Quasi-Metric Learning for Bilateral Person-Job Fit

Yingpeng Du, Hongzhi Liu, Hengshu Zhu, *Senior Member, IEEE*, Yang Song, Zhi Zheng, and Zhonghai Wu

***Abstract*—Matching suitable jobs with qualified candidates is crucial for online recruitment. Typically, users (i.e., candidates and employers) have specific expectations in the recruitment market, making them prefer similar jobs or candidates. Metric learning technologies provide a promising way to capture the similarity propagation between candidates and jobs. However, they rely on symmetric distance measures, failing to model users' asymmetric relationships in two-way selection. Additionally, users' behaviors (e.g., candidates) are highly affected by the feedback from their counterparts (e.g., employers), which can hardly be captured by the existing person-job fit methods that primarily explore homogeneous and undirected graphs. To address these problems, we propose a quasi-metric learning framework to capture the similarity propagation between candidates and jobs while modeling their asymmetric relations for bilateral person-job fit. Specifically, we propose a quasi-metric space that not only satisfies the triangle inequality to capture the fine-grained similarity between candidates and jobs, but also incorporates a tailored asymmetric measure to model users. two-way selection process in online recruitment. More importantly, the proposed quasi-metric learning framework can theoretically model recruitment rules from *similarity* and *competitiveness* perspectives, making it seamlessly align with bilateral person-job fit scenarios. To explore the mutual effects of two-sided users, we first organize candidates, employers, and their different-typed interactions into a heterogeneous relation graph, and then propose a relation-aware graph convolution network to capture users. mutual effects through their bilateral behaviors. Extensive experiments on several real-world datasets demonstrate the effectiveness of the proposed methods.**

***Index Terms*—Bilateral person-job fit, heterogeneous relations, quasi-metric learning, asymmetric distance.**

I. Introduction

RECENTLY, bilateral person-job fit methods [1], [2] have gained popularity for enhancing both user satisfaction and platform profit in online recruitment, aiming to match candidates with employers and look forward to the agreement for both parties. For example, online recruitment platforms such as LinkedIn and BOSS Zhipin have numerous bilateral users, including candidates (also known as job seekers or persons) and employers (who offer job positions). These platforms provide candidates with suitable jobs and match employers with qualified candidates to facilitate agreements. In recruitment scenarios, candidates and employers typically have specific expectations for job positions and candidates, which often leads them to favor similar job positions and candidates. For example, candidates usually possess specific job skills, residential addresses, and educational backgrounds, leading them to prefer job positions that offer similar responsibilities, are located nearby, and match their qualification levels. Specifically, several illustrative data statistics are conducted in Section III-B, which indicate that users in recruitment scenarios tend to favor the "similar items" compared to other application scenarios. Therefore, capturing the fine-grained similarities among candidates and job positions is crucial for improved bilateral person-job fit.

Fortunately, the metric learning technology provides a promising way for modeling such fine-grained similarity among candidates and job positions, which can capture the similarity propagation behind their interactions [3], [4], [5], [6], [7], [8]. As shown in Fig. 1(a), the triangle inequality axiom $d(A,B) \leq d(A,C) + d(C,B)$ of metric space will pull the two entities (e.g., A and B) close to each other if they are similar to one common entity (e.g., C), facilitating capture the fine-grained similarity propagation among entities. Such similarity propagation can not be well captured by existing inner product or multi-layer perception (MLP) based methods [3], [9] as they violate the triangle inequality axiom. However, it remains challenging to directly model person-job fit with existing metric learning technologies for several reasons. First, bilateral person-job fit is essentially a two-way selection process [10]. Existing metric learning methods rely on the symmetric distance $d(\cdot,\cdot)$ between users (i.e., $d(c,j) = d(j,c)$), thus failing to model asymmetrically relations of candidates and employers for bilateral person-job fit. For example, a candidate c applied for a job j but might get rejected by the employer, indicating the asymmetrical nature of their relationship. Second, removing the symmetric axioms in metric space is not well-explored in machine learning, making it challenging to model the asymmetric relationships that seamlessly align with bilateral person-job fit scenarios.

In addition, the bilateral interactions between candidates and job positions are naturally exhibited with the relationship diversity [11], leading to varying mutual influences among them in

Received 30 August 2023; revised 26 June 2024; accepted 25 January 2025. Date of publication 4 February 2025; date of current version 3 April 2025. This work was supported in part by National Key Research and Development Program of China under Grant 2022YFB2703301 and in part by S&T Program of Hebei under Grant 246W7703D. Recommended for acceptance by B. Hammer. *(Corresponding authors: Hongzhi Liu; Hengshu Zhu; Zhonghai Wu.)*

Yingpeng Du is with the Nanyang Technological University, Singapore 639798, and also with Peking University, Beijing 100871, China (e-mail: dyp1993@pku.edu.cn).

Hongzhi Liu is with the School of Software and Microelectronics, Peking University, Beijing 100871, China (e-mail: liuhz@pku.edu.cn).

Hengshu Zhu is with the Computer Network Information Center, Chinese Academy of Sciences, China (e-mail: zhuhengshu@gmail.com).

Yang Song is with the BOSS Zhipin Inc., Beijing 100028, China (e-mail: songyang@kanzhun.cn).

Zhi Zheng is with the University of Science and Technology of China, Hefei 101127, China (e-mail: zhengzhi97@mail.ustc.edu.cn).

Zhonghai Wu is with the National Engineering Center of Software Engineering, Key Lab of High Confidence Software Technologies (MOE), Peking University, Beijing 100871, China (e-mail: wuzh@pku.edu.cn).

Digital Object Identifier 10.1109/TPAMI.2025.3538774

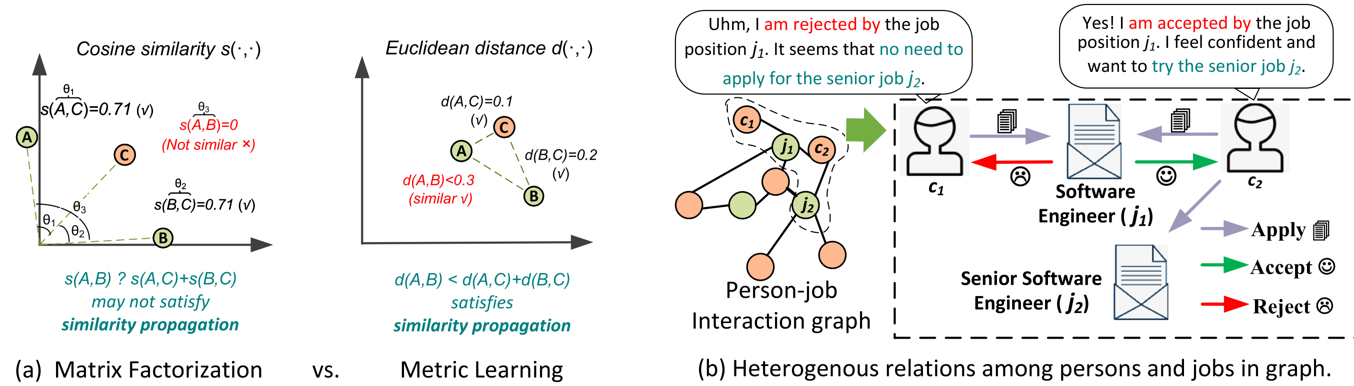


Fig. 1. (a) Matrix factorization with cosine similarity $s(\cdot,\cdot)$ versus Metric learning with Euclidean distance $d(\cdot,\cdot)$. (b) An illustrative example of online recruitment for candidates c_1, c_2 and job positions j_1, j_2, where arrows with different colors denote the different-typed interactions among them.

recruitment scenarios. As evidence in social psychology shows, positive feedback increases people's confidence to pursue their goals, while negative feedback undermines their confidence in pursuing their goals and their expectations of success [12]. Specifically, Fig. 1(b) depicts a bilateral relationship graph between candidates and job positions for online recruitment, where candidates c_1 and c_2, despite having similar preferences, applied for job position j_1 and received different feedback (e.g., rejection and acceptance). This may further affect candidates' behaviors, e.g., candidate c_2 is more likely to apply for the senior job j_2 compared to candidate c_1 because candidate c_2 is more confident to be accepted by job j_2. However, such effects can hardly be captured by most of the existing person-job fit methods that primarily rely on homogeneous and undirected graphs [2], [13], [14], [15], as they fail to model the heterogeneous relations between users across their different-typed interactions. To this end, we explore the heterogeneous relations between users (candidates and employers) to capture their mutual effects on each other, contributing to more accurate matching results.

In this paper, we propose a quasi-metric learning framework for bilateral person-job fit. To alleviate the dilemma of existing methods, we come up with a quasi-metric space with an asymmetric distance to measure the bilateral preferences of two-sided users. On the one hand, we propose capturing the similarity propagation among candidates and job positions by benefiting from the triangle inequality axiom [3] of quasi-metric space. On the other hand, we propose conducting quasi-metric learning to model the two-way selection process of users with an asymmetric distance. The proposed quasi-metric space is designed with thorough consideration of the similarity and competitiveness rules for online recruitment scenarios, ensuring it theoretically and seamlessly aligns with bilateral person-job fit scenarios. Specifically, the similarity rule models the relatedness of candidates and job positions, e.g., matching candidates with relative job positions. The competitiveness rule models the competitiveness relationship in the recruitment market, e.g., employers may prefer competitive candidates with stronger educational backgrounds and more relevant experiences. These rules contribute to quasi-metric learning to model asymmetric relationships between two-sided users. To explore users' mutual effects on each other, we propose a relation-aware graph convolution network to model the heterogeneous relations between candidates and employers, which enables the capturing of how users' actions influence others and how they are influenced by actions of others. Extensive experiments on four real-world datasets validate the effectiveness of the proposed method.

The main contributions of this paper can be summarized as follows

- We propose a quasi-metric space with an asymmetric distance measure for bilateral person-job fit. To the best of our knowledge, this is the first time to explore the quasi-metric space on person-job fit and recommendation tasks. We also establish two recruitment rules to construct the quasi-metric space, making it align well with real recruitment scenarios theoretically.
- We propose to explore the heterogeneous relations between candidates and employers with a relation-aware graph convolution network, which can capture their mutual effects behind different-typed interactions.
- We evaluate our model on four real-world datasets, and the experimental results demonstrate that the proposed method consistently outperforms state-of-the-art methods for bilateral users. Ablation experiments and case studies further validate the motivations behind our approach.

II. Related Work

A. Metric Learning in Recommender Systems

Recommender systems help users discover the items that they need and prefer from numerous choices, which can enhance the satisfaction of users and the profit of platforms. Among the methods used in recommender systems, collaborative filtering (CF) based methods such as matrix factorization (MF) [16] gain popularity due to their good performance and elegant theory, which are augmented with delegate modification such as utilizing deep neural networks [17], [18], [19], [20] and graph models [21], [22], [23]. However, these methods fail to capture the fine-grained similarity propagation of users and items as they violate the triangle inequality [3].

Metric learning techniques [24], [25] aim to learn an appropriate data distribution with the distance function (e.g., Euclidean distance), which have been utilized in a wide range of applications, e.g., computer vision [26], [27], [28], knowledge graphs [29], and recommender systems [3], [9]. In recommender systems, one of the most representative works, CML [3], argued that the inner product violated the triangle inequality and failed to capture the fine-grained preferences of users. As demonstrated by [9], even if a user interacted with two similar items, MF or MLP-based methods will place these two items close to the user but will not necessarily place these items close to each other. To this end, various methods have embraced metric learning approaches to effectively capture similarity propagation with the triangle inequality, which demonstrates highly competitive performance in recommendation tasks [3], [4], [6], [7], [30]. Significant effort has been devoted to improving metric learning technology for recommendations with flexible margins [5], [8], sampling strategies [31], and distance measure [32], [33]. However, these methods focus on exploring the symmetric relationship between users and items based on symmetric distance measures, which have difficulty in modeling the asymmetric two-way selection process for bilateral person-job fit.

B. Person-Job Fit in Recommender Systems

Reciprocal recommendation is a type of recommender system that suggests users to other users rather than products. It is widely adopted in several areas such as recruitment [34], [35] and online dating [36], [37]. In this paper, we focus on a special case of reciprocal recommendation, called bilateral person-job fit, which aims to accurately match suitable job positions with desirable candidates, ultimately leading to mutually beneficial agreements [38]. According to the ways of modeling bilateral relations, existing person-job fit methods can be classified into symmetric and asymmetric methods.

Symmetric methods focus on the undirected matching for candidates and job positions, such as text-based matching [39], [40], [41], [42], [43], [44], collaborative filtering [36], [37], [45], [46], [47]. To exploit the rich textual semantic information in resumes and job requirements, several methods have adopted text-matching strategies or text enhancement methods, such as CNN [40], [48], RNN [39], and memory network [49], for person-job fit. For example, APJFNN [39] utilized hierarchical attention RNN models the word-level semantic representation for both candidate resumes and job descriptions for person-job fit. Moreover, several methods have been investigated to enhance the expressive power of text encoders such as adversarial learning [50], co-teaching mechanisms [51], and transfer learning [52]. *Asymmetric methods attempt to model the asymmetric relationship between candidates and job positions*. Several methods adopted an undirected graph [2] or separate bipartite graphs (matrices) [13], [14] to model the asymmetric relationship between candidates and job positions. For example, DPGNN [2] utilized a unified dual-perspective interaction graph to model the active and passive aspects of candidates and jobs. Several methods adopt the multi-task framework to model the asymmetric relationship between candidates and jobs [15], [53], [54]. For example, IPJF [53] adopted a multi-task optimization approach to learn the asymmetric perspectives/intentions of employers and candidates based on shared text features. However, most of these methods fall short of adequately capturing similarity propagation and modeling the diverse relations of users. To address these limitations, we first propose a quasi-metric learning framework to capture both similarity propagation and asymmetric relations between candidates and job positions, which can be enhanced by the proposed relation-aware GCNs to capture the heterogeneous relations between users.

III. Preliminary

A. Problem Formulation

Let $\mathcal{C} = \{c_1, \ldots, c_N\}$ and $\mathcal{J} = \{j_1, \ldots, j_M\}$ represent the sets of N candidates and M jobs, respectively. Each candidate or job is associated with a text document that describes the resume or the job requirements. We suppose to know the bilateral relations between candidates and jobs, including the direction $\varrho_{ik} \in \{\rightarrow, \leftarrow\}$ (i.e., proactive or passive) and feedback $\pi_{ik} \in \{\surd, \times\}$ (i.e., accept or reject). We summarize the heterogeneous relations between users as follows:

- In the candidate side, if a candidate c_i applied for a job j_k but received a rejection or was ignored by the employer, the corresponding interaction set is denoted as $\mathcal{A}_c^{\rightarrow} = \{(c_i, j_k) | \varrho_{ik} = \rightarrow \wedge \pi_{ik} = \times\}$. Alternatively, if a candidate c_i rejected or ignored the interview of job j_k, the corresponding interaction set is denoted as $\mathcal{A}_c^{\leftarrow} = \{(c_i, j_k) | \varrho_{ik} = \leftarrow \wedge \pi_{ik} = \times\}$.
- In the employer side, if an employer who posted job j_k reached out to the candidate c_i but got a rejection, the corresponding interaction set is denoted as $\mathcal{A}_j^{\rightarrow} = \{(j_k, c_i) | \varrho_{ik} = \rightarrow \wedge \pi_{ik} = \times\}$. Alternatively, if the employer who posted job j_k rejected the job application of a candidate c_i, the corresponding interaction set is denoted as $\mathcal{A}_j^{\leftarrow} = \{(j_k, c_i) | \varrho_{ik} = \leftarrow \wedge \pi_{ik} = \times\}$.
- If the candidate and the employer reach the agreement, we denote the corresponding interaction sets $\mathcal{A}_c^{\leftrightarrow} = \{(c_i, j_k) | \pi_{ik} = \surd\}$ and $\mathcal{A}_j^{\leftrightarrow} = \{(j_k, c_i) | \pi_{ik} = \surd\}$ for the candidate and job sides, respectively.

In this paper, our goal is to effectively match job positions with suitable candidates, ultimately facilitating mutually beneficial agreements. This process involves providing bilateral (two-way) recommendations for both candidates and employers. Formally, we aim to accomplish this goal by developing a matching function, denoted as $m(c_i, j_k)$, based on their textual descriptions and bilateral relations. Taking the task of providing a candidate c_i with job positions as an example (i.e., to the candidate side), we aim to rank the job list for the candidate c_i as $R_{c_i} = \{j_{(1)}, \ldots, j_{(M)}\}$, where $j_{(k)}$ denotes the kth job position in the ranking list R_{c_i}. Symmetrically, the task of recommending candidates to a job position j_k with candidates (i.e., to the employer side) involves ranking the candidates list as $R_{j_k} = \{c_{(1)}, \ldots, c_{(N)}\}$. The main notations used throughout the paper are summarized in Table I.

B. Data Analysis

To analyze the behavioral patterns of users, we conducted a statistical analysis in different scenarios as shown in

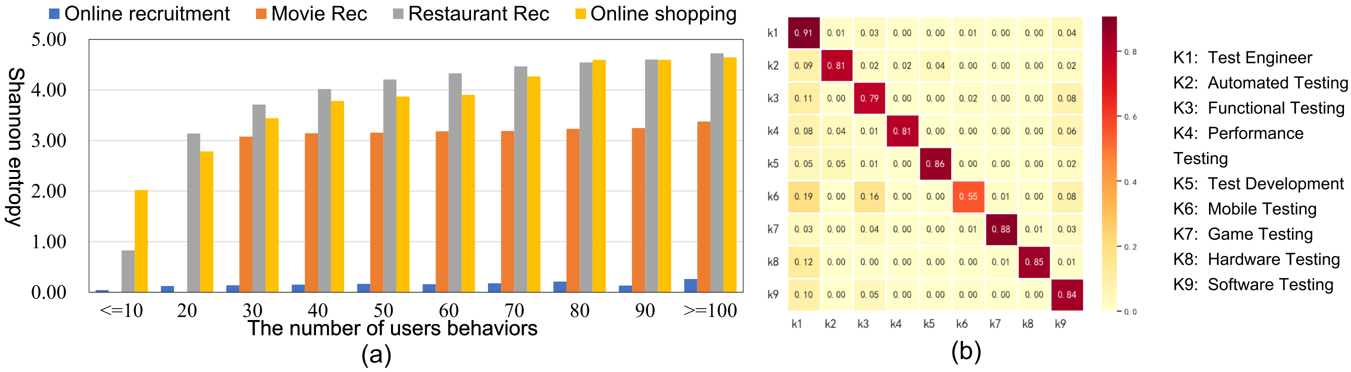


Fig. 2. (a) The statistics on users' preference concentration among different recommendation scenarios. (b) Transitions of users' application behaviors among keywords related to "IT testing" job positions in online recruitment.

TABLE I
SYMBOLS AND NOTATIONS

Notation	Description
N, M	The number of candidates and job positions.
$\mathcal{C} = \{c_1, \cdots, c_N\}$	The whole set of candidates.
$\mathcal{J} = \{j_1, \cdots, j_M\}$	The whole set of jobs.
T_{c_i}, T_{j_k}	Resume of candidate c_i, job requirements of job position j_k.
$\varrho \in \{\rightarrow, \leftarrow\}$	The direction of interactions.
$\pi \in \{\surd, \times\}$	The feedback, e.g., accept ($\surd$) and reject ($\times$).
$r(i,k) = (\varrho_{ik}, \pi_{ik})$	Asymmetric relation between candidate (or job) i and job (or candidate) k.
$\mathcal{A}_c^*$, $\mathcal{A}_j^*$	Different-typed interactions for candidates and jobs
P, Q	Embedding matrices of candidates and jobs.
W_t	Text encoding by the BERT model.
M_v	Relation-aware diagonal matrix.
x_{c_i}, x_{j_k}	The GCN based embeddings.
$q(\cdot, \cdot)$	The asymmetric distance in quasi-metric space.
$(\mathcal{S}, q)$, $(\mathcal{S}, d)$	Quasi-metric space and metric space with distance functions q and d, respectively.
d_e	The dimension of latent space.

Fig. 2(a). Specifically, we gather statistics on users' preference concentration grouped by the number of users' behaviors with different scenarios, including recruitment (Boss Zhipin[1]), movie recommendation (ML-1 M dataset[2]), online shopping (Amazon clothing, shoes and jewelry dataset[3]), and restaurant recommendation (Dianping[4]). The users' concentration is calculated by the Shannon entropy of their behavior distribution across different categories of items,

$$H = \frac{1}{N} \sum_{u=1}^{N} \left(- \sum_{i=1}^{R} p_{u,r} \log_2 p_{u,r} \right),$$

where N is the number of users, $\{p_{u,r} | r = 1, \ldots, R\}$ denotes the behaviors' distribution over R categories of user u which satisfied $p_{u,r} = Num(u,r)/\sum_{r=1}^{R} Num(u,r)$. $Num(u,r)$ denotes the number of user u's interactions on items that belong to category r. The higher entropy generally signifies a more diverse behavior pattern across various categories, whereas the lower entropy suggests a more concentrated behavior pattern within fewer categories. We observe that users in the online recruitment scenario exhibit a higher concentration among categories, which implies that candidates prefer job positions that share similar categories. Furthermore, we examine users' sequential behaviors in the online recruitment scenario within the same category (e.g., "IT testing" related jobs) as shown in Fig. 2(b). It shows that users are inclined to apply for job positions that share similar keywords in their sequential behaviors, indicating users tend to gravitate toward similar job positions even within the same category. Therefore, the users in recruitment scenarios tend to favor "similar items" compared to other recommendation scenarios, which verifies our motivation to capture the fine-grained similarity of candidates and job positions for bilateral person-job fit.

[1][Online]. Available: https://www.zhipin.com/
[2][Online]. Available: https://www.movielens.org/
[3][Online]. Available: https://jmcauley.ucsd.edu/data/amazon
[4][Online]. Available: https://www.dianping.com/

IV. The Proposed Method

In this section, we first present a quasi-metric space for bilateral person-job fit, which can meet the nature of the bilateral person-job fit scenario in a theoretical way. Then, we propose a quasi-metric learning framework to model the asymmetric relationship between candidates and job positions, Finally, we propose a relation-aware graph convolutional network to augment the proposed quasi-metric learning framework by exploring the heterogeneous relations among users.

A. Quasi-Metric Space for Person-Job Fit

Intuitively, the quasi-metric space can be seen as a generalization of the metric space, which can be defined as follows:

Definition 1: Let $\mathcal{S}$ be a nonempty set. The $(\mathcal{S}, q)$ is a quasi-metric space with an asymmetric distance function $q \colon \mathcal{S} \times \mathcal{S} \rightarrow \mathbb{R}$ if it satisfies the following conditions for all $s, t \in \mathcal{S}$, i.e.,

- $q(s,t) \geq 0$ (the non-negative axiom);
- $q(s,t) = q(t,s) = 0 \Leftrightarrow s = t$ (the identity axiom);
- $q(s,t) + q(t,z) \geq q(s,z)$ (the triangle inequality axiom).

Therefore, a metric space must be a quasi-metric space, but the converse is not necessarily true because a quasi-metric space does not need to satisfy the symmetry axiom (i.e., $q(s,t) = q(t,s)$) required in a metric space. In online recruitment scenarios, we use the distance function $q(\cdot,\cdot)$ to model the bilateral preference of candidates and jobs, allowing us to model the asymmetric relationship among users for bilateral person-job fit. For example, we can model a situation where a candidate c_j applies for a job j_k but gets rejection by $q(c_i, j_k) < \tau < q(j_k, c_i)$. Specifically, $q(c_i, j_k)$ denotes the preference of candidate c_i for the job j_k, and $q(j_k, c_i)$ represents the preference of the employer who posts job j_k for candidate c_i, where the close distance $q(c_i, j_k)$ indicates the high preference of c_i for j_k. τ denotes the threshold below which a candidate is willing to apply for a job and an employer is willing to accept a candidate's application. By considering such asymmetric relations, the quasi-metric learning framework can capture a more nuanced understanding of candidates' and employers' preferences and behavior, leading to more accurate and effective recommendation results.

However, constructing a quasi-metric space with an appropriate function q is a crucial but challenging problem for effective bilateral person-job fit. To align the quasi-metric space with the inherent nature of recruitment scenarios, we adopt two recruitment rules for the quasi-metric space construction, namely similarity rule and competitive rule for person-job fit:

- *Similarity rule:* a bilateral good job-person fit system should prioritize the similarity between candidates and job positions, e.g., providing software engineers with job positions related to software development.
- *Competitiveness rule:* a good bilateral job-person fit system should also consider the competitiveness of candidates and job positions in a two-way selection process. Specifically, candidates may prioritize job positions that offer higher salaries, benefits, and other perks, while employers may prefer candidates with impressive educational backgrounds and relevant experiences. For example, if two candidates c_1 and c_2 apply for the same job position j_1, and candidate c_1 is accepted while candidate c_2 is rejected, it suggests that candidate c_1 is more competitive than candidate c_2, even if they are both "similar" to job position j_1.

To effectively capture both these rules in the two-way selection process, we propose to model the mutual preferences between candidates and employers in a quasi-metric space $(\mathcal{S}, q)$ that satisfies the following condition:

$$q(s,t) - q(t,s) = w(s) - w(t), \tag{1}$$

where $q(s,t)$ denotes a distance measure to model the preference of s for t, and $w(\cdot) : \mathcal{S} \to \mathbb{R}$ denotes a real-valued function to measure the "competitive score" of a candidate or a job position within the same space. To demonstrate how (1) facilitates modeling the similarity and competitiveness rules for bilateral person-job fit, we present the following theorem about the proposed quasi-metric space.

Theorem 1: Every quasi-metric space $(\mathcal{S}, q)$ which satisfies (1) can be constructed based on a metric space $(\mathcal{S}, d)$ by

$$q(s,t) = \underbrace{d(s,t)}_{\text{Similarity}} + \frac{1}{2} \cdot \underbrace{[w(s) - w(t)]}_{\text{Competitiveness}}, \tag{2}$$

where $s, t \in \mathcal{S}$. $d(\cdot,\cdot)$ denotes the symmetric distance in the metric space, which measures the similarity between candidates and job positions. $w(\cdot)$ denotes a function that maps from $\mathcal{S}$ to a real value, measuring the "competitive score" of a candidate or a job position and satisfying with $(w(s) - w(t))/2 + d(s,t) \geq 0$.

Proof: For any quasi-metric space $(\mathcal{S}, q)$ that satisfies $q(s,t) - q(t,s) = w(s) - w(t)$, we define a symmetric function $d : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ based on $q(\cdot,\cdot)$:

$$d(s,t) = \frac{q(s,t) + q(t,s)}{2}.$$

We can prove that $(\mathcal{S}, d)$ is a metric space as follows:

- Non-negative axiom: $d(s,t) = [q(s,t) + q(t,s)]/2 \geq 0$;
- Symmetric axiom: $d(s,t) = [q(s,t) + q(t,s)]/2 = [q(t,s) + q(s,t)]/2 = d(t,s)$;
- Identity axiom: Sufficiency, if $d(s,t) = 0$ we have $q(s,t) = q(t,s) = 0$ as $q(s,t), q(t,s) \geq 0$. Necessity is apparently established;
- Triangle inequality axiom: $d(s,t) + d(t,z) = [q(s,t) + q(t,s) + q(t,z) + q(z,t)]/2 \geq [q(s,z) + q(z,s)]/2 = d(s,z)$.

Furthermore, we can reformulate (1) as

$$\begin{aligned}
&q(s,t) - q(t,s) = w(s) - w(t) \Leftrightarrow q(s,t) + w(t) = q(t,s) + w(s)\\
&\Leftrightarrow q(s,t) + q(s,t) + w(t) = q(s,t) + q(t,s) + w(s)\\
&\Leftrightarrow 2 \cdot q(s,t) + w(t) = 2 \cdot d(s,t) + w(s)\\
&\Leftrightarrow q(s,t) = d(s,t) + 1/2 \cdot [w(s) - w(t)].
\end{aligned}$$

Therefore, we can construct any quasi-metric space $(\mathcal{S}, q)$ that satisfies $q(s,t) - q(t,s) = w(s) - w(t)$ in a way of (2). Due to the non-negative axiom of the quasi-metric space, we have $d(s,t) + 1/2 \cdot (w(s) - w(t)) \geq 0$. □

This theorem states that every quasi-metric space $(\mathcal{S}, q)$ which satisfies (1) can be constructed in a way of (2) based on a metric space $(\mathcal{S}, d)$. First, we can model the *similarity rule* by measuring the similarity scores of candidates and job positions with a distance function $d(s,t) : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ in a metric space $(\mathcal{S}, d)$. The metric space satisfies the triangle inequality and symmetric axioms, which can capture the propagation of similarity among candidates and jobs in online recruitment scenarios [3]. Second, the *competitiveness rule* is easily established based on the quasi-metric distance, as shown in (2). Specifically, suppose candidate c_1 shares the same similarity w.r.t. job position j_1 and job position j_2 (i.e., $d(c_1, j_1) = d(c_1, j_2)$), and job position j_1's competitive score is higher than job position j_2's (i.e., $w(j_1) - w(j_2) > 0$), we have

$$\begin{aligned}
q(c_1, j_1) - q(c_1, j_2) = w(j_2) - w(j_1) < 0\\
\Leftrightarrow q(c_1, j_1) < q(c_1, j_2).
\end{aligned}$$

Therefore, candidate c_1 will prefer job position j_1 to job position j_2, which verifies the competitiveness rule in the proposed quasi-metric space.

To intuitively understand the proposed quasi-metric space, we present an illustrative example of quasi-metric space for online recruitment as shown in Fig. 3. In this quasi-metric space, the brown plane $x - y$ represents the metric space $(\mathcal{S}, d)$ which can

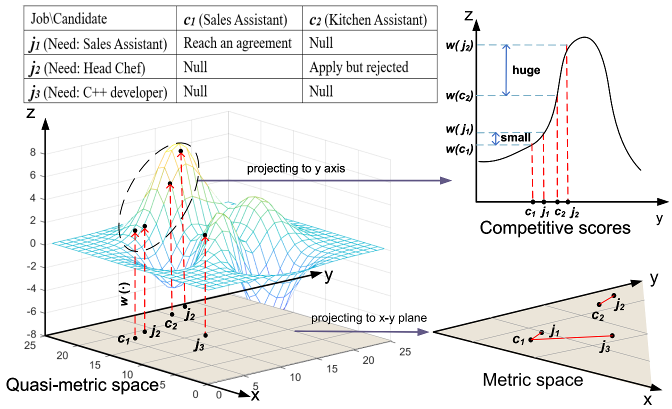


Fig. 3. An illustrative example of quasi-metric space for online recruitment.

be used to measure the similarity between two-sided users, while the curve interwoven by lines represents the competitive score $s(\cdot)$ mapping from the metric space $\mathcal{S}$ to real values $\mathbb{R}$. Suppose we have two candidates (c_1 and c_2) and three job positions (j_1, j_2 and j_3), where their interaction records are shown in the upper left side of Fig. 3. First, candidate c_1 and job position j_1 can reach an agreement because of the high similarity between them and the small gap between their competitive scores within the competitive market. Second, although candidate c_2 and job position j_2 meet the similarity rule for each other, i.e., matching the cook candidate with the chef job position, they may not reach an agreement due to the large competitive score gap between them. Specifically, candidate c_2 prefers job position j_2 and decides to apply for it, but he/she may get rejected due to being less competitive, possibly due to a lack of experience or inadequate skills. Third, there is no interaction between the two candidates (c_1 and c_2) and job position j_3 because of the huge dissimilarity between them.

In this paper, we introduce an easy and efficient implementation of the asymmetric distance $q(s,t)$ for quasi-metric space, which can align with the nature of online recruitment scenarios by considering both similarity and competitiveness rules.

Corollary 1: Given any function $w : \mathcal{S} \to \mathbb{R}$ on space $\mathcal{S}$, we can construct a quasi-metric function $q : \mathcal{S} \times \mathcal{S} \to \mathbb{R}$ based on the p-norm space where $p \geq 1$, i.e.,

$$q(s,t) = ||s-t||_p + \max(w(s) - w(t), 0), \tag{3}$$

where $||s-t||_p$ denote the p-norm distance between s and t.

Proof: First, the function $q(s,t)$ satisfies the similarity and competitiveness rules in (1) due to

$$\begin{aligned} q(s,t) - q(t,s) &= ||s-t||_p + \max(w(s) - w(t), 0) \\ &\quad - ||t-s||_p - \max(w(t) - w(s), 0) \\ &= \max(w(s) - w(t), 0) \\ &\quad - \max(w(t) - w(s), 0) \\ &= 1/2 \cdot [|w(s) - w(t)| + w(s) - w(t)] \\ &\quad - 1/2 \cdot [|w(t) - w(s)| + w(t) - w(s)] \\ &= w(s) - w(t). \end{aligned}$$

Second, we can use the function $q(s,t)$ to generate a quasi-metric space which satisfies the axioms of Definition 1.

- Non-negative axiom: $q(s,t) = ||s-t||_p + \max(w(s) - w(t), 0) \geq 0$;
- Identity axiom: If $q(s,t) = q(t,s) = 0$, we have $||s-t||_p + \max(w(s) - w(t), 0) = ||t-s||_p + \max(w(t) - w(s), 0) = 0$, therefore we have $s = t$ as $||s-t||_p = ||t-s||_p = 0$. Necessity is apparently established, i.e., $q(s,s) = 0$;
- Triangle inequality axiom:

$$\begin{aligned} q(s,t) + q(t,z) &= ||s-t||_p + \max(w(s) - w(t), 0) \\ &\quad + ||t-z||_p + \max(w(t) - w(z), 0) \\ &\geq ||s-z||_p + \max(w(s) - w(t), 0) \\ &\quad + \max(w(t) - w(z), 0) \\ &\geq ||s-z||_p + \max(w(s) - w(z), 0) = q(s,z). \end{aligned}$$

□

B. Quasi-Metric Learning Framework

1) Embedding for Candidates and Job Positions: To learn representations of candidates and job positions in the quasi-metric space, we first embed all candidates and job positions with two embedding matrices $P \in \mathbb{R}^{N \times d_e}$ and $Q \in \mathbb{R}^{M \times d_e}$, respectively. d_e denotes the dimension of latent space. Then, we employ the one-hot encoding to retrieve representations $P_i \in \mathbb{R}^{d_e}$ and $Q_k \in \mathbb{R}^{d_e}$ for candidate c_i and job position j_k,

$$P_i = Onehot(c_i) \cdot P, \quad Q_k = Onehot(j_k) \cdot Q,$$

where $Onehot(c_i) \in \mathbb{R}^{1 \times N}$ and $Onehot(j_k) \in \mathbb{R}^{1 \times M}$ denote the one-hot encoding of candidate c_i and job position j_k, respectively. To make use of the descriptive text associated with a candidate or job position, we adopt the BERT model to encode them into the text representations $W_t \in \mathbb{R}^{d_e}$ as in [2]. Specifically, we first maintain the original text order and place a unique token $[CLS]$ before the text, then we feed the combined sequence into the BERT model and use the output of token $[CLS]$ as the semantic representation of the descriptive text. To align the dimensions of the user and text representations, we use the PCA strategy to reduce the dimension of text representations W_t into d_e, i.e., $W'_t = PCA(W_t)$. To enhance the person-job fit modeling, we fuse the representations of candidates (or jobs) w.r.t. their text representations as the hybrid representations for quasi-metric learning, i.e.,

$$x_{c_i} = P_i + \alpha \cdot W'_{T_i}, \quad x_{j_k} = Q_k + \alpha \cdot W'_{T_k}, \tag{4}$$

where W'_{T_i} and W'_{T_k} denote the text representations of T_i and T_k respectively, after processing through BERT encoding and PCA reduction. T_i and T_k denote the descriptive text of candidate c_i and job position j_k. The parameter α is a learnable coefficient used to combine these two types of representations.

2) Learning Asymmetric and Symmetric Relations in Quasi-Metric Space: In real-world scenarios, both asymmetric and symmetric relationships are prevalent in the recruitment market. For example, asymmetric relations occur when a candidate applies for a job but receives a rejection, while symmetric relations occur when a candidate and an employer mutually agree on employment. Recognizing and modeling these diverse relationships is essential for accurate bilateral person-job fit, we propose learning asymmetric and symmetric relations in the quasi-metric space for online recruitment.

To learn the asymmetric relations between candidates and job positions, we employ a pair-wise loss function in the quasi-metric space inspired by the large margin nearest neighbor method [55]. The loss function is formulated as follows:

$$L_{\text{asym}} = \sum_{(c_i, j_k)\in\mathcal{A}_c \wedge (c_i, j'_k)\notin\mathcal{A}_c} [q(x_{c_i}, x_{j_k}) - q(x_{c_i}, x_{j'_k}) + \kappa]_+ + \sum_{(j_k, c_i)\in\mathcal{A}_j \wedge (j_k, c'_i)\notin\mathcal{A}_j} [q(x_{j_k}, x_{c_i}) - q(x_{j_k}, x_{c'_i}) + \kappa]_+, \quad (5)$$

where $\mathcal{A}_c = \mathcal{A}_c^{\rightarrow} \cup \mathcal{A}_c^{\leftrightarrow}$ denotes the action set of candidates when they apply or accept for a job interview, and $\mathcal{A}_j = \mathcal{A}_j^{\rightarrow} \cup \mathcal{A}_j^{\leftrightarrow}$ denotes the action set of employers when they invite candidates for a job interview or accept a job application from a candidate. (c_i, j'_k) and (j_k, c'_i) are the negative samples of (c_i, j_k) and (j_k, c_i), satisfying $(c_i, j'_k) \notin \mathcal{A}_c$ and $(j_k, c'_i) \notin \mathcal{A}_j$. κ denotes the margin constant between positive and negative samples. $[\cdot]_+ = \max(\cdot, 0)$ denotes the standard hinge loss.

To learn the symmetric relations when candidates and employers reach agreements, we first calculate their matching score (distance) by considering both candidate-to-job and job-to-candidate preferences to model their two-way selection process, i.e.,

$$\begin{aligned} m(c_i, j_k) &= q(x_{c_i}, x_{j_k}) + q(x_{j_k}, x_{c_i}) \\ &= ||x_{c_i} - x_{j_k}||_p + \max(w(x_{c_i}) - w(x_{j_k}), 0) \\ &\quad + ||x_{j_k} - x_{c_i}||_p + \max(w(x_{j_k}) - w(x_{c_i}), 0) \\ &= 2 \cdot \underbrace{||x_{c_i} - x_{j_k}||_p}_{similarity} + \underbrace{|w(x_{c_i}) - w(x_{j_k})|}_{co-equity}. \end{aligned} \quad (6)$$

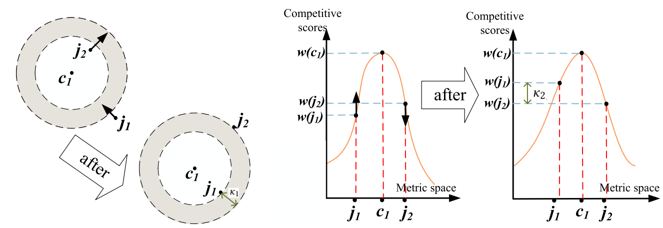
Intuitively, it suggests that candidates and job positions are more likely to come to an agreement when they are similar and co-equity to each other. Then, we adopt the triplet-wise loss in the quasi-metric distance, i.e.,

$$L_{\text{sym}} = \sum_{\substack{(c_i, j_k)\in\mathcal{A}_c^{\leftrightarrow} \wedge \\ (c'_i, j'_k)\in N^{\leftrightarrow}(c_i, j_k)}} [m(c_i, j_k) - m(c_i, j'_k) - m(c'_i, j_k) + \kappa]_+, \quad (7)$$

where $(c'_i, j'_k) \in N^{\leftrightarrow}(c_i, j_k)$ is the negative sample of (c_i, j_k) for triplet-wise loss function, which satisfies $N^{\leftrightarrow}(c_i, j_k) = \{(c'_i, j'_k) | (c_i, j'_k) \notin \mathcal{A}_c^{\leftrightarrow} \wedge (c'_i, j_k) \notin \mathcal{A}_c^{\leftrightarrow}\}$.

To leverage the rich asymmetric and symmetric relations for quasi-metric learning, we adopt the multi-task learning framework which combines both asymmetric and symmetric goals to jointly learn the representation of candidates and job positions, i.e.,

$$\min L_{\text{QML}} = L_{\text{asym}} + L_{\text{sym}} + \lambda_{\Theta}||\Theta||^2, \quad (8)$$



(A) Metric space for similarity, where candidate c_1 is more similar to job j_1 than job j_2.

(B) Score mapping for competitiveness, where job j_1 is more attractive than job j_2.

Fig. 4. An illustrative example of quasi-metric learning in online recruitment.

where Θ denotes all parameters need to be learned in the proposed model. λ_Θ is the regularization coefficient of the L2 norm $|| \cdot ||^2$.

To make an intuitive understanding of quasi-metric learning, we illustrate the model learning process resulting from the asymmetric and symmetric loss functions. For the *asymmetric loss function* as in (5), we can divide it into two parts which correspond to the similarity and competitiveness rules as follows:

$$\begin{aligned} &[q(x_{c_1}, x_{j_1}) - q(x_{c_1}, x_{j_2}) + \kappa]_+ \\ &\le \underbrace{[d(x_{c_1}, x_{j_1}) - d(x_{c_1}, x_{j_2}) + \kappa_1]_+}_{Reason1} \\ &+ \underbrace{[\max(w(x_{c_1}) - w(x_{j_1}), 0) - \max(w(x_{c_1}) - w(x_{j_2}), 0) + \kappa_2]_+}_{Reason2}, \end{aligned} \quad (9)$$

where $\kappa_1 + \kappa_2 = \kappa$. Specifically, suppose that candidate c_1 applies for job position j_1 but not for job position j_2, there may be two reasons for this behavior: *Reason 1: Job position* j_1 *is more similar to candidate* c_1 *than job position* j_2. In this case, as shown in Fig. 4(a), job j_2 will move away from candidate c_1 and job j_1 will move closer to candidate c_1 until there is a safe margin κ_1 in the metric space. *Reason 2: Job position* j_1 *is more attractive than the similar job position* j_2. In this case, as shown in Fig. 4(b), job j_1's competitive score will increase while job j_2's competitive score will decrease until there is a safe margin κ_2. Both of these reasons can be captured according to (9) in the quasi-metric learning. For the *symmetric loss function* as in (7), if candidate c_1 and job position j_1 reach an agreement, they are very likely to be similar and co-equality according to (6), which is consistent with the inherent nature of recruitment scenarios.

C. Relation-Aware Graph Convolution Network Based Quasi-Metric Learning Framework

To capture the complex and diverse mutual effects among users, we propose to explore heterogeneous relations among users for bilateral person-job fit. Specifically, we organize candidates, jobs, and their relations in a unified graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. The nodes $\mathcal{V}$ in the graph represent candidates and jobs, i.e., $\mathcal{V} = \{\mathcal{C} \cup \mathcal{J}\}$. The edges $\mathcal{E}$ are constructed based on the different-typed interactions between candidates and jobs, i.e., $\mathcal{E} = \mathcal{A}_c^{\rightarrow} \cup \mathcal{A}_c^{\leftarrow} \cup \mathcal{A}_j^{\rightarrow} \cup \mathcal{A}_j^{\leftarrow} \cup \mathcal{A}_c^{\leftrightarrow} \cup \mathcal{A}_j^{\leftrightarrow}$. Then, we propose a

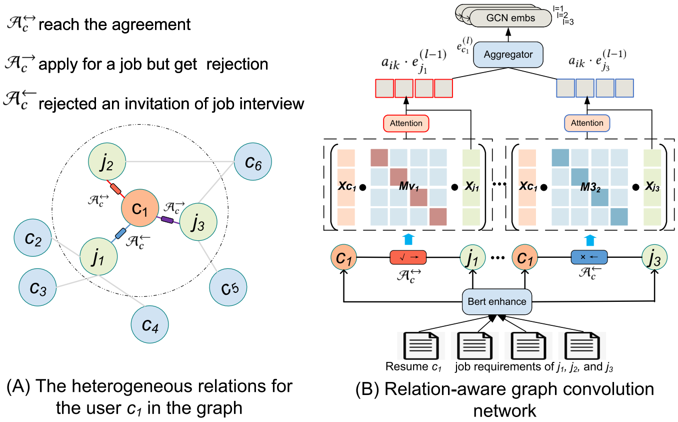


Fig. 5. The architecture of relation-aware graph convolution network.

relation-aware graph convolution network to capture the mutual effects of candidates and job positions with their different-typed interactions.

