Yuye Wang, Jing Yang, and Jianpei Zhan. 2021. Differentially Private Attributed Network Releasing Based on Early Fusion. *Security and Communication Networks* 2021 (2021).

[53] Qian Xiao, Rui Chen, and Kian-Lee Tan. 2014. Differentially private network data release via structural inference. In *SIGKDD*. 911–920.

[54] Xiaohang Xu, Hao Peng, Md Zakirul Alam Bhuiyan, Zhifeng Hao, Lianzhong Liu, Lichao Sun, and Lifang He. 2021. Privacy-Preserving Federated Depression Detection From Multisource Mobile Health Data. *IEEE Transactions on Industrial Informatics* 18, 7 (2021), 4788–4797.

[55] Ming Yan, Jitao Sang, Tao Mei, and Changsheng Xu. 2013. Friend transfer: Cold-start friend recommendation with cross-platform transfer learning of social knowledge. In *ICME*. IEEE, 1–6.

[56] Ming Yan, Jitao Sang, Changsheng Xu, and M Shamim Hossain. 2016. A unified video recommendation by cross-network user modeling. *TOMM* 12, 4 (2016), 1–24.

[57] Carl Yang, Haonan Wang, Ke Zhang, Liang Chen, and Lichao Sun. 2021. Secure Deep Graph Generation with Link Differential Privacy. In *IJCAI*. 3271–3278.

[58] Xin Yao, Rui Zhang, and Yanchao Zhang. 2021. Differential Privacy-Preserving User Linkage across Online Social Networks. In *IWQOS*. IEEE, 1–10.

[59] Xiang Yue, Minxin Du, Tianhao Wang, Yaliang Li, Huan Sun, and Sherman S. M. Chow. 2021. Differential Privacy for Text Analytics via Natural Text Sanitization. In *ACL-IJCNLP*.

[60] Jiawei Zhang, Xiangnan Kong, and Philip S. Yu. 2013. Predicting social links for new users across aligned heterogeneous social networks. In *ICDM*. IEEE, 1289–1294.

[61] Jiawei Zhang, Congying Xia, Chenwei Zhang, Limeng Cui, Yanjie Fu, and Philip S. Yu. 2017. BL-MNE: emerging heterogeneous social network embedding through broad learning with aligned autoencoder. In *ICDM*. IEEE, 605–614.

[62] Shijie Zhang, Hongzhi Yin, Tong Chen, Zi Huang, Lizhen Cui, and Xiangliang Zhang. 2021. Graph Embedding for Recommendation against Attribute Inference Attacks. In *WWW*. 3002–3014.

[63] Wei Zhang, Wen Wang, Jun Wang, and Hongyuan Zha. 2018. User-guided hierarchical attention network for multi-modal social image popularity prediction. In *WWW*. 1277–1286.