Anchor User Oriented Accordant Embedding for User Identity Linkage

Xiang Li[1,2,3], Yijun Su[2,3], Neng Gao[1,3], Wei Tang[1,2,3], Ji Xiang[3(✉)], and Yuewu Wang[1,2,3]

[1] State Key Laboratory of Information Security, Institute of Information Engineering, CAS, Beijing, China
{lixiang9015,gaoneng,tangwei,wangyuewu}@iie.ac.cn

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

[3] Institute of Information Engineering, Chinese Academy of Sciences, Beijing, China
{suyijun,xiangji}@iie.ac.cn

Abstract. User Identity Linkage is to find users belonging to the same real person in different social networks. Besides, anchor users refer to matching users known in advance. However, how to match users only based on network information is still very difficult and existing embedding methods suffer from the challenge of **error propagation**. Error propagation means the error occurring in learning some users' embeddings may be propagated and amplified to other users along with edges in the network. In this paper, we propose the Anchor UseR ORiented Accordant Embedding (AURORAE) method to learn the vector representation for each user in each social network by capturing useful network information and avoiding error propagation. Specifically, AURORAE learns the potential relations between anchor users and all users, which means each user is directly connected to all anchor users and the error cannot be propagated without paths. Then, AURORAE captures the useful local structure information into final embeddings under the constraint of accordant vector representations between anchor users. Experimental results on real-world datasets demonstrate that our method significantly outperforms other state-of-the-art methods.

1 Introduction

Nowadays, the fast development of communication technology makes people easy to enjoy "second life" in Internet. To meet different needs, people have been used to appear on multiple social networks. User Identity Linkage (UIL) aims to find users belonging to the same person in different social networks [12]. UIL is a significant task due to its ability of completing user portrait and fusing diverse information based on matching users. Hence, much more useful information can be delivered to many sequent applications, such as cross-network recommendation [2,11,14,15], link prediction [1,17,18] and topic analysis [6].

However, how to solve the UIL problem purely based on network information is still very hard. Existing methods apply the network embedding technique

T. Gedeon et al. (Eds.): ICONIP 2019, CCIS 1143, pp. 561–572, 2019.
https://doi.org/10.1007/978-3-030-36802-9_60

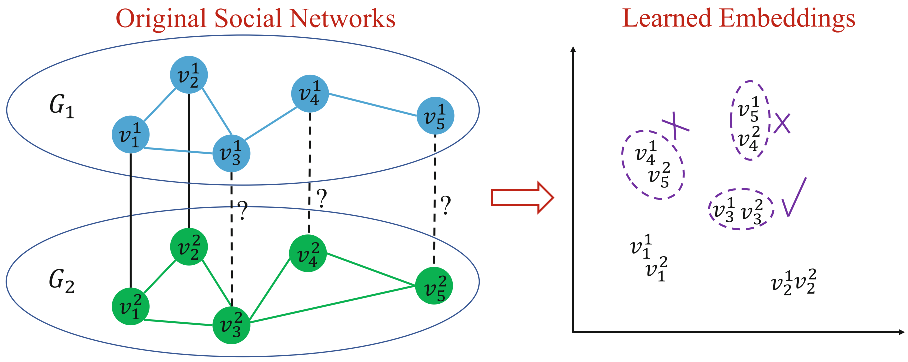


Fig. 1. A tiny example of error propagation. Given two social networks G_1 and G_2, the user space is formed by probable learned vector representations. v_i^j refers to i-th user in G_j and we already know matching user pairs (v_1^1, v_1^2) and (v_2^1, v_2^2). UIL aims to predict true matching user pairs (v_3^1, v_3^2), (v_4^1, v_4^2) and (v_5^1, v_5^2).

to learn the vector representation for each user with enough local structure information preserved [7–9,16,22]. These methods most pay much attention to the way of preseving enough useful information while ignore the main reason of misleading the learning process - **Error Propagation**. The challenge of error propagation means the error occurring in learning some users' embedding may be propagated and amplified along with the paths in networks. When the learned embeddings of two true matching users cannot correctly demonstrate they are a same real person, we think some errors happen during the process of learning embeddings. As shown in Fig. 1, from the user space, we cannot predict v_4^1 and v_4^2 are a same person. As a result, it's hard to correctly predict v_5^1 and v_5^2 are a same person due to error propagation. When user is far away from anchor users, the error would be amplified and we would not be able to predict correctly.

In this paper, to address the challenge of error propagation, we propose the **A**nchor **U**se**R** **OR**iented **A**ccordant **E**mbedding (AURORAE) method to learn the vector representation for each user in each social network by capturing useful network information and avoiding error propagation. Anchor user refers to user whose matching user has been given in advance. AURORAE learns potential graphs oriented at anchor users, which means each user is connected to all anchor users directly and not connected to users not being anchor users. By this way, due to full confidence of information from anchor users, error occurring during learning process has no direct path to be propagated and amplified. Then, when learning the user embeddings, AURORAE tends to capture useful local structure information and learn accordant embeddings to make users belonging to a same person much closer. Our contributions can be summarized as follows:

- We propose a new Anchor User Oriented Accordant Embedding method to solve the challenge of error propagation. To the best of our knowledge, we are the first to explicitly study the effect of error propagation and provide an effective solution purely based on network information.
- When handling error propagation, we give a direct but effective idea, which is to cut off the path propagating the error. Besides, a carefully devised

optimization algorithm is proposed to solve the final optimization problem formed by combining our idea with traditional embedding methods.
- Experiments on real-world Data sets demonstrate the effectiveness of our proposed method. Extension experiments show our method is superior to state-of-the-art methods under different conditions.

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

2 Related Work

In this section, we review two main kinds of methods only using network information. Firstly, we introduce the traditional propagation methods. Then, we discuss recent work of embedding methods.

Propagation methods discover unknown user pairs in an iterative way from anchor users [12]. Formally, they firstly find probable matching user pairs directly connected to anchor users. Then, they view these user pairs as anchor users and predict probable matching user pairs directly connected to them. Finally, when all users have been scanned, these methods will be terminated [3,23]. Another kind of propagation methods is dividing all users into several subsets and then matching user pairs among same subset [4]. The core of propagation methods is to design reasonable similarity measures centering around anchor users [4,10]. For example, the similarity of one user pair can be computed using the number of shared identified friends [4,23]. Other metrics such as common neighbors, Jaccard's coefficient and Adamic/Adar score are extended to measure the neighborhood similarities as well [3]. However, we can find the error propagation occurring in the prediction of each propagation can explicitly influence the next prediction, which can be avoided by our AURORAE.

Recently, embedding methods have attracted much attention due to its ability of preserving network structure information. Spectral embedding has been studied at early [16]. PALE preserves neighbor links in users' representations and learns the linear/non-linear mapping among anchor users [9]. IONE models the followee/follower relationship and learn multiple representations for each user [7]. ABNE utilizes attention mechanism to distinguish different effects of neighbors [8]. DeepLink introduces the deep neural network based on the learned users' representations by random walk [22]. Existing embedding methods focus on preserving useful information and cannot avoid error propagation during learning, which is the main problem solved by our proposed AURORAE.

3 Proposed Method

We use $G_i = (V_i, A_i)$ 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. n_i means the total

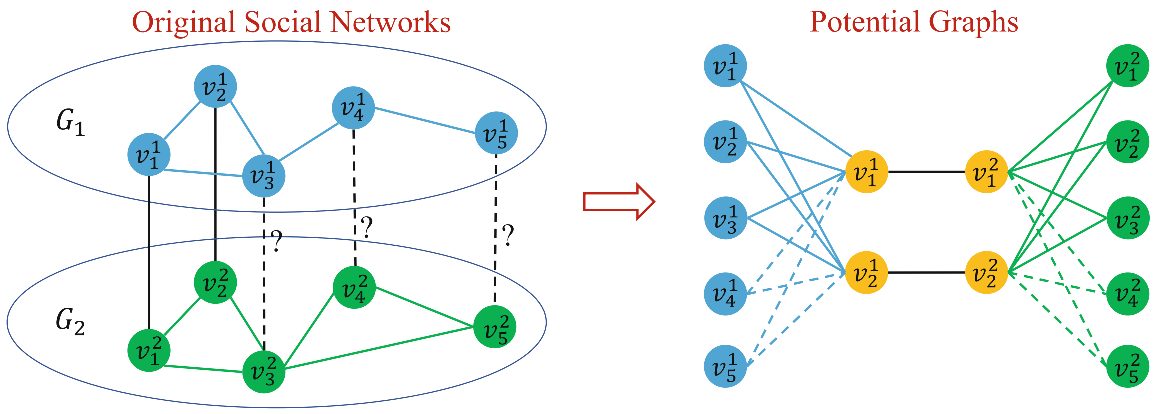


Fig. 2. Illustration of learning potential graphs oriented at anchor users. In the right, blue dash line and green dash line means edges to be learned. Blue solid line and green solid line means edges known. Besides, each edge owns a weight value to be learned, which is not shown in the figure. (Color figure online)

number of users in G_i. V_i refers to the set of all users and v_i^j means the i-th user in j-th social network. Besides, we use $H_i \in R^{n_i \times l}$ to represent the connection matrix between all users V_i and all anchor users. l is the number of all anchor users. m is the dimension of learned vector representation. In the sequel, we firstly present the way of learning potential graphs oriented at anchor users. Then, we discuss the way of learning accordant embeddings.

3.1 Learning Potential Graph Oriented at Anchor Users

In this section, we re-think the challenge of error propagation and solve it from an intuitive view. When learning the user representations, existing studies often think the information from anchor users are fully trustworthy. Besides, paths through questionable users is the essential reason of propagating and amplifying error. Hence, if questionable users cannot influence users connected to them, the error would have no path to be propagated and amplified. Naturally, as shown in Fig. 2, if all users are only connected to trustworthy anchor users and not directly connected to each other, the challenge of error propagation can be effectively avoided.

Formally, we can get the vector representations for the i-th social network by recovering original network:

$$\min_{U_i, Q_i} \frac{1}{2}||H_i - U_i Q_i^T||_F^2 + \frac{\beta}{2}||U_i||_F^2 \,, \tag{1}$$

where $U_i \in R^{n_i \times m}$ represents users' vector representations and $Q_i \in R^{l \times m}$ represents vector representations of all anchor users. β is to control the complexity of U_i and $||\cdot||_F$ stands for Frobenius norm.

Noting that objective (1) tries to recover connection matrix H_i exactly by user embeddings, which means only solid lines in Fig. 2 can be recovered. However, users connected to anchor users are so few that H_i is sparse and many users may be not connected to any anchor users such as users connected by dash lines in Fig. 2. In our experiment settings, we select 30% matching users as anchor users. Then, the ratio of nonzero values in H_i is $0.73\%, 0.34\%, 0.18\%$ and 0.24% for social network dataset Twitter, BlogCatalog, Douban Online and Douban Offline respectively. Hence, the final user embeddings may contain many zero values and the information learned in the vector representations is not sufficient.

To solve the problem of sparse connections, we need to complete unknown connections. As shown in Fig. 2, we want to learn the weight of each dash line. Formally, we view the zero in H_i (e.g., dash line in Fig. 2) as missing edge rather than no connection. Then, we use collective matrix factorization technique [13] to complete missing edges and restrict the value in $[0, 1]$. Finally, the optimization problem for the i-th social network can be written as:

$$\min_{0 \le U_i \le 1, 0 \le Q_i \le 1} \frac{1}{2} ||H_i \odot (H_i - U_i Q_i^T)||_F^2 + \frac{\beta}{2} ||U_i||_F^2 , \qquad (2)$$

where $\odot$ is the Hadamard (element-wise) product. To provide more information for learning the potential graph by (2), we need to preserve some local structure information:

$$\min_{0 \le U_i \le 1, 0 \le Q_i \le 1} \frac{1}{2} ||H_i \odot (H_i - U_i Q_i^T)||_F^2 + \frac{\alpha_i}{2} ||A_i - U_i U_i^T||_F^2 + \frac{\beta}{2} ||U_i||_F^2 , \qquad (3)$$

where α_i is to control how much information to be preserved. Objective (3) can make the reconstruction loss of observed edges minimized and complete missing edges with the help of local structure information.

3.2 Learning Accordant Embedding

Optimization problem (3) can learn potential graph oriented at anchor users for each social network. As shown in Fig. 2, we can view it as two potential bipartite graphs connected by anchor users. Hence, we cannot independently learn the user embeddings for different social networks. Specifically, we should learn the accordant embedding under the constraint of anchor users.

We have two kinds of constraints of anchor users from two different aspects. On one hand, when we consider the interaction between two different social networks, we should restrict the distance between embeddings of users belonging to a same person as closer as possible. On the other hand, among each bipartite graph, we can find the the vector representations U_i also contain the embeddings of anchor users. Then, we should make the embeddings of anchor users in U_i is similar to Q_i. Finally, the optimization problem can be rewritten as:

$$\min_{0 \le U_i \le 1, 0 \le Q_i \le 1} ||Q_1 - Q_2||_F^2 + ||T_1 U_1 - Q_1||_F^2 + ||T_2 U_2 - Q_2||_F^2 , \qquad (4)$$

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

In conclusion, the final optimization problem can be formulated as:

$$\min_{0\le U_i\le 1, 0\le Q_i\le 1} \sum_i \frac{1}{2}||H_i \odot (H_i - U_iQ_i^T)||_F^2 + \frac{\alpha_i}{2}||A_i - U_iU_i^T||_F^2 + \frac{\beta}{2}||U_i||_F^2 \\ \frac{\gamma}{2}(||Q_1 - Q_2||_F^2 + ||T_1U_1 - Q_1||_F^2 + ||T_2U_2 - Q_2||_F^2) \;, \tag{5}$$

where γ is penalty parameter and we set it to a large value such as 10 compared to the small value of other parameters.

4 Optimization

Due to the nonconvexity of (5), we cannot get the optimal solution. Therefore, we utilize an alternative way to update U_i, Q_i by stochastic gradient method with multiplicative updating rules, which can ensure the nonnegativity of U_i and Q_i. After each update step, we apply the projection technique [5,19] to project elements greater than 1 in U_i and Q_i to 1. The whole algorithm is shown in Algorithm 1.

Optimize U_1, U_2*:* The partial derivatives of objective (5) *w.r.t.*$\{U_i\}$ are

$$\frac{\partial L}{\partial U_1} = H_1 \odot (U_1Q_1^T - H_1)Q_1 + \alpha_1(U_1U_1^T - A_1)U_1 + \beta U_1 + \gamma T_1^T(T_1U_1 - Q_1) \\ \frac{\partial L}{\partial U_2} = H_2 \odot (U_2Q_2^T - H_2)Q_2 + \alpha_2(U_2U_2^T - A_2)U_2 + \beta U_2 + \gamma T_2^T(T_2U_2 - Q_2) \;. \tag{6}$$

Similar to classic nonnegative matrix factorization, we use following updating rules:

$$U_1 = U_1 \odot \sqrt{\frac{H_1Q_1 + \alpha_1A_1U_1 + \gamma T_1^TQ_1}{(H_1 \odot (U_1Q_1^T))Q_1 + \alpha_1U_1U_1^TU_1 + \beta U_1 + \gamma T_1^TT_1U_1}} \tag{7}$$

$$U_2 = U_2 \odot \sqrt{\frac{H_2Q_2 + \alpha_2A_2U_2 + \gamma T_2^TQ_2}{(H_2 \odot (U_2Q_2^T))Q_2 + \alpha_2U_2U_2^TU_2 + \beta U_2 + \gamma T_2^TT_2U_2}} \;. \tag{8}$$

We can find that it is easy to update U_1 and U_2 parallelly.

Optimize Q_1, Q_2*:* The partial derivates of objective (5) *w.r.t.*$\{Q_i\}$ are

$$\frac{\partial L}{\partial Q_1} = H_1^T \odot (Q_1U_1^T - H_1^T)U_1 + \gamma(Q_1 - Q_2) + \gamma(Q_1 - T_1U_1) \\ \frac{\partial L}{\partial Q_2} = H_2^T \odot (Q_2U_2^T - H_2^T)U_2 + \gamma(Q_2 - Q_1) + \gamma(Q_2 - T_2U_2) \;. \tag{9}$$

Similar to U_1, U_2, we update Q_1, Q_2 by

$$Q_1 = Q_1 \odot \sqrt{\frac{H_1^TU_1 + \gamma T_1U_1 + \gamma Q_2}{(H_1^T \odot (Q_1U_1^T))U_1 + 2\gamma Q_1}} \tag{10}$$

$$Q_2 = Q_2 \odot \sqrt{\frac{H_2^T U_2 + \gamma T_2 U_2 + \gamma Q_1}{(H_2^T \odot (Q_2 U_2^T)) U_2 + 2\gamma Q_2}}. \tag{11}$$

Algorithm 1. Anchor User Oriented Accordant Embedding (AURORAE)

Input: $G_1 = (V_1, A_1), G_2 = (V_2, A_2)$, anchor users, parameters α_i, β, γ, maximal number of iterations *maxiter*
Output: U_1, U_2 and Q_1, Q_2

1: Initialize U_1, U_2, Q_1, Q_2 with $(0, 1)$ uniform distribution
2: **for** t=1:*maxiter* **do**
3: Update U_1 by (7)
4: Update U_2 by (8)
5: Update Q_1 by (10)
6: Update Q_2 by (11)
7: **if** objective (5) converge **then** break

5 Experiment Study

In this section, we study the performance of our proposed AURORAE compared to several state-of-the-art methods. Besides, we evaluate the effect of different ratios of anchor users. Finally, experiment results are reported under different testing settings.

5.1 Experimental Settings

Compared Methods. To evaluate the performance of AURORAE, we select three state-of-the-art methods:

Table 1. Statistical information of datasets

Dataset	Type	#Users	#Edges	#Matching users
Douban online	Undirected	3,906	8,164	1,118
Douban offline	Undirected	1,118	1,511	
Twitter	Directed	5,120	164,919	1,609
Foursquare	Directed	5,313	76,972	

- IONE [7] models the followee and follower relationships and learns multiple representations for each user.
- ABNE [8] distinguishes the different effects of neighbors of each user and represents all users in a same vector space.
- DeepLink [22] gets the initial representations by random walk and feeds it to a deep neural network.
- AURORAE is proposed to solve the challenge of error propagation and learn accordant embeddings for users in each social network.

Datasets. We use following two pairs of datasets to evaluate different methods: (1) *Douban Online-Offline*, which is provided by [20]. Online users mean users who surf and interact with others in online Douban website and Offline users mean users who attend the offline activities. (2) *Twitter-Foursquare*, which is provided by [7,8]. Different from the first dataset, Twitter and Foursquare used in the experiments are directed networks, which means the adjacency matrix is not symmetric due to the emergence of followee/follower relationships. The statistical information of datasets is shown in Table 1.

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 [21]. Specially, *Hit Precision@k* allocates different weights for different rank k:

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

where $hit(x)$ 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 $\frac{\sum_i^N h(x_i)}{N}$.

Experiment Setups. All compared methods have provided their source codes. For reproducibility, we will also provide full codes publicly. The final dimension of user representation in compared methods is set according to original papers. For our proposed AURORAE method, we set $\beta = 0.4$ and $\gamma = 10$ for all datasets. For each α_i, we tune it in $\{0.1, 0.2, 0.3, 0.4, 0.5\}$. Due to the core idea of our method, we denote r_o as the ratio of anchor users among all matching users. When testing, we set $k = 5$ for metric *Hit Precision@k*. Besides, we test each pair of datasets by different orders and report the average performance.

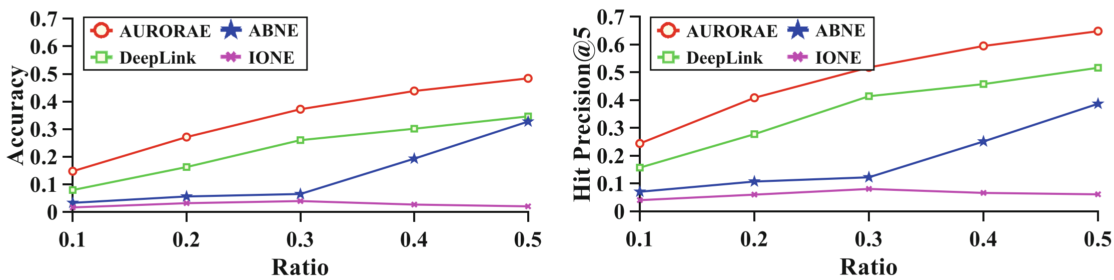
5.2 Experimental Results

Table 2. Overall prediction performance on different datasets with $r_o = 30\%$

Metric	Method	Online-Offline	Twitter-Foursquare
Accuracy	IONE	3.96	5.28
	ABNE	6.52	8.39
	DeepLink	26.02	1.02
	AURORAE	**37.21**	**10.66**
Hit Precision@5	IONE	8.07	10.37
	ABNE	12.26	14.86
	DeepLink	41.43	1.89
	AURORAE	**51.79**	**18.15**

Overall Prediction Performance. We evaluate the overall prediction performance for compared methods. The ratio of anchor users is 30%. As shown in Table 2, our proposed AURORAE method always behaves much better than other methods. IONE and ABNE shows much worse performance than DeepLink on Online-Offline. By contrary, DeepLink shows much worse than IONE and ABNE on Twitter-Foursquare. Hence, we can find IONE and ABNE are good at modeling directed networks while DeepLink is good at handling undirected networks. It is reasonable because IONE and ABNE are designed to handle the directed networks and DeepLink is not specially designed to handle the directed relationship. However, no matter whether networks are directed, our proposed AURORAE can achieve best performance. Compared to other methods not specially using graphs oriented at anchor users, the good performance of AURORAE demonstrates the effectiveness and generalization of learning accordant embedding based on the potential graphs oriented at anchor users.

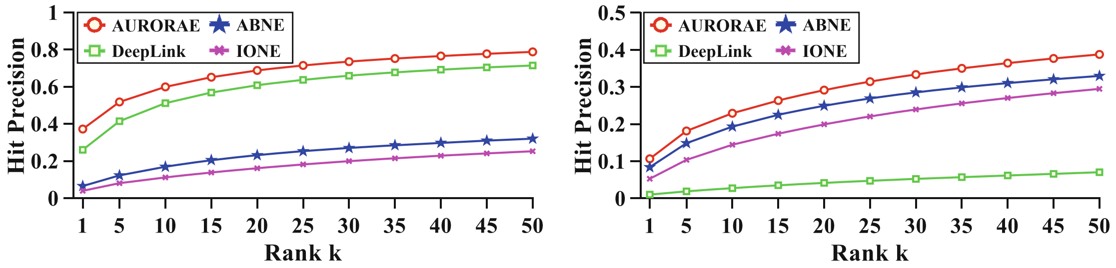
Effect of Anchor Users. Because the core idea of our method to avoid error propagation is based on anchor users, we should study the effect of different ratios of anchor users. Hence, we report the accuracy and hit precision@5 for compared methods on dataset Douban Online-Offline with the r_o varying in $\{0.1, 0.2, 0.3, 0.4, 0.5\}$. As shown in Fig. 3, with the increasement of r_o, the performances of all compared methods rise explicitly. When we only know a few anchor users, our proposed AURORAE still shows better performance than other methods. When we know more anchor users, the performance of AURORAE increases quickly. Though ABNE and DeepLink behaves better when the number of anchor users grows, our proposed AURORAE exceeds DeepLink about 9%–13% on Hit Precision@5 and 7%–14% on Accuracy when the ratio of anchor users varying. Therefore, using potential graphs oriented at anchor users is a good way to solve the challenge of error propagation even the number of anchor users is not large.



