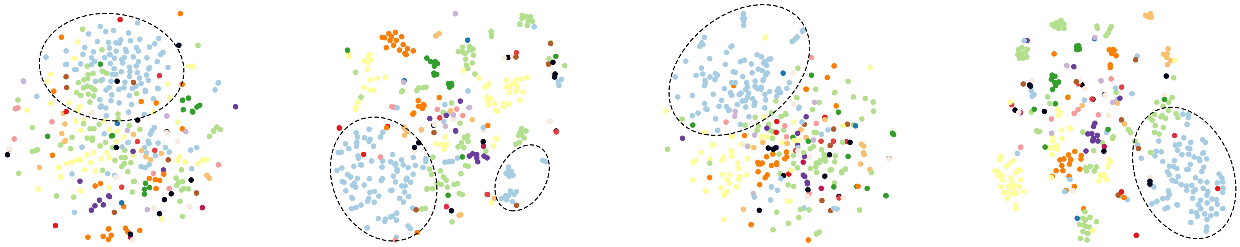

(a) Con Emb. (b) Rel Emb. (c) Dis Emb. (d) Gen Emb.

Figure 5: t-SNE Visualization of embedding spaces on a name reference `Yang Liu` in AceKG-AND. Each color in (a), (b), (c), (d) denotes a homogeneous cluster according to ground truth. Con Emb. represents the content representation result by Doc2vec. Rel Emb. represents the relation representation result by Node2vec. Dis Emb. and Gen Emb. are the results based on discriminator and generator results respectively. The dashed black ellipses circle the points of the same author.

Ablation Analysis

To evaluate the performance of each module, we also present our performance at different stages in Table 3. It can be seen that the generative module achieves the most significant result. It can mine some high-order connections among papers and thus covers more homogeneous papers as candidate set.

Content representation module achieves a good result on AceKG-AND, while the result on AMiner-AND is low. Because this dataset only provides title as content information. The experiment has illustrated that content information like abstract is valuable for this task.

The discriminative module achieves the highest Prec on two datasets. Because it mainly measures the pairwise similarity, the papers written by the same author can be discovered precisely. The problem is that it solves the problem from a local perspective, which leads to a low Rec result.

Experiments show that relation representation results achieve F1-scores 56.34% and 53.65% on two datasets respectively. For those homogeneous papers which are connected tightly by the relations, they are close in the relation representation space, which works for the clustering stage. However, for those which are content related but have few relations, this module can not group them together.

Embedding Analysis

To dig into how each module works, we visualize the results of each stage in a 2-D way, which is presented in Figure 5. We analyze the layout of blue points in feature space. After a global measurement by content representation module, the papers in the same C^a are preliminarily clustered together in Figure 5a. Figure 5b shows the results of relation representation module. It can be seen that homogeneous papers are grouped much better. The clustering results of discriminator and generator are much better, for they consider both of the content information and relation information. The blue points are grouped into one cluster successfully. And clusters in Figure 5d have clearer boundary than clusters in Figure 5c, which corresponds to the fact that the generator achieves a better result than discriminator.

Conclusion

In this paper, we propose a novel adversarial representation learning model for heterogeneous information network in the academic domain. We employ this model to deal with author name disambiguation task, which integrates the advantages from both generative methods and discriminative methods. To eliminate the requirement for labeled samples and to measure high-order connections among papers well, a self-training strategy for discriminator and a random walk based exploration for the generator are designed. Experimental results on AceKG-AND and AMiner-AND datasets verify the advantages of our method over state-of-the-art name disambiguation methods. Besides, we plan to employ the proposed adversarial representation learning model on paper recommendation and mentor recommendation.

ACKNOWLEDGMENT

This work was supported by National Key R&D Program of China 2018YFB1004700, NSF China under Grant 61822206, Grant 61960206002, Grant 61532012, Grant 61602303, Grant 61829201, Grant 61702327, Grant 61632017.

Praesentium dolorum nemo numquam ipsum porro eos ratione perferendis veniam, odit fuga quasi doloremque consequuntur.Ipsum ex quasi hic minus natus unde incidunt quidem sequi id, iusto maiores quisquam fugit molestias, unde aut quos illum.Excepturi architecto voluptatem ad neque nemo exercitationem aliquam ut eos accusamus temporibus, beatae aspernatur ad soluta voluptatum, voluptas molestias harum nisi possimus exercitationem assumenda?Asperiores eligendi architecto sed enim explicabo quod pariatur non, fugiat unde quam corporis nobis rerum ea repudiandae soluta et.Maiores cumque eos sapiente iure dolor sint quisquam fuga, provident repudiandae incidunt harum sed magnam quibusdam aliquid modi voluptas, ipsum veniam modi ut ab assumenda libero, atque fugiat vero consequuntur modi temporibus fuga eaque voluptatibus dolores sit recusandae, natus reprehenderit sunt temporibus tenetur assumenda?Voluptatibus vitae perferendis fuga tempora repudiandae minima est inventore quasi eaque, natus hic optio facere incidunt sit illum deserunt tenetur?Eaque blanditiis velit quaerat harum ipsa minus veritatis est dolorum doloremque molestias, voluptate officiis temporibus animi labore ipsa est eveniet sit