



Aligning Users Across Social Networks by Joint User and Label Consistence Representation

Xiang Li^{1,2,3}, Yijun Su^{2,3}, Neng Gao^{1,3}, Wei Tang^{1,2,3}, Ji Xiang³(✉),
and Yuewu Wang^{1,2,3}

¹ State Key Laboratory of Information Security,

Institute of Information Engineering, CAS, Beijing, China

² School of Cyber Security, University of Chinese Academy of Sciences,
Beijing, China

³ Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China
{lixiang9015,suyijun,gaoneng,tangwei,xiangji,wangyuewu}@iie.ac.cn

Abstract. Aligning users belonging to the same person in different social networks has attracted much attention. Recently, embedding methods have been proposed to represent users from different social networks into vector spaces with same dimension. To handle the challenge of *vector space diversity*, existing methods usually make vectors of known aligned users closer/consistent and overlap different vector spaces. However, compared to large amount of users in each social network, the consistence constraint on aligned users is not enough to reduce the diversity. Besides, missing edges/labels may provide incorrect information and affect the effect of the overlap between learned vector spaces. Therefore, we propose the OURLACER method, i.e, jOint UseR and LAbel Consistence Representation, to jointly represent each user and label under the consistence constraints of know aligned users and shared labels. Specifically, OURLACER utilizes collective matrix factorization to complete missing edges and labels for each user, which can provide sufficient information to distinguish each user. Moreover, OURLACER adds the consistence constraint on shared labels in different social networks. Because each user has own labels, label consistence can restrict each user and greatly reduce the diversity between learned vector spaces. Extensive experiments conducted on real-world datasets demonstrate that our method significantly outperforms other state-of-the-art methods.

1 Introduction

With the rapid development of Internet, social networks (e.g., Facebook, Twitter and YouTube) have played important roles in our daily life. Nowadays, people are accustomed to surfing on multiple social networks at the same time. According

to the statistical data from Pew Research Center report¹, more than half of the users tend to read news from multiple social media sites.

Nevertheless, existing social networks are provided by different companies and isolated from one another, which hinders the positive experience for users across different social networks. User Identity Linkage is to align users belonging to the same person in different social networks and has attracted much attention. Benefited from aligned users, we can complete and integrate users' information for sequent applications such as cross-network recommendation [5, 14, 22, 23], link prediction [1, 27, 28] and topic analysis [7].

Recently, several embedding methods have been proposed to solve the problem of user identity linkage [9, 11, 12, 18, 19, 30–32] by mapping users from each social network into a vector space with same dimension. Then, to give correct prediction, these methods overlap different vector spaces by making the vector representations of known aligned users closer or totally same (also called consistence constraint). Similarly, when we don't know any aligned users in advance, making the vector distributions similar can also make effect [8]. In conclusion, reducing **vector space diversity** can produce better vector spaces and overlap probable aligned users in different social networks.

However, existing embedding methods handling the challenge of vector space diversity still have following problems: (1) Missing edges and labels may mislead the process of learning good vector space for each social network, which makes space diversity hard to be reduced. (2) Consistence constraint on known aligned users may not come into effect. For example, two learned vector spaces may only overlap known aligned users while other users are all non-overlapped and the space diversity is still large.

In this paper, to address above problems in the challenge of vector space diversity, we propose the OURLACER method, i.e, **jOint UseR and LAbel ConsistencE Representation**, to jointly represent each user and label under the consistence constraints of know aligned users and shared labels. Specifically, OURLACER learns a good vector space for each social network with missing edges and labels completed by collective matrix factorization. Besides, to reduce the diversity between vector spaces, OURLACER not only utilizes the consistence constraint between known aligned users but also adds the consistence constraint between shared labels in different social networks. Because each user has unique labels, label consistence constraint can restrict each user and reduce the space diversity greatly.

The rest of this paper is organized as follows: We review related work in Sect. 2. Section 3 presents proposed OURLACER approach in detail and optimization algorithm is proposed in Sect. 4. Experimental evaluation and comparison are shown in Sect. 5. At last, Sect. 6 concludes the paper with a brief discussion.

¹ <http://www.pewresearch.org/fact-tank/2017/11/02/more-americans-are-turning-to-multiple-social-media-sites-for-news/>.

2 Related Work

In this section, we review the main lines of works on user identity linkage. Firstly, we briefly introduce traditional methods. Then, we discuss the progress of embedding methods.

Traditional methods have paid much attention to extract useful features and compute reasonable similarity. The first work on UIL problem utilizes usernames [24]. More specifically, they study the behavior patterns during selecting usernames and construct totally more than four hundreds features [25, 26]. Moreover, spatio-temporal information has been specially studied for extracting useful features [2, 3]. For content information, topic distribution has been demonstrated the effect [13]. Furthermore, based on pairwise similarity of artificial features, a new discrimination model has been proposed to promote the performance by viewing user identity linkage as a classification problem [10].

Considering the cost of artificial features, embedding methods have attracted much attention and made great progress. Different embedding methods have been proposed for different types of information. For network information, PALE preserves neighbor links in users' representations and learns the linear/non-linear mapping among known aligned users [11]. IONE models the followee/follower relationship and learn multiple representations for each user [9]. DeepLink introduces the deep neural network based on the learned users' representations by random walk [32]. Besides network information, label information has also been studied. MAH constructs hypergraph by labels for capturing high-order relation [19]. Based on MAH, UMAH emphasizes the effect of shared labels among different social networks and automatically learns the weights of different types of labels [31]. Besides, MASTER utilizes matrix factorization to factorize pre-computed similarity matrices into users' representations with kernel tricks [18]. MEgo2Vec views user identity linkage as a classification problem and capture user's attributes and ego network in the user's representation [30]. However, above embedding methods haven't thought over the effect of missing edges and labels when learning users' representations, which can be solved by proposed OURLACER method.

When learning the vector space for each social network, an inevitable challenge is how to reduce the diversity between vector spaces, which means probable aligned users are closer. Existing embedding methods mainly utilize the consistence constraint based on known aligned users [9, 11, 18, 19, 30–32]. Moreover, ULink modifies consistence constraint by making aligned users closer than non-aligned users [12]. However, compared the large amount of users in each social network, the consistence constraint based on limited known aligned users is not enough to reduce the space diversity. In this paper, the proposed OURLACER considers the emergence of same labels in different social networks and adds the consistence constraint based on these shared labels to restrict each user.

3 Proposed Method

In this section, we firstly introduce the basic notations. Then, we present the way of completing missing edges and labels. Finally, we show the two types of consistence constraint and give the final optimization objective.

We use $G_i = (A_i, L_i)$ to represent i -th social network. $A_i \in R^{n_i \times n_i}$ is the adjacency matrix, where 1 represents two users is connected. Different from previous methods, we use 0 to represent missing edge rather than no connection. Besides, $L_i \in R^{n_i \times d_i}$ refers to the label matrix, where each row means the labels of one user. Similarly, we use 0 to represent missing label. n_i means the total number of users in G_i . The final dimension of user representation is m .

3.1 Collective Matrix Factorization

In real life, users in social networks usually own a fraction of labels and links, which means some real labels/links are missing in the social networks. Hence, the vector space learned by existing embedding methods cannot capture full and useful information in fact. To learn a good vector space, we should take into account the missing labels and edges.

As demonstrated in the work of network embedding [16], some classical methods, such as DeepWalk [15], LINE [21], PTE [20] and node2vec [4], can be unified into the matrix factorization framework with closed forms. Therefore, we also apply matrix factorization to learn the final vector representations. Noting that we own not only the adjacency matrix but also the label matrix. Then, we factorize these two matrices jointly. For the i -th social network, we can express the problem as

$$\min_{U_i, V_i} \frac{1}{2} \|A_i - U_i U_i^T\|_F^2 + \frac{1}{2} \|L_i - U_i V_i^T\|_F^2 + \frac{\alpha}{2} (\|U_i\|_F^2 + \|V_i\|_F^2), \quad (1)$$

where $U_i \in R^{n_i \times m}$ represents users' vector representations and $V_i \in R^{d_i \times m}$ represents labels' vector representations. α is to control the complexity of U_i and V_i . $\|\cdot\|_F$ stands for Frobenius norm.

Though objective (1) can learn a vector space preserving enough network information and label information, we still cannot complete the missing edges and labels because objective (1) tend to recover original edges in A_i and original labels in L_i exactly. As a result, 0 in A_i is seen as no edge rather than missing edges. Therefore, analogous to transfer learning based collaborative filtering, we use collective matrix factorization [17] to complete missing edges and labels and learn a good vector space for social network G_i by following optimization problem

$$\min_{U_i, V_i} \frac{1}{2} \|I_i^A \odot (A_i - U_i U_i^T)\|_F^2 + \frac{1}{2} \|I_i^L \odot (L_i - U_i V_i^T)\|_F^2 + \frac{\alpha}{2} (\|U_i\|_F^2 + \|V_i\|_F^2), \quad (2)$$

where \odot is the Hadamard (element-wise) product and I_i^A is an indicator matrix. $I_i^A(p, q) = 1$ if $A_i(p, q)$ is observed, and otherwise $I_i^A(p, q) = 0$. Similarly,

$I_i^L(p, q) = 1$ if $L_i(p, q)$ is observed, and otherwise $I_i^L(p, q) = 0$. Noting that the normal value in A_i and L_i should be equal to 0 or 1. We change the value from discrete value into the continuous value in $[0, 1]$. Hence, we add new constraints on U_i and V_i and the optimization problem can be written as

$$\begin{aligned} \min_{U_i, V_i} & \frac{1}{2} \|I_i^A \odot (A_i - U_i U_i^T)\|_F^2 + \frac{1}{2} \|I_i^L \odot (L_i - U_i V_i^T)\|_F^2 \\ & + \frac{\alpha}{2} (\|U_i\|_F^2 + \|V_i\|_F^2) \\ \text{s.t.} & 0 \leq U_i \leq 1, 0 \leq V_i \leq 1. \end{aligned} \quad (3)$$

By above optimization problem, we can learn a good vector space for each social network with missing edges and labels completed.

3.2 Consistence Constraint

When learning the good vector space for each social network, we should make the diversity between different vector spaces as small as possible. In this paper, we apply the user consistence constraint widely used in traditional methods and propose a new label consistence constraint, which can restrict each user effectively.

User Consistence Constraint. In real life, we often know some aligned users in different social networks. A direct intuition is to make the representation of same user in different social network closer or totally same. Then, by preserving the network information, different vector spaces can be overlapped. Formally, we get following optimization problem

$$\min_{U_1, U_2} \|T_1 U_1 - T_2 U_2\|_F^2, \quad (4)$$

where $T_i \in R^{a \times n_i}$ is the indicator matrix. $T_i(p, q) = 1$ if the q -th user belongs to the p -th real person. a is the number of known aligned users and all know aligned users are re-numbered from 1 to a .

However, though preserving the network information, we only can restrict neighbors connected to known aligned users and users far away may suffer from the error propagation. Therefore, we should seek to other consistence constraint to bind each user.

Label Consistence Constraint. To restrict each user, we should add constraint on the information owned by each user. Naturally, each user owns unique labels and label consistence constraint is reasonable. Formally, the label consistence constraint can be formulated as

$$\min_{V_1, V_2} \|M_1 V_1 - M_2 V_2\|_F^2, \quad (5)$$

where $M_i \in R^{l \times d_i}$ is the indicator matrix. $M_i(p, q) = 1$ if the q -th label is the p -th shared label. l is the number of shared labels and all shared labels are re-numbered from 1 to l .

Finally, with above two types of consistence constraint, we can get the final optimization problem

$$\begin{aligned} \min_{U_i, V_i} \sum_i \frac{1}{2} \|I_i^A \odot (A_i - U_i U_i^T)\|_F^2 &+ \frac{1}{2} \|I_i^L \odot (L_i - U_i V_i^T)\|_F^2 \\ &+ \frac{\alpha}{2} (\|U_i\|_F^2 + \|V_i\|_F^2) + \frac{\beta}{2} (\|T_1 U_1 - T_2 U_2\|_F^2 + \|M_1 V_1 - M_2 V_2\|_F^2) \\ \text{s.t. } 0 \leq U_i \leq 1, 0 \leq V_i \leq 1, \end{aligned} \quad (6)$$

where β is a penalty term to control the importance of consistence constraints.

4 Optimization

In this section, we present the optimization algorithm to solve (6). It is hard to get the optimal solution due to the nonconvexity of optimization objective (6). Therefore, we utilize stochastic gradient method with multiplicative updating rules to ensure the nonnegativity of U_i and V_i . Besides, we use an alternative way to update U_1, U_2, V_1, V_2 . The whole algorithm is shown in Algorithm 1.

Optimize U_1, U_2 : The partial derivatives of objective (6) *w.r.t.* U_1, U_2 are

$$\begin{aligned} \frac{\partial L}{\partial U_1} &= I_1^A \odot (U_1 U_1^T - A_1) U_1 + I_1^L \odot (U_1 V_1^T - L_1) V_1 + \alpha U_1 \\ &\quad + \beta T_1^T (T_1 U_1 - T_2 U_2) \\ \frac{\partial L}{\partial U_2} &= I_2^A \odot (U_2 U_2^T - A_2) U_2 + I_2^L \odot (U_2 V_2^T - L_2) V_2 + \alpha U_2 \\ &\quad + \beta T_2^T (T_2 U_2 - T_1 U_1). \end{aligned} \quad (7)$$

Using the Karush-Kuhn-Tucker (KKT) complementarity conditions, we can obtain the following updating rules:

$$U_1 = U_1 \odot \sqrt{\frac{(I_1^A \odot A_1) U_1 + (I_1^L \odot L_1) V_1 + \beta T_1^T T_2 U_2}{(I_1^A \odot U_1 U_1^T) U_1 + (I_1^L \odot U_1 V_1^T) V_1 + \alpha U_1 + \beta T_1^T T_1 U_1}} \quad (8)$$

$$U_2 = U_2 \odot \sqrt{\frac{(I_2^A \odot A_2) U_2 + (I_2^L \odot L_2) V_2 + \beta T_2^T T_1 U_1}{(I_2^A \odot U_2 U_2^T) U_2 + (I_2^L \odot U_2 V_2^T) V_2 + \alpha U_2 + \beta T_2^T T_2 U_2}}. \quad (9)$$

Optimize V_1, V_2 : The partial derivatives of objective (6) *w.r.t.* V_1, V_2 are

$$\begin{aligned} \frac{\partial L}{\partial V_1} &= (U_1^T I_1^L \odot (U_1 V_1^T - L_1))^T + \alpha V_1 + \beta M_1^T (M_1 V_1 - M_2 V_2) \\ \frac{\partial L}{\partial V_2} &= (U_2^T I_2^L \odot (U_2 V_2^T - L_2))^T + \alpha V_2 + \beta M_2^T (M_2 V_2 - M_1 V_1). \end{aligned} \quad (10)$$

Similar to U_1, U_2 , we update V_1, V_2 by

$$V_1 = V_1 \odot \sqrt{\frac{((I_1^L)^T \odot L_1^T)U_1 + \beta M_1^T M_2 V_2}{((I_1^L)^T \odot V_1 U_1^T)U_1 + \alpha V_1 + \beta M_1^T M_1 V_1}} \quad (11)$$

$$V_2 = V_2 \odot \sqrt{\frac{((I_2^L)^T \odot L_2^T)U_2 + \beta M_2^T M_1 V_1}{((I_2^L)^T \odot V_2 U_2^T)U_2 + \alpha V_2 + \beta M_2^T M_2 V_2}}. \quad (12)$$

Algorithm 1. Joint User and Label Consistence Representation (OURLACER)

Input: $G_1 = (A_1, L_1), G_2 = (A_2, L_2)$, a known aligned users, l shared labels, parameters α, β , maximal number of iterations $maxiter$

Output: U_1, U_2 and V_1, V_2

- 1: Initialize U_1, U_2, V_1, V_2 with $(0, 1)$ uniform distribution
 - 2: **for** $t=1:maxiter$ **do**
 - 3: Update U_1 by (8)
 - 4: Update U_2 by (9)
 - 5: Update V_1 by (11)
 - 6: Update V_2 by (12)
 - 7: **if** objective (6) converge **then** break
-

Considering the value of U_i and V_i cannot exceed 1, we utilize the projection technique [6, 29] to project elements greater than 1 in U_i and V_i to 1 after each update process.

5 Experiment Study

In this section, we evaluate the performance compared to state-of-the-art methods. The main compared methods used in experiments include:

- Global Method (GM) [26]: By constructing spectral embedding for each user, this algorithm learns a linear transformation between known aligned users and this method can be seen as a basic version of PALE [11].
- MAH [19]: By constructing hypergraphs by labels and edges, each user owns a vector representation while known aligned users in different social networks own totally same vector representation.
- UMAH [31]: Based on MAH, this method considers the effect of shared labels and automatically learns the weights of different types of shared labels.
- OURLACER: Our proposed OURLACER method can learn the vector representation for each user and label with user and label consistence constraint.

Datasets. We use two real-world datasets to evaluate the performance: (1) Twitter *vs.* BlogCatalog: This dataset is provided by [31] and contains 2710 aligned users in both networks. For each user, this dataset has friendship, user-name and location information. For location information, 6.38% users do not reveal their location information in both networks and 31.03% only publish location information in one network. In the remaining users (62.59%), only 14.39% users input exactly the same location information in the two networks. (2) DBLP 2015 *vs.* 2016: We use “Yoshua Bengio” as the center node, and then crawl the co-authors that can be reached from the center node with no more than two hops. This process was repeated for authors published papers in 2015 and 2016 independently. Then, we can get two co-author networks. Besides, the conferences/journals published at least once in one year are used as the labels of users in that year. Finally, we have 2845 users in 2015, 3234 users in 2016 and 2169 aligned users between two networks. For label information, user in 2015 and 2016 respectively owns 882 and 1005 unique labels. Except unique labels, the number of shared labels is 945.

Performance Metric. To evaluate the performance of comparison methods, *Accuracy* and *Hit Precision@k* are used to evaluate the exact prediction and top- k prediction [31]. Specially, *Hit Precision@k* allocates different weights for different rank k :

$$h(x) = \frac{k - (\text{hit}(x) - 1)}{k},$$

where $\text{hit}(x)$ is the position of correct linked user in the returned top- k users. Then, *Hit Precision@k* can be computed on N test users by $\frac{\sum_i^N h(x_i)}{N}$. During experiments, we set $k = 5$.

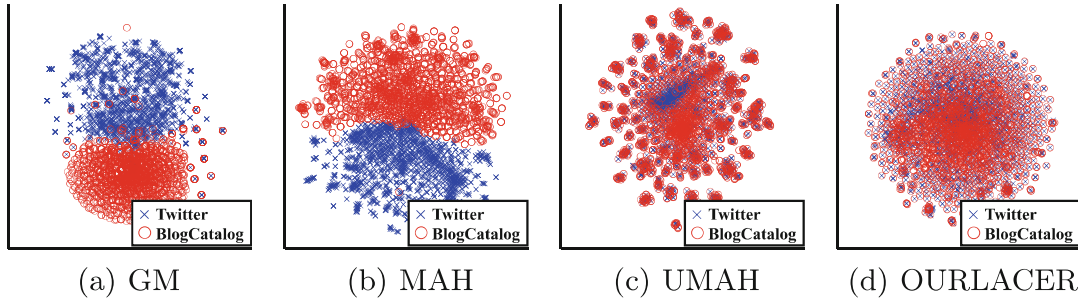
Experiment Setups. Compared methods except MAH have provided their source codes. For MAH, we implement it by matlab according to original paper and the implement of UMAH. We use same training ratio and test setting in UMAH. Noting that we only use the friendship among 2710 users in dataset Twitter-BlogCatalog. Considering existing study on the effect of dimension, we set the dimension of user representation to a big value such as 500. We denote r_o as the ratio of known aligned users among all aligned users. When setting the parameter of our method, we set β to a bigger value such as 10 to make loss of consistence constraint as small as possible. For parameter α , we set it to a same value for both two datasets. For parameters of other compared methods, we set them to reasonable values according to original papers. Because spectral embedding in GM can only capture structure information, we concat normalized label information with spectral embedding to form new user representation.

Overall Prediction Performance. We evaluate the overall prediction performance for compared methods. The ratio of known aligned users is 30%. As shown in Table 1, OURLACER always behaves better than other methods. Compared

Table 1. Overall prediction performance on two datasets with $r_o = 30\%$

Metric	Method	Twitter-BlogCatalog	DBLP 2015–2016
Accuracy	GM	3.54	2.93
	MAH	9.39	5.40
	UMAH	47.27	6.45
	OURLACER	52.63	11.59
Hit Precision@5	GM	6.72	5.40
	MAH	17.40	12.06
	UMAH	52.13	11.55
	OURLACER	55.88	19.61

to GM, other methods utilizing label information carefully show better performance, which demonstrate the potential good effect of labels. Furthermore, by comparing MAH and UMAH, we can find modeling the shared labels in two networks simultaneously is better than modeling labels separately. Finally, OURLACER is still much better than UMAH, which means the great effects of filling missing labels/edges and label consistence constraint.

**Fig. 1.** Visualization of user representations learned by different methods. We plot the overlap results on dataset Twitter-BlogCatalog with $r_o = 30\%$.

Visualization of User Representations. To vividly understand the effect of different methods, we visualize the learned user representations of dataset Twitter-BlogCatalog as shown in Fig. 1. Noting that we only use the representations of testing users. From Fig. 1(a) and (b), we can find the diversity between learned two vector spaces are very big. Besides, Fig. 1(c) and (d) demonstrates that the diversity between two vector spaces can be reduced greatly by carefully modeling label information. Furthermore, Fig. 1(c) shows UMAH tends to learn clustered representations, which means it can restrict users with coarse-grained. By contrary, Fig. 1(d) shows proposed OURLACER can learn the representations more uniformly, which means it can approximately restrict each user by proposed label consistence constraint. From Table 2, the gap between UMAH and

Table 2. Prediction performance by different consistence constraints

Metric	Constraint	Twitter-BlogCatalog	DBLP 2015–2016
Accuracy	user	11.51	8.66
	label	45.89	3.71
	user+label	52.63	11.59
Hit Precision@5	user	17.40	15.07
	label	47.93	6.73
	user+label	55.88	19.61

OURLACER on Hit Precision@5 is smaller than it on Accuracy, which demonstrates UMAH learns clustered vector space while OURLACER learns uniform vector space.

Effect of User and Label Consistence Constraint. Besides overall prediction performance, we also study the effect of user consistence constraint and label consistence constraint. The ratio of known aligned users is also set to 30% and the ratios of shared labels are 100% for Twitter-BlogCatalog and 33.37% for DBLP 2015–2016. As shown in Table 2, only one consistence constraint is always worse than two consistence constraints. Hence, our proposed label consistence constrain can make great effect. When the ratio of shared labels increases, the performance of only using label consistence constraint rises greatly. Specifically, the performance of only using label consistence constraint is much higher than only using user consistence constraint for Twitter-BlogCatalog while much smaller for DBLP 2015–2016, which means the effect of proposed label consistence constraint can be enhanced with the increasement of shared labels.

6 Conclusion

Vector space diversity is a great challenge for existing methods. Traditional methods try to learn a vector space for each social network while ignore the effect of missing edges/labels and label consistence among different social networks. Therefore, we propose the jOint UseR and LAbel ConsistencE Representation (OURLACER) method to learn a good space for each social network and greatly reduce the diversity between different vector spaces. Specially, OURLACER learns the vector space by using collective matrix factorization to complete the missing edges and labels. Besides, we propose the label consistence constraint to restrict each user and reduce the vector space diversity. Experiment results demonstrate the effectiveness of OURLACER. Future directions include the consideration of automatically learning the different importances of network information and label information.

Acknowledgments. This work is supported by the National Key Research and Development Program of China, and National Natural Science Foundation of China (No. U163620068).

References

1. Cao, X., Chen, H., Wang, X., Zhang, W., Yu, Y.: Neural link prediction over aligned networks. In: Proceedings of the 32th AAAI Conference on Artificial Intelligence, pp. 249–256 (2018)
2. Chen, W., Yin, H., Wang, W., Zhao, L., Hua, W., Zhou, X.: Exploiting spatio-temporal user behaviors for user linkage. In: Proceedings of the 26th ACM on Conference on Information and Knowledge Management, pp. 517–526 (2017)
3. Chen, W., Yin, H., Wang, W., Zhao, L., Zhou, X.: Effective and efficient user account linkage across location based social networks. In: Proceedings of the 34th IEEE International Conference on Data Engineering, pp. 1085–1096 (2018)
4. Grover, A., Leskovec, J.: node2vec: scalable feature learning for networks. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 855–864 (2016)
5. Hu, G., Zhang, Y., Yang, Q.: CoNet: collaborative cross networks for cross-domain recommendation. In: Proceedings of the 27th ACM International Conference on Information and Knowledge Management, pp. 667–676 (2018)
6. Koutra, D., Tong, H., Lubensky, D.: Big-Align: fast bipartite graph alignment. In: Proceedings of the 13th IEEE International Conference on Data Mining, pp. 389–398 (2013)
7. Lee, R.K.W., Hoang, T.A., Lim, E.P.: On analyzing user topic-specific platform preferences across multiple social media sites. In: Proceedings of the 26th International Conference on World Wide Web, pp. 1351–1359 (2017)
8. Li, C., et al.: Distribution distance minimization for unsupervised user identity linkage. In: Proceedings of the 27th ACM International Conference on Information and Knowledge Management, pp. 447–456 (2018)
9. Liu, L., Cheung, W.K., Li, X., Liao, L.: Aligning users across social networks using network embedding. In: Proceedings of the 25th International Joint Conference on Artificial Intelligence, pp. 1774–1780 (2016)
10. Liu, S., Wang, S., Zhu, F., Zhang, J., Krishnan, R.: HYDRA: large-scale social identity linkage via heterogeneous behavior modeling. In: Proceedings of the 37th ACM SIGMOD International Conference on Management of Data, pp. 51–62 (2014)
11. Man, T., Shen, H., Liu, S., Jin, X., Cheng, X.: Predict anchor links across social networks via an embedding approach. In: Proceedings of the 25th International Joint Conference on Artificial Intelligence, pp. 1823–1829 (2016)
12. Mu, X., Zhu, F., Lim, E.P., Xiao, J., Wang, J., Zhou, Z.H.: User identity linkage by latent user space modelling. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1775–1784 (2016)
13. Nie, Y., Jia, Y., Li, S., Zhu, X., Li, A., Zhou, B.: Identifying users across social networks based on dynamic core interests. *Neurocomputing* **210**, 107–115 (2016)
14. Perera, D., Zimmermann, R.: LSTM networks for online cross-network recommendations. In: Proceedings of the 27th International Joint Conference on Artificial Intelligence, pp. 3825–3833 (2018)

15. Perozzi, B., Al-Rfou, R., Skiena, S.: DeepWalk: online learning of social representations. In: Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 701–710 (2014)
16. Qiu, J., Dong, Y., Ma, H., Li, J., Wang, K., Tang, J.: Network embedding as matrix factorization: Unifying DeepWalk, LINE, PTE, and node2vec. In: Proceedings of the 11th ACM International Conference on Web Search and Data Mining, pp. 459–467 (2018)
17. Singh, A.P., Gordon, G.J.: Relational learning via collective matrix factorization. In: Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 650–658 (2008)
18. Su, S., Sun, L., Zhang, Z., Li, G., Qu, J.: MASTER: across multiple social networks, integrate attribute and structure embedding for reconciliation. In: Proceedings of the 27th International Joint Conference on Artificial Intelligence, pp. 3863–3869 (2018)
19. Tan, S., Guan, Z., Cai, D., Qin, X., Bu, J., Chen, C.: Mapping users across networks by manifold alignment on hypergraph. In: Proceedings of the 28th AAAI Conference on Artificial Intelligence (2014)
20. Tang, J., Qu, M., Mei, Q.: PTE: predictive text embedding through large-scale heterogeneous text networks. In: Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1165–1174 (2015)
21. Tang, J., Qu, M., Wang, M., Zhang, M., Yan, J., Mei, Q.: LINE: large-scale information network embedding. In: Proceedings of the 24th International Conference on World Wide Web, pp. 1067–1077 (2015)
22. Yan, M., Sang, J., Xu, C.: Mining cross-network association for youtube video promotion. In: Proceedings of the 22nd ACM International Conference on Multimedia, pp. 557–566 (2014)
23. Yan, M., Sang, J., Xu, C., Hossain, M.S.: A unified video recommendation by cross-network user modeling. *ACM Trans. Multimed. Comput. Commun. Appl.* **12**, 53:1–53:24 (2016)
24. Zafarani, R., Liu, H.: Connecting corresponding identities across communities. In: Proceedings of the 3rd International AAAI Conference on Weblogs and Social Media (2009)
25. Zafarani, R., Liu, H.: Connecting users across social media sites: a behavioral-modeling approach. In: Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 41–49 (2013)
26. Zafarani, R., Tang, L., Liu, H.: User identification across social media. *ACM Trans. Knowl. Discov. Data* **10**(2), 16:1–16:30 (2015)
27. Zhang, J., Chen, J., Zhi, S., Chang, Y., Yu, P.S., Han, J.: Link prediction across aligned networks with sparse and low rank matrix estimation. In: Proceedings of the 33rd IEEE International Conference on Data Engineering, pp. 971–982 (2017)
28. Zhang, J., Kong, X., Yu, P.S.: Predicting social links for new users across aligned heterogeneous social networks. In: Proceedings of the 13th IEEE International Conference on Data Mining, pp. 1289–1294 (2013)
29. Zhang, J., Yu, P.S.: Multiple anonymized social networks alignment. In: Proceedings of the 15th IEEE International Conference on Data Mining, pp. 599–608 (2015)
30. Zhang, J., et al.: MEgo2Vec: embedding matched ego networks for user alignment across social networks. In: Proceedings of the 27th ACM International Conference on Information and Knowledge Management, pp. 327–336 (2018)

31. Zhao, W., et al.: Learning to map social network users by unified manifold alignment on hypergraph. *IEEE Trans. Neural Netw. Learn. Syst.* **29**, 5834–5846 (2018)
32. Zhou, F., Liu, L., Zhang, K., Trajcevski, G., Wu, J., Zhong, T.: DeepLink: a deep learning approach for user identity linkage. In: *Proceedings of the 37th IEEE Conference on Computer Communications*, pp. 1313–1321 (2018)