(a) Accuracy (b) Hit Precision@5

Fig. 3. Effect of different ratios of anchor users on Douban Online-Offline.

Effect of Rank k. When testing the performance, we focus on the top-k prediction performance. In practical scenario, algorithms can provide top-k list for experts to economize time and match users effectively. Hence, we evaluate the



(a) Online-Offline (b) Twitter-Foursquare

Fig. 4. Performance of different rank k on different datasets with $r_o = 30\%$.

performance for different rank k with $r_o = 30\%$. From Fig. 4, our AURORAE shows better performance than other methods. With the increasement of rank k, the performance of AURORAE can rise a lot on both two datasets, which means AURORAE can provide more accurate candidate users for sequent artificial identification.

6 Conclusion

Error Propagation is a great challenge for the task of User Identity Linkage. However, existing methods mainly seek to capture more useful information into user embeddings and lack the study of the essential reason for error propagation. In this paper, we study the reason leading to this challenge. Then, we propose the Anchor UseR ORiented Accordant Embedding (AURORAE) method to avoid error propagation by connecting each user to all anchor users and learn the accordant embedding with consistence constraint. Future work is to combine our method with more information such as label and content 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. 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)
3. Kong, X., Zhang, J., Yu, P.S.: Inferring anchor links across multiple heterogeneous social networks. In: Proceedings of the 22nd ACM International Conference on Information and Knowledge Management, pp. 179–188 (2013)

4. Korula, N., Lattanzi, S.: An efficient reconciliation algorithm for social networks. In: PVLDB, vol. 7, no. 5, pp. 377–388 (2014)
5. 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)
6. 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)
7. 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)
8. Liu, L., Zhang, Y., Fu, S., Zhong, F., Hu, J., Zhang, P.: ABNE: an attention-based network embedding for user alignment across social networks. IEEE Access **7**, 23595–23605 (2019)
9. 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)
10. Narayanan, A., Shmatikov, V.: De-anonymizing social networks. In: Proceedings of the 30th IEEE Symposium on Security and Privacy, pp. 173–187 (2009)
11. 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)
12. Shu, K., Wang, S., Tang, J., Zafarani, R., Liu, H.: User identity linkage across online social networks: a review. SIGKDD Explor. **18**(2), 5–17 (2016)
13. 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)
14. 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)
15. Yan, M., Sang, J., Xu, C., Hossain, M.S.: A unified video recommendation by cross-network user modeling. ACM Trans. Multimedia Comput. Commun. Appl. **12**, 53:1–53:24 (2016)
16. Zafarani, R., Tang, L., Liu, H.: User identification across social media. ACM Trans. Knowl. Discov. Data **10**(2), 16:1–16:30 (2015)
17. 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)
18. 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)
19. 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)
20. Zhang, S., Tong, H.: Final: fast attributed network alignment. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1345–1354 (2016)
21. 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)

22. 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)
23. Zhou, X., Liang, X., Zhang, H., Ma, Y.: Cross-platform identification of anonymous identical users in multiple social media networks. IEEE Trans. Knowl. Data Eng. **28**(2), 411–424 (2016)