To explore the heterogeneous relationships between candidates and job positions, we embed each type of relationship into distinct relation-aware diagonal matrices, e.g., modeling the v-typed relation by a matrix $M_v \in \mathbb{R}^{d_e \times d_e}$. With nodes and edges in the graph, we propose a relation-aware GCN method to model the triplet of candidate, job, and their different-typed interaction as shown in Fig. 5. Unlike traditional GCN methods that primarily focus on homogeneous relationships in a graph, our relation-aware GCN model is capable of modeling heterogeneous effects among nodes through message propagation based on different-typed interactions. Specifically, with the $(l-1)$th layer representation of graph convolution of nodes, we can obtain the next layer (lth) representation of candidate c_i's by aggregating his/her neighborhood $N(c_i)$ (i.e., job positions) in the unified heterogeneous relation graph,

$$e_{c_i}^{(l)} = e_{c_i}^{(l-1)} + \sigma\left(\sum_{j_k \in N(c_i)} a_{ik} \cdot e_{j_k}^{(l-1)}\right),$$

where a_{ik} controls the weights of neighbors for aggregation, which can be formulated as

$$a_{ik} = \frac{\exp(-[e_{c_i}^{(l)} - e_{j_k}^{(l)}]^\top \cdot M_{v(i,k)} \cdot [e_{c_i}^{(l)} - e_{j_k}^{(l)}])}{\sum_{j_{k'} \in N(c_i)} \exp(-[e_{c_i}^{(l)} - e_{j_{k'}}^{(l)}]^\top \cdot M_{v(i,k')} \cdot [e_{c_i}^{(l)} - e_{j_{k'}}^{(l)}])}, \tag{10}$$

where $M_{v(i,k)} \in \mathbb{R}^{d_e \times d_e}$ denotes the relation-aware matrix to model the v-type interaction between candidate c_i and job j_k. Therefore, the attention weight a_{ik} measures the "distance" between candidate c_i and job position j_k measured by the v-typed relationship. The closer the distance between the adjacent nodes, the more information from the neighbor node is propagated into the target node. Inspired by the convolution scheme in LightGCN [21], we adopt the identity activation function $\sigma(x) = x$ without linear transformations during the message passing. The final GCN representation for candidate c_i is obtained by averaging the representations from L layers as follows:

$$x_{c_i}^{gcn} = AveragePool([e_{c_i}^{(0)}, \ldots, e_{c_i}^{(L)}]),$$

where $e_{c_i}^{(0)} = x_{c_i}$. Symmetrically, the relation-aware GCN representation for job position j_k can be constructed similarly,

$$x_{j_k}^{gcn} = AveragePool([e_{j_k}^{(0)}, \ldots, e_{j_k}^{(L)}]),$$

$$e_{j_k}^{(l)} = e_{j_k}^{(l-1)} + \sigma\left(\sum_{c_i \in N(j_k)} a_{ki} \cdot e_{c_i}^{(l-1)}\right),$$

$$a_{ki} = \frac{\exp(-[e_{j_k}^{(l)} - e_{c_i}^{(l)}]^\top \cdot M_{v(k,i)} \cdot [e_{j_k}^{(l)} - e_{c_i}^{(l)}])}{\sum_{c_{i'} \in N(j_k)} \exp(-[e_{j_k}^{(l)} - e_{c_{i'}}^{(l)}]^\top \cdot M_{v(k,i')} \cdot [e_{j_k}^{(l)} - e_{c_{i'}}^{(l)}])}. \tag{11}$$

Finally, we combine the proposed relation-aware GCNs-based (rG) representations with the proposed Quasi-Metric Learning (QML) framework for bilateral person-job fit, i.e.,

$$\min L_{\text{rGQML}} = L_{\text{asym}}^{gcn} + L_{\text{sym}}^{gcn} + \lambda_\Theta ||\Theta||^2, \tag{12}$$

where L_{asym}^{gcn} and L_{sym}^{gcn} denote the asymmetric and symmetric loss functions by adopting $x_{c_i} = x_{c_i}^{gcn}$ and $x_{j_k} = x_{j_k}^{gcn}$ as in (4). Θ denotes all parameters need to be learned in the proposed model, including the embedding of candidates P, the embedding of jobs Q, combination parameter α, and relation matrices $\{M_v | v \in \mathcal{V}\}$. λ_Θ is the regularization coefficient of the L2 norm $|| \cdot ||^2$. As the objective function is differentiable, we optimize it by stochastic gradient descent (SGD) and adaptively adjust the learning rate by AdamGrad, which can be automatically implemented by PyTorch. The pseudo-code of rGQML is shown in Algorithm 1.

Algorithm 1: Quasi Metric Learning Method rGQML for Online Recruitment Recommendation.

Require: Candidate set $\mathcal{C}$, job set $\mathcal{J}$, their associated text documents and interactions, initial learning rate η, and regularization coefficient λ_Θ.

Randomly initialize all parameters Θ need to be learned.
Construct a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ for candidates and job positions, and gain their embeddings (i.e., $x_{c_i}^{gcn}$ and $x_{j_k}^{gcn}$) based on the relation-aware GCN method.
repeat
  Randomly draw three samples from different-typed interactions, i.e., $(c_{i_1}, j_{k_1}) \in \mathcal{A}_c$, $(j_{k_2}, c_{i_2}) \in \mathcal{A}_j$, and $(c_{i_3}, j_{k_3}) \in \mathcal{A}_c^{\leftrightarrow}$, and randomly draw corresponding negative samples, i.e., $(c_i, j_k') \notin \mathcal{A}_c$, $(j_k, c_i') \notin \mathcal{A}_j$, and $(c_i', j_k') \in N^{\leftrightarrow}(c_i, j_k)$, respectively.
  Update the parameters: $\Theta \leftarrow \Theta - \eta \cdot \frac{\partial L_{\text{rGQML}}}{\partial \Theta}$;
  Adjust the learning rate η by AdamGrad;
until Convergence
return Model parameters Θ

V. Experiments

In this section, we aim to evaluate the performance and effectiveness of the proposed method rGQML. Specifically, we

TABLE II
STATISTICS OF THE EXPERIMENTAL DATASETS

Datasets	Designs	Sales	Tech	Zhilian
#Seekers	12,290	15,854	56,634	4,500
#Jobs	9,143	12,772	48,090	22,467
#Seekers-action	1,200,590	213,860	3,749,807	12,685
#Job-action	76,287	2,077,560	907,087	39,888
#Agreement	166,270	145,066	925,193	15,774

conduct several experiments to study the following research questions:

- *RQ1:* Does the proposed method rGQML outperform state-of-the-art methods for bilateral person-job fit?
- *RQ2:* Does the proposed quasi-metric learning in rGQML can effectively model asymmetric relationships between candidates and job positions for bilateral person-job fit?
- *RQ3:* Does the proposed method rGQML benefit from capturing the fine-grained similarity between candidates and job positions by leveraging the triangle inequality in a quasi-metric space?
- *RQ4:* Does the proposed method rGQML benefit from multi-task learning by considering both asymmetric and symmetric goals for bilateral person-job fit?
- *RQ5:* Do the proposed relation-aware GCNs in rGQML can effectively capture heterogeneous relations between candidates and jobs to enhance bilateral person-job fit?
- *RQ6:* How do different configurations of key hyper-parameters impact the performance of the proposed method rGQML?

A. Experimental Setup

1) Datasets: We evaluated the proposed method rGQML on four real-world datasets provided by two popular online recruitment platforms. The first three datasets were collected from the Boss Zhipin platform[5] during 106 days of real online logs of person-job fit in the designer, sales, and technology industries, respectively. The last dataset Zhilian was collected from the Zhaopin platform,[6] which is publicly available in Tianchi.[7] These datasets contain rich interactions between candidates and employers, such as the actions (i.e., a candidate applying for a job, an employer reaching out to a candidate) and feedback (i.e., accept, reject). In addition, all these datasets also contain text document information, e.g., the resume of candidates and the description of job positions. The characteristics of these datasets are summarized in Table II.

2) Evaluation Methodology and Metrics: For the first three datasets, We split the interaction records of each dataset chronologically into training, validation, and test sets. Following prior study [2], we used the first 84 days records as the train set, and we equally divided the following 22 days records into the validation set and the test set. As there is no timestamp information in the Zhilian dataset, we equally divided it into training, validation, and test sets. Experimental results were recorded as the average of five runs with different random initialization of model parameters.

To evaluate the performance of different methods, we adopted four widely used evaluation metrics for top-n recommendation [56], including recall ($r@n$), precision ($p@n$), normalized discounted cumulative gain ($ndcg@n$), mean reciprocal rank, and (mrr), where n was set as 5 empirically. We sampled 20 negative instances for each positive instance from the user's non-interacted records by following the previous study [2]. We utilized these metrics to simultaneously evaluate the bilateral person-job fit tasks, i.e., ranking job positions for candidates and ranking candidates for employers, which was consistent with actual online recruitment scenarios.

3) Baselines: We took the following state-of-the-art methods as the baselines, including content-based methods [39], collaborative filtering based methods [13], [17], [57], and hybrid methods [1], [2], [21], [53] for bilateral person-job fit.

- *BPJFNN [39]:* It adopts a hierarchical attentional RNN model to learn the word-level semantic representation of candidates' resumes and jobs' descriptions.
- *BPR [57]:* It learns the latent representations of users and items based on the pairwise loss with the sigmoid activation function.
- *NCF [17]:* It enhances collaborative filtering with deep neural networks, which adopt an MLP to explore the non-linear interaction between users and items.
- *LFRR [13]:* It adopts matrix factorization based on separate candidate-to-job and candidate-to-job bipartite graphs, and then combines them for reccomendation.
- *LightGCN$_{bert}$:* It enhances the LightGCN [21] with text information about candidates and jobs for person-job fit.
- *IPJF [53]:* It adopts a multi-task optimization approach to learn the perspectives/intentions of employers and candidates based on their text information.
- *PJFFF [1]:* It fuses the representations of the explicit and implicit intentions of candidates and employers based on their historical application records.
- *DPGNN [2]:* It designs a unified dual-perspective interaction graph to model the active and passive aspects of candidates and jobs.

4) Implementation Details: We adopted a simple linear regression mapping function $w(t) = W \cdot t$ for competitive score modeling, where $W \in \mathbb{R}^{1\times d}$. For a fair comparison, all methods were optimized by the Adam optimizer with the same latent space dimension (i.e., 64), batch size (i.e., 1024), learning rate (i.e., 0.001), and regularization coefficient (i.e., 0.0001) for all the experiments. We adopted the 3-layer of relational-aware GCNs and 2-norm distance in the quasi-metric space. For other special hyper-parameters in baseline models, we carefully searched for their best performance. Early stopping was used with the patience of 50 epochs.

B. Model Comparison (RQ1)

Table III shows the performance of different methods for bilateral person-job fit. To make the table more notable, we bold

[5][Online]. Available: https://www.zhipin.com/
[6][Online]. Available: https://www-al.zhaopin.com/
[7][Online]. Available: https://tianchi.aliyun.com/dataset/31623

TABLE III
PERFORMANCE OF THE PROPOSED AND BASELINE METHODS FOR PERSON-JOB FIT

Dataset	Direction	Metric	BPJFNN	BPR	NCF	LFRR	$\text{LGCN}_{\text{bert}}$	IPJF	PJFFF	DPGNN	rGQML	imprv.
Designs	To candidates	r@5	0.3452	0.3720	0.5266	0.5430	0.5612	0.4048	0.5110	0.5714	**0.6794***	18.90%
		p@5	0.1016	0.1136	0.1684	0.1690	0.1778	0.1208	0.1614	0.1800	**0.2178***	21.00%
		ndcg@5	0.3094	0.3488	0.5196	0.5128	0.5416	0.3682	0.4962	0.5496	**0.6408***	16.59%
		mrr	0.3052	0.3464	0.4984	0.4930	0.5184	0.3608	0.4734	0.5256	**0.6056***	15.22%
	To employers	r@5	0.2984	0.3166	0.4362	0.4854	0.4696	0.3416	0.4800	0.4956	**0.5640***	13.80%
		p@5	0.1148	0.1386	0.1954	0.2268	0.2160	0.1484	0.2224	0.2362	**0.2888***	22.27%
		ndcg@5	0.3008	0.3750	0.4958	0.5492	0.5300	0.4004	0.5502	0.5898	**0.6986***	18.45%
		mrr	0.2880	0.3662	0.4766	0.5312	0.5116	0.3816	0.5182	0.5636	**0.6720***	19.23%
Sales	To candidates	r@5	0.2504	0.3514	0.4170	0.4560	0.4622	0.1872	0.4076	0.4724	**0.5862***	24.09%
		p@5	0.0730	0.1106	0.1320	0.1436	0.1470	0.0550	0.1288	0.1498	**0.1886***	25.90%
		ndcg@5	0.2088	0.3408	0.4010	0.4302	0.4446	0.1604	0.3888	0.4548	**0.5652***	24.27%
		mrr	0.2196	0.3410	0.3906	0.4140	0.4296	0.1860	0.3774	0.4396	**0.5376***	22.29%
	To employers	r@5	0.2696	0.4058	0.4566	0.4866	0.4950	0.2632	0.4534	0.5068	**0.5422***	6.99%
		p@5	0.0842	0.1388	0.1584	0.1744	0.1758	0.0864	0.1566	0.1834	**0.1950***	6.32%
		ndcg@5	0.2344	0.4016	0.4494	0.4886	0.4868	0.2490	0.4422	0.5188	**0.5532***	6.63%
		mrr	0.2430	0.3880	0.4336	0.4682	0.4676	0.2588	0.4254	0.4992	**0.5318***	6.53%
Tech	To candidates	r@5	0.4966	0.4504	0.5960	0.6276	0.6508	0.5758	0.6430	0.7034	**0.8290***	17.86%
		p@5	0.1640	0.1576	0.2350	0.2442	0.2458	0.2028	0.2472	0.2662	**0.3288***	23.52%
		ndcg@5	0.4226	0.4352	0.6428	0.6594	0.6782	0.5404	0.6744	0.7234	**0.8350***	15.43%
		mrr	0.3822	0.4118	0.6266	0.6462	0.6528	0.5008	0.6484	0.6950	**0.8090***	16.40%
	To employers	r@5	0.4436	0.3896	0.5714	0.5862	0.6038	0.5024	0.6324	0.6334	**0.7050***	11.30%
		p@5	0.1892	0.1872	0.3024	0.3182	0.3240	0.2444	0.3362	0.3406	**0.4020***	18.03%
		ndcg@5	0.4590	0.4762	0.6598	0.6750	0.7250	0.5782	0.7344	0.7542	**0.8364***	10.90%
		mrr	0.4152	0.4498	0.6422	0.6654	0.6998	0.5334	0.7038	0.7288	**0.8194***	12.43%
Zhilian	To candidates	r@5	0.1290	0.0650	0.4874	0.5399	0.5922	0.1481	0.3438	0.6196	**0.6867***	10.83%
		p@5	0.0313	0.0169	0.1394	0.1552	0.1679	0.0376	0.0949	0.1795	**0.1993***	11.03%
		ndcg@5	0.0883	0.0492	0.4224	0.4617	0.5055	0.1104	0.2851	0.5359	**0.5960***	11.21%
		mrr	0.1171	0.0735	0.3910	0.4234	0.4704	0.1326	0.2793	0.4986	**0.5544***	17.86%
	To employers	r@5	0.1962	0.0733	0.6477	0.7133	0.7645	0.2189	0.5108	0.7732	**0.8000***	3.47%
		p@5	0.0436	0.0165	0.1536	0.1696	0.1811	0.0493	0.1192	0.1845	**0.1902***	3.09%
		ndcg@5	0.1285	0.0492	0.5625	0.6149	0.6512	0.1497	0.3963	0.6689	**0.6920***	3.45%
		mrr	0.1508	0.0731	0.5363	0.5815	0.6137	0.1642	0.3777	0.6349	**0.6532***	2.88%

* indicates that the improvements are significant at the level of 0.01 with paired t-test.

the best results for each dataset with one specific evaluation metric. From the experimental results, we can get the following conclusions. First, the proposed method rGQML consistently outperforms all baseline methods in all cases, demonstrating the effectiveness of the proposed method. Second, the hybrid methods (e.g., DPGNN) that consider both interaction and text information perform well in most cases, emphasizing the importance of exploring both text descriptions and interactions of candidates and jobs. Third, DPGNN and LightGCN$_{bert}$ show good performance among the baseline methods, indicating the superiority of exploring structural information in the candidate-and-job graph. Fourth, the method LFRR, which adopts two simple matrix factorization techniques [57] to model the candidate-to-job and employer-to-candidate preferences, achieves competitive performance among baseline methods, which indicates the necessity of modeling two-way selection in reciprocal person-job fit. Finally, it should be noted that the content-based method BPJFNN exhibits poor performance compared to others. This can be attributed to the low quality of textual descriptions of candidates and jobs, which may contain default, meaningless, or erroneous information.

C. Ablation Studies

To evaluate the effectiveness of the module design in the proposed method rGQML, we compare it to several special cases which include:

- *woAsys:* We replace the asymmetric distance with symmetric Euclidean metric distance, i.e., $q(s,t) = d(s,t)$, which cannot capture the asymmetric relationship between candidates and jobs. Also, the ablation method fails to model the competitiveness rule in person-job fit scenarios as $w(s) = w(t) = 0$.
- *woTri:* We remove the triangle inequality property in quasi-metric space. Specifically, we replace (3) with $q(s,t) =< s,t > + \max(W \cdot t - W \cdot s, 0)$, i.e., adopting the similarity measure based on the popular matrix factorization [16], which fails to model the similarity rule in person-job fit scenarios.
- *woMT:* We eliminate the multi-task learning in the proposed method which aims to model both asymmetric and symmetric relations between candidates and employers. Specifically, we only conduct the symmetric learning $L_{\text{symmetric}}$ for bilateral person-job fit.

TABLE IV
PERFORMANCE OF VARIANT rGQML FOR ABLATION STUDIES

Dataset	Direction	Method	r@5	p@5	ndcg@5	mrr
Designs	To candidates	woAsys	0.6116	0.1951	0.5839	0.5539
		woMT	0.5786	0.1756	0.5073	0.4713
		woTri	0.5617	0.1732	0.5118	0.4841
		woRGCN	0.6207	0.1984	0.5983	0.5717
		woHetero	0.6583	0.2062	0.6011	0.5603
		rGQML	**0.6794**	**0.2178**	**0.6408**	**0.6056**
	To employers	woAsys	0.5339	0.2683	0.6588	0.6336
		woMT	0.4678	0.2164	0.5524	0.5235
		woTri	0.4864	0.2211	0.5464	0.5196
		woRGCN	0.4934	0.2423	0.6116	0.5914
		woHetero	0.5228	0.2613	0.6499	0.6244
		rGQML	**0.5640**	**0.2888**	**0.6986**	**0.6720**
Sales	To candidates	woAsys	0.4963	0.1579	0.4772	0.4595
		woMT	0.3877	0.1165	0.3477	0.3404
		woTri	0.4796	0.1485	0.4430	0.4272
		woRGCN	0.4567	0.1450	0.4497	0.4390
		woHetero	0.5227	0.1625	0.4835	0.4609
		rGQML	**0.5862**	**0.1886**	**0.5652**	**0.5376**
	To employers	woAsys	0.4862	0.1732	0.4906	0.4739
		woMT	0.3800	0.1268	0.3592	0.3521
		woTri	0.4934	0.1737	0.4888	0.4687
		woRGCN	0.4720	0.1700	0.4839	0.4650
		woHetero	0.5239	0.1847	0.5190	0.4971
		rGQML	**0.5422**	**0.1950**	**0.5532**	**0.5318**

– *woRGCN:* We eliminate the relation-aware graph convolution network in method rGQML, which degenerates into method QML as in (8) for bilateral person-job fit.
– *woHetero:* We eliminate the heterogeneous relations in the graph. Specifically, we replace the relation-aware matrices $M_{v(\cdot)}$ with the identity matrix E in (10) and (11), which fails to capture the heterogeneous mutual effects of candidates and employers within their different-typed behaviors.

Table IV shows the performance of the proposed method rGQML and ablation methods including woAsys, woTri, woMT, woRGCN, and woHetero. From the experimental results, we can get the following conclusions:

- RQ2: The proposed method rGQML consistently outperforms the variant **woAsys** in all cases, which highlights the superiority of the proposed quasi-metric learning framework for bilateral person-job fit. Specifically, the proposed method rGQML benefits from modeling the competitiveness rule in the recruitment market with quasi-metric learning. Compared to the existing metric learning methods which only model the symmetric relations of users, the proposed method rGQML can model their asymmetric relations with the similarity rule and competitiveness rule based on quasi-metric learning for bilateral person-job fit.
- RQ3: The proposed method rGQML shows significant improvement to the variant **woTri** in all cases. It indicates that the triangle inequality property can capture the fine-grained similarity propagation between candidates and jobs for bilateral person-job fit. Specifically, the statistical analysis in Section III-B indicates that the users tend to favor the "similar items" in recruitment scenarios, and the ablation experiment verifies the effectiveness of the proposed method for capturing similarity propagation among candidates and jobs.
- RQ4: The variant **woMT** performs badly among all methods in most cases, indicating the necessity of conducting multi-task learning to capture both asymmetric and symmetric relationships between candidates and employers. Specifically, besides the symmetric agreement relations, it is necessary to learn from users' asymmetric relationships to explore the rich similarity and competitiveness information between candidates and job positions.
- RQ5: The variant **woRGCN**, also known as method QML, performs badly among these ablation methods, which demonstrates the importance of enhancing the proposed quasi-metric learning framework with the relation-aware graph convolution network. In addition, the proposed method rGQML outperforms the variant **woHetero** in all cases, indicating that the proposed method benefits from capturing heterogeneous relations between candidates and jobs across different-typed interactions. Specifically, the relation-aware GCNs module enables the capturing of how users influence others and how they are influenced by their neighbors within different-typed interactions. The attention weights of the relation-aware GCNs measure the "distance" among users based on the relation-aware matrix, which controls the information diffusion in the graph.

D. Hyper-Parameter Study (RQ6)

There are two key parameters for the proposed method rGQML, including the number of layers L for the relation-aware GCNs and the norm p for the quasi-metric space. In this subsection, we investigate how these parameters influence the performance of the proposed method rGQML.

First, we investigate how the number of GCN layers L influences the performance of the proposed method. Fig. 6 shows the performance of rGQML and DPGNN with various number of GCN layers. It shows that the proposed method rGQML outperforms the baseline model DPGNN with different numbers of GCN layers. For the proposed method rGQML, it indicates that too large or too small L layers may degrade the performance of rGQML. With consideration of both predictiveness and computational complexity, we suggest setting a moderate L for the proposed method. Second, we explore how the p-norm in the quasi-metric space influences the proposed method. Fig. 6(b) shows that the proposed method rGQML performs the best with the 2-norm (i.e., Euclidean distance) in all cases. Therefore, we suggest using the Euclidean distance to capture the similarity ropagation in the proposed method.

E. Case Study

As shown in the ablation study, the proposed method rGQML can model heterogeneous relations between candidates and job position for better recommendation results. To investigate whether rGQML can capture our claims at the case level, we conducted a case study between the proposed method rGQML and the baseline model DPGNN to illustrate why rGQML can

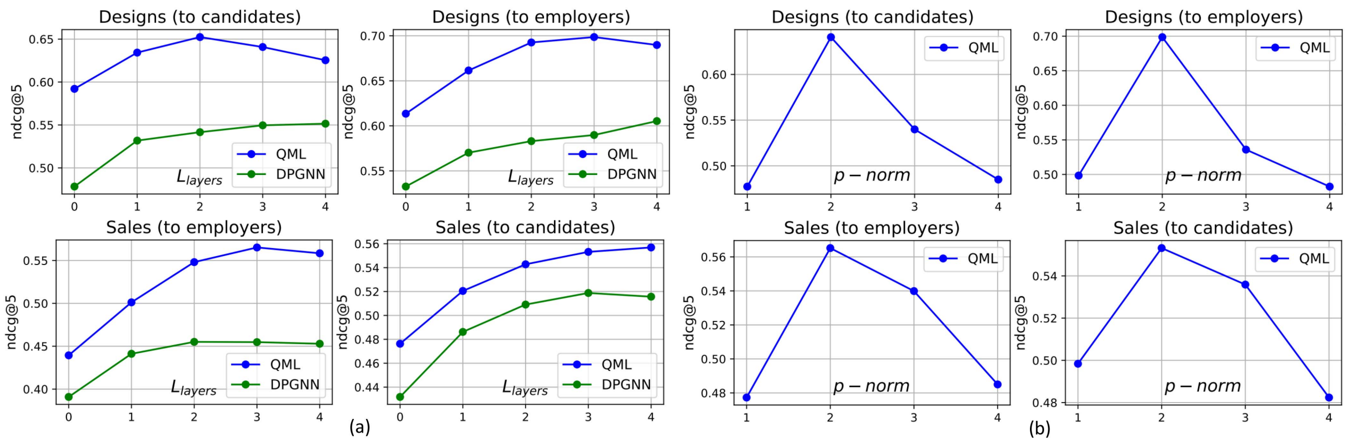


Fig. 6. (a) Performance comparison of rGQML and DPGNN w.r.t. different number of layers. (b) Performance of rGQML w.r.t. different p-norm.

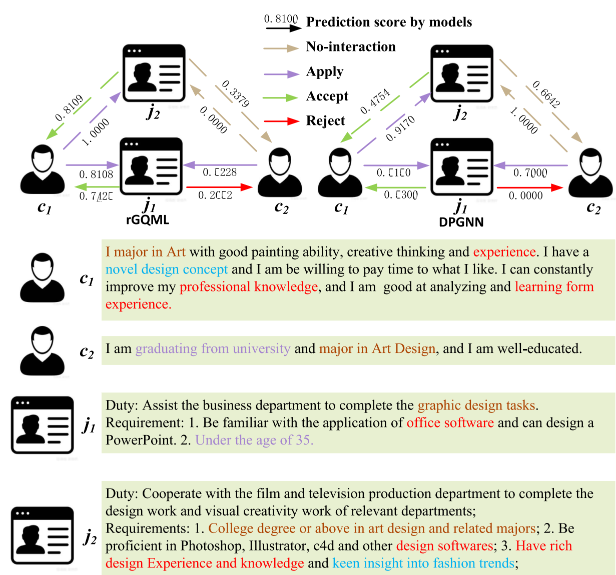


Fig. 7. Case studies demonstrate how the proposed method rGQML contributes to improved recommendation results. The solid arrow and the dashed arrow denote the relations in the train set and the test set, respectively. The numbers around the arrows denote the prediction scores based on the proposed method rGQML and method DPGNN. The colored text denotes the matching text between candidates' resumes and jobs' descriptions.

perform better for bilateral person-job fit. Specifically, we sampled two candidates (c_1, c_2) and two job positions (j_1, j_2) as shown in Fig. 7. First, both candidates c_1 and c_2 applied for job position j_1, but they received acceptance and rejection feedback, respectively. Second, we sampled a similar job position j_2 w.r.t. job position j_1. We predicted the person-job fit relations between the two candidates and the job position j_2, which were then verified based on the test set.

Fig. 7 depicts the prediction scores for person-job fit between candidates and job positions based on the proposed method rGQML and the baseline model DPGNN. As the scores predicted by rGQML and DPGNN are based on distance and similarity measures respectively, we transform their values using max-min normalization to make them comparable, where the larger score indicates the higher preference.

On the one hand, rGQML and DPGNN share similar predictions on the relations between these two candidates and job positions j_1, which can capture the relations in the train set. On the other hand, there is a significant discrepancy between the rGQML and DPGNN methods in predicting whether candidate c_2 will apply for the job position j_2. The method DPGNN gives a positive response to this question because it believes that candidate c_2 may apply for a similar job position j_2. However, in the real-world scenario, there is no interaction between them, and this is wrongly predicted by the method DPGNN. The main reason for candidate c_2 showing less interest in job j_2 lies in the fact that he has received negative feedback from job j_1, which is not captured by the method DPGNN. The proposed method rGQML can explore the heterogeneous relations and model the competitiveness rule in the recruitment market, resulting in a better prediction in this case. Specifically, job position j_2 has higher requirements than job position j_1, such as work experience at the semantic level, which makes candidate c_2 be less confident to be accepted by the senior job j_2. Candidate c_1 has higher competitiveness compared to candidate c_2, such as rich experience and professional knowledge, making it easier for candidate c_1 to be accepted. All of these facts are effectively captured by the proposed method rGQML.

VI. Conclusion

In this paper, we propose a quasi-metric learning framework for bilateral person-job fit, which can capture the similarity propagation between candidates and job positions while modeling their asymmetric relations. We propose to capture the fine-grained similarity between candidates and job positions by a quasi-metric space, and design an asymmetric measure to model the two-way selection process of candidates and employers. To model the mutual effects of two-sided users on each other, we organize users and their heterogeneous relations into a unified graph, and propose a relation-aware graph convolution network to capture the different mutual effects among them. The proposed method outperforms state-of-the-art baseline methods, which demonstrates the superiority of capturing both the similarity and competitiveness rules for bilateral person-job fit. The ablation studies show the importance of each component of the proposed method, including triangle inequality, asymmetric distance, multi-task learning, and relation-aware GCNs. The case studies further illustrate the superiority of the proposed method in capturing the heterogeneous relations and competitiveness rule for person-job fit. In the future, we hope to further explore the sequential patterns of candidates' and employers' behaviors within the quasi-metric learning framework, including designing a more delicate sequence mining module for the quasi-metric learning framework and predicting long-term user behaviors.

References

[1] J. Jiang, S. Ye, W. Wang, J. Xu, and X. Luo, "Learning effective representations for person-job fit by feature fusion," in *Proc. ACM Int. Conf. Inf. Knowl. Manage.*, 2020, pp. 2549–2556.

[2] C. Yang, Y. Hou, Y. Song, T. Zhang, J.-R. Wen, and W. X. Zhao, "Modeling two-way selection preference for person-job fit," in *Proc. ACM Conf. Recommender Syst.*, 2022, pp. 102–112.

[3] C.-K. Hsieh, L. Yang, Y. Cui, T.-Y. Lin, S. Belongie, and D. Estrin, "Collaborative metric learning," in *Proc. Int. Conf. World Wide Web*, 2017, pp. 193–201.

[4] Y. Tay, L. Anh Tuan, and S. C. Hui, "Latent relational metric learning via memory-based attention for collaborative ranking," in *Proc. Int. Conf. World Wide Web*, 2018, pp. 729–739.

[5] M. Li et al., "Symmetric metric learning with adaptive margin for recommendation," in *Proc. AAAI Conf. Artif. Intell.*, 2020, pp. 4634–4641.

[6] X. Zhou, D. Liu, J. Lian, and X. Xie, "Collaborative metric learning with memory network for multi-relational recommender systems," in *Proc. Int. Joint Conf. Artif. Intell.*, 2019, pp. 4454–4460.

[7] Z. Liu, X. Wang, Y. Ma, and X. Yang, "Relational metric learning with high-order neighborhood interactions for social recommendation," *Knowl. Inf. Syst.*, vol. 64, no. 6, pp. 1525–1547, 2022.

[8] R. Matsui, T. Naito, S. Yaginuma, and K. Nakata, "Confident collaborative metric learning," in *Proc. Int. Conf. Data Mining Workshops*, 2021, pp. 246–253.

[9] C. Park, D. Kim, X. Xie, and H. Yu, "Collaborative translational metric learning," in *Proc. Int. Conf. Data Mining*, 2018, pp. 367–376.

[10] J. M. McNamara and E. Collins, "The job search problem as an employer–candidate game," *J. Appl. Probab.*, vol. 27, no. 4, pp. 815–827, 1990.

[11] L. Xia, Y. Xu, C. Huang, P. Dai, and L. Bo, "Graph meta network for multi-behavior recommendation," in *Proc. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2021, pp. 757–766.

[12] A. Fishbach, T. Eyal, and S. R. Finkelstein, "How positive and negative feedback motivate goal pursuit," *Social Pers. Psychol. Compass*, vol. 4, no. 8, pp. 517–530, 2010.

[13] J. Neve and I. Palomares, "Latent factor models and aggregation operators for collaborative filtering in reciprocal recommender systems," in *Proc. ACM Conf. Recommender Syst.*, 2019, pp. 219–227.

[14] J. Neve and I. Palomares, "Hybrid reciprocal recommender systems: Integrating item-to-user principles in reciprocal recommendation," in *Proc. Web Conf.*, 2020, pp. 848–853.

[15] B. Fu et al., "Market-aware dynamic person-job fit with hierarchical reinforcement learning," in *Proc. Int. Conf. Database Syst. Adv. Appl.*, Springer, 2022, pp. 697–705.

[16] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, Aug. 2009.

[17] X. He, L. Liao, H. Zhang, L. Nie, X. Hu, and T.-S. Chua, "Neural collaborative filtering," in *Proc. Int. Conf. World Wide Web*, 2017, pp. 173–182.

[18] S. Zhang, L. Yao, A. Sun, and Y. Tay, "Deep learning based recommender system: A survey and new perspectives," *ACM Comput. Surv.*, vol. 52, no. 1, pp. 1–38, 2019.

[19] S. Zhang et al., "Personalized latent structure learning for recommendation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 8, pp. 10285–10299, Aug. 2023.

[20] X. Wang, H. Chen, Y. Zhou, J. Ma, and W. Zhu, "Disentangled representation learning for recommendation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 1, pp. 408–424, Jan. 2023.

[21] X. He, K. Deng, X. Wang, Y. Li, Y. Zhang, and M. Wang, "LightGCN: Simplifying and powering graph convolution network for recommendation," in *Proc. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2020, pp. 639–648.

[22] X. Wang, X. He, M. Wang, F. Feng, and T.-S. Chua, "Neural graph collaborative filtering," in *Proc. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2019, pp. 165–174.

[23] X. Wang, X. He, Y. Cao, M. Liu, and T.-S. Chua, "KGAT: Knowledge graph attention network for recommendation," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2019, pp. 950–958.

[24] B. Kulis et al., "Metric learning: A survey," *Found. Trends Mach. Learn.*, vol. 5, no. 4, pp. 287–364, 2013.

[25] P. Zadeh, R. Hosseini, and S. Sra, "Geometric mean metric learning," in *Proc. Int. Conf. Mach. Learn.*, PMLR, 2016, pp. 2464–2471.

[26] X. Liu, B. V. K. Vijaya Kumar, J. You, and P. Jia, "Adaptive deep metric learning for identity-aware facial expression recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, 2017, pp. 20–29.

[27] W. Ge, "Deep metric learning with hierarchical triplet loss," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 269–285.

[28] J. Wang, F. Zhou, S. Wen, X. Liu, and Y. Lin, "Deep metric learning with angular loss," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 2593–2601.

[29] G. Ji, S. He, L. Xu, K. Liu, and J. Zhao, "Knowledge graph embedding via dynamic mapping matrix," in *Proc. Annu. Meeting Assoc. Comput. Linguistics, Int. Joint Conf. Natural Lang. Process.*, 2015, pp. 687–696.

[30] P. Li and A. Tuzhilin, "Dual metric learning for effective and efficient cross-domain recommendations," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 1, pp. 321–334, Jan. 2023.

[31] J. Park, Y.-C. Lee, and S.-W. Kim, "Effective and efficient negative sampling in metric learning based recommendation," *Inf. Sci.*, vol. 605, pp. 351–365, 2022.

[32] T. Wei, J. Ma, and T. W. Chow, "Collaborative residual metric learning," in *Proc. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2023, pp. 1107–1116.

[33] L. Qu, Y. Lin, W. Yuan, X. Wan, Y. Shi, and H. Yin, "Automated similarity metric generation for recommendation," 2024, *arXiv:2404.11818*.

[34] Y. Su, M. Bayoumi, and T. Joachims, "Optimizing rankings for recommendation in matching markets," in *Proc. Int. Conf. World Wide Web*, 2022, pp. 328–338.

[35] Z. Gong, Y. Song, T. Zhang, J.-R. Wen, D. Zhao, and R. Yan, "Your career path matters in person-job fit," in *Proc. AAAI Conf. Artif. Intell.*, 2024, pp. 8427–8435.

[36] L. Pizzato, T. Rej, T. Chung, I. Koprinska, and J. Kay, "RECON: A reciprocal recommender for online dating," in *Proc. ACM Conf. Recommender Syst.*, 2010, pp. 207–214.

[37] P. Xia, B. Liu, Y. Sun, and C. Chen, "Reciprocal recommendation system for online dating," in *Proc. Int. Conf. Adv. Social Netw. Anal. Mining*, 2015, pp. 234–241.

[38] Y. Mashayekhi, N. Li, B. Kang, J. Lijffijt, and T. De Bie, "A challenge-based survey of e-recruitment recommendation systems," *ACM Comput. Surveys*, vol. 56, 2022, Art. no. 252.

[39] C. Qin et al., "Enhancing person-job fit for talent recruitment: An ability-aware neural network approach," in *Proc. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2018, pp. 25–34.

[40] C. Zhu et al., "Person-job fit: Adapting the right talent for the right job with joint representation learning," *ACM Trans. Manage. Inf. Syst.*, vol. 9, no. 3, pp. 1–17, 2018.

[41] Y. Hou et al., "Leveraging search history for improving person-job fit," in *Proc. Int. Conf. Database Syst. Adv. Appl.*, Springer, 2022, pp. 38–54.

[42] Y. Du et al., "Enhancing job recommendation through LLM-based generative adversarial networks," in *Proc. AAAI Conf. Artif. Intell.*, 2024, pp. 8363–8371.

[43] Z. Zheng, Z. Qiu, X. Hu, L. Wu, H. Zhu, and H. Xiong, "Generative job recommendations with large language model," 2023, *arXiv:2307.02157*.

[44] K. Yao, J. Zhang, C. Qin, P. Wang, H. Zhu, and H. Xiong, "Knowledge enhanced person-job fit for talent recruitment," in *Proc. IEEE Int. Conf. Data Eng.*, 2022, pp. 3467–3480.

[45] L. Wu, Z. Qiu, Z. Zheng, H. Zhu, and E. Chen, "Exploring large language model for graph data understanding in online job recommendations," in *Proc. AAAI Conf. Artif. Intell.*, 2024, pp. 9178–9186.

[46] X. Hu, Y. Cheng, Z. Zheng, Y. Wang, X. Chi, and H. Zhu, "BOSS: A bilateral occupational-suitability-aware recommender system for online recruitment," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2023, pp. 4146–4155.

[47] B. Zheng, Y. Hou, W. X. Zhao, Y. Song, and H. Zhu, "Reciprocal sequential recommendation," in *Proc. ACM Conf. Recommender Syst.*, 2023, pp. 89–100.

[48] D. Shen, H. Zhu, C. Zhu, T. Xu, C. Ma, and H. Xiong, "A joint learning approach to intelligent job interview assessment," in *Proc. Int. Joint Conf. Artif. Intell.*, 2018, pp. 3542–3548.

[49] R. Yan, R. Le, Y. Song, T. Zhang, X. Zhang, and D. Zhao, "Interview choice reveals your preference on the market: To improve job-resume matching through profiling memories," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2019, pp. 914–922.

[50] Y. Luo, H. Zhang, Y. Wen, and X. Zhang, "ResumeGAN: An optimized deep representation learning framework for talent-job fit via adversarial learning," in *Proc. ACM Int. Conf. Inf. Knowl. Manage.*, 2019, pp. 1101–1110.

[51] S. Bian et al., "Learning to match jobs with resumes from sparse interaction data using multi-view co-teaching network," in *Proc. ACM Int. Conf. Inf. Knowl. Manage.*, 2020, pp. 65–74.

[52] S. Bian, W. X. Zhao, Y. Song, T. Zhang, and J.-R. Wen, "Domain adaptation for person-job fit with transferable deep global match network," in *Proc. Conf. Empir. Methods Natural Lang. Process., Int. Joint Conf. Natural Lang. Process.*, 2019, pp. 4810–4820.

[53] R. Le, W. Hu, Y. Song, T. Zhang, D. Zhao, and R. Yan, "Towards effective and interpretable person-job fitting," in *Proc. ACM Int. Conf. Inf. Knowl. Manage.*, 2019, pp. 1883–1892.

[54] B. Fu, H. Liu, Y. Zhu, Y. Song, T. Zhang, and Z. Wu, "Beyond matching: Modeling two-sided multi-behavioral sequences for dynamic person-job fit," in *Proc. Int. Conf. Database Syst. Adv. Appl.*, Springer, 2021, pp. 359–375.

[55] K. Q. Weinberger and L. K. Saul, "Distance metric learning for large margin nearest neighbor classification," *J. Mach. Learn. Res.*, vol. 10, no. 2, pp. 207–244, 2009.

[56] Z. Sun et al., "DaisyRec 2.0: Benchmarking recommendation for rigorous evaluation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 7, pp. 8206–8226, Jul. 2023.

[57] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme, "BPR: Bayesian personalized ranking from implicit feedback," in *Proc. Conf. Uncertainty Artif. Intell.*, 2009, pp. 452–461.

Yingpeng Du received the PhD degree in software engineering from Peking University, Beijing, China, in 2023. He is currently a research fellow with the College of Computing and Data Science, Nanyang Technological University. His research interests include recommender systems and ensemble learning. He has published more than 20 papers in top-tier journals and conferences, such as *Journal of Machine Learning Research*, *Pattern Recognition*, AAAI, SIGKDD, ICDM, IJCAI, and WWW.

Hongzhi Liu received the PhD degree in computer science from Peking University, Beijing, China, in 2012. He is currently a full professor with the School of Software and Microelectronics, Peking University. His research interests include recommender systems and ensemble learning. He has published more than 70 papers in refereed journals and conferences, including *Journal of Machine Learning Research*, *ACM Transactions on Architecture and Code Optimization*, *IEEE Transactions on Knowledge and Data Engineering*, SIGKDD, AAAI, IJCAI, and WWW.

Hengshu Zhu (Senior Member, IEEE) received the BE and PhD degrees both in computer science from the University of Science and Technology of China (USTC), China, in 2009 and 2014, respectively. He is currently is a Professor at the Computer Network Information Center (CNIC), Chinese Academy of Sciences (CAS), where he leads the Data Intelligence for Scientific Innovation Lab (DI4Science Lab). His general area of research is data mining and machine learning, with a focus on developing advanced data analysis techniques for innovative business applications. He has published prolifically in refereed journals and conference proceedings, such as *Nature Communications*, *IEEE Transactions on Knowledge and Data Engineering*, *IEEE Transactions on Mobile Computing*, *ACM Transactions on Information Systems*, SIGKDD, SIGIR, and NeurIPS. He served as the program co-chair of KDD CUP-2019 Regular ML Track, the industry chair of PRICAI-2022, the area chair of AAAI and IJCAI, and regularly as the (senior) program committee members in numerous top conferences. He was the recipient of the Distinguished Dissertation Award of CAS (2016), the Distinguished Dissertation Award of CAAI (2016), the Special Prize of President Scholarship for Postgraduate Students of CAS (2014), the Best Student Paper Award of KSEM-2011, WAIM-2013, CCDM-2014, and the Best Paper Nomination of ICDM-2014 and WSDM-2022. He is the senior member of ACM, CAAI, and CCF.

Yang Song received the PhD degree in computer science from Peking University, in 2013. He is currently a research fellow with BOSS Zhipin. His research interests include natural language processing and recommendation system. In recent years, he has published several papers in major conferences and journals, such as KDD, ACL, and EMNLP.

Zhi Zheng received the BE degree from the University of Science and Technology of China. He is currently working toward the PhD degree with the School of Data Science, University of Science and Technology of China (USTC). His research interests include data mining and recommendation system. He has published several papers in top venues, such as SIGKDD, WWW, *ACM Transactions on Information Systems*.

Zhonghai Wu received the BSc, MSc and PhD degrees in computer science from Zhejiang University, China, in 1990, 1993 and 1997 respectively. He is currently a full professor and the dean with the School of Software and Microelectronics, Peking University, Beijing, China. His research interests include machine learning, cloud computing, and context-aware services.