Generative Adversarial Network Based Heterogeneous Bibliographic Network Representation for Personalized Citation Recommendation

Xiaoyan Cai, Junwei Han,* Libin Yang
School of Automation, Northwestern Polytechnical University, China
{xiaoyanc, jhan, libiny}@nwpu.edu.cn

Abstract
Network representation has been recently exploited for many applications, such as citation recommendation, multi-label classification and link prediction. It learns low-dimensional vector representation for each vertex in networks. Existing network representation methods only focus on incomplete aspects of vertex information (i.e., vertex content, network structure or partial integration), moreover they are commonly designed for homogeneous information networks where all the vertices of a network are of the same type. In this paper, we propose a deep network representation model that integrates network structure and the vertex content information into a unified framework by exploiting generative adversarial network, and represents different types of vertices in the heterogeneous network in a continuous and common vector space. Based on the proposed model, we can obtain heterogeneous bibliographic network representation for efficient citation recommendation. The proposed model also makes personalized citation recommendation possible, which is a new issue that a few papers addressed in the past. When evaluated on the AAN and DBLP datasets, the performance of the proposed heterogeneous bibliographic network based citation recommendation approach is comparable with that of the other network representation based citation recommendation approaches. The results also demonstrate that the personalized citation recommendation approach is more effective than the non-personalized citation recommendation approach.

Introduction
With the rapid growth in the number of scientific papers, researchers might find it hard to find appropriate and necessary work to cite. Researchers usually retrieve papers from web search engines based on certain keywords, manually review them and decide which paper should be cited. However, it is labor-intensive as well as time-consuming, and especially difficult for the beginning researchers. Citation recommendation which can recommend a list of reference papers that are relevant to the researchers' information need, is an essential technology to overcome this problem. A variety of citation recommendation approaches have been proposed in the literature (He et al. 2010; He et al. 2011; Ren et al. 2014; Huang et al. 2014). These approaches are either global or local. Global recommendation (He et al. 2011; Meng et al. 2013; Ren et al. 2014) recommends a list of references for a given manuscript. Local recommendation (He et al. 2010), on the other hand, aims to recommend citations for specific context of each place where a citation should be made. We focus on global citation recommendation in this work.

Existing global citation recommendation approaches fall into three categories: collaborative filtering (CF) (Hernando et al. 2016), content-based filtering (CBF) (Nascimento et al. 2011) and graph-based approaches (Meng et al. 2013). CF makes citation recommendation by finding correlations among other researchers with similar research interests. CBF recommends a reference paper based on words and/or topic features of a manuscript and the identity of a researcher. Graph-based approaches often consider citation recommendation as a link prediction problem and solve the problem using properties of random walks.

Recently, network representation has been exploited for graph-based citation recommendation (Gupta and Varma 2017). Most network representation methods are based on network structure. Inspired by deep learning techniques in natural language processing (Mikolov et al. 2013), Perozzi et al. (2014) proposed DeepWalk approach, which learns feature vectors for vertices from a corpus of random walks generated from networks by employing neural network models. Grover and Leskovec (2016) proposed Node2Vec approach, which maximizes the likelihood of preserving network neighborhoods of vertices to learn feature representation. However, they only consider network structure, ignoring content information associated with each vertex. Pan et al. (2016) proposed TriDNR approach, which uses information from three parties, i.e. vertex structure, vertex
content and vertex labels, to jointly learn vertex representation. But this approach ignores inter-relationship among heterogeneous vertices.

In this paper, we propose a generative adversarial network based model to learn heterogeneous bibliographic network representation by modeling both vertex content and network structure. The distributed representation obtained using the model in turn can be used to calculate similarity scores. Finally the top ranked scientific papers are generated as the citation recommendation list. The contributions of this paper are lists as follows:

1. A bibliographic network is constructed to model different relationships among heterogeneous objects (i.e., scientific papers, authors, manuscript and author of the manuscript);
2. A generative adversarial network based heterogeneous bibliographic network representation (GAN-HBNR) model is developed, which incorporates bibliographic network structure and content of different kinds of objects to learn optimal representations of these objects;
3. A novel personalized citation recommendation approach based on the GAN-HBNR model is proposed, and the thorough experimental studies are conducted to verify the effectiveness of the proposed approach.

Related Work

Graph-based Citation Recommendation

Recent studies employed graph-based approaches to investigate the citation recommendation problem (Strohman et al. 2007; Zhou et al. 2008; Meng et al. 2013; Pan et al. 2016; Gupta and Varma 2017). Strohman et al. (2007) deemed citation recommendation as link prediction problem. They represented each paper as a vertex, the citation relationship as the link between vertices and a new paper as a vertex without any in-link and out-link. Zhou et al. (2008) measured paper similarities by combining the author-paper graph, the paper-venue graph and the paper citation graph. Then they recommended reference papers by treating some known citations as positive labels and applying semi-supervised learning on the combined graphs. Gori and Pucci (2006) proposed to recommend research papers by a random-walk based approach. Meng et al. (2013) presented a personalized citation recommendation approach, which incorporated different kinds of information, such as content of papers, authorship and citation etc., into a unified graph model. Pan et al. (2016) proposed an academic paper recommendation approach based on a heterogeneous graph containing various kinds of features. Gupta and Varma (2017) proposed a scientific paper recommendation combining distributed representations of paper’s content and distributed representations of the graph constructed from the bibliographic network. We propose a heterogeneous bibliographic network representation based citation recommendation approach, which recommends the top ranked scientific papers to the author of the manuscript based on the similarity scores among the representation of different objects (manuscript, scientific papers, the author of manuscript and authors of the scientific papers).

Network Representation

Network representation was first proposed by Hoff et al. (2002), and it was later followed by many approaches include multi-dimensional scaling (MDS) (Kruskal and Wish 1978), Laplacian Eigenmap (Belkin and Niyogi 2001), local linear embedding (LLE) (Roweis and Saul 2000) and IsoMap (Tenenbaum et al. 2000), which treats eigenvectors as representations. However, these approaches are not applicable to large scale networks due to computational complexity. Perozzi et al. (2014) proposed DeepWalk by using local information from truncated random walks as input and learning latent representation of vertices in a network. Tang et al. (2015a) proposed a network embedding method called LINE to preserve the local and global network structures. Grover and Leskovec (2016) proposed Node2Vec approach to learn vertex representation by maximizing the likelihood of preserving network neighborhoods of vertices. But DeepWalk, LINE, Node2Vec only focus on homogeneous networks. PTE (Tang et al. 2015b) extends the LINE to handle with heterogeneous network embedding problem, but PTE only leverages network structure, ignoring vertex content information. Pan et al. (2016) proposed TriDNR, a tri-party deep network representation model, to simultaneously learn network structure and vertex content. But this model cannot capture highly non-linear structures. Moreover, the structure information they use is not comprehensive (e.g., ignoring relationships among heterogeneous objects).

Generative Adversarial Network based Heterogeneous Bibliographic Network Representation (GAN-HBNR) Model

Heterogeneous Bibliographic Network Construction

In this section, we first construct a heterogeneous bibliographic network containing papers and authors as $G=<V, E, C>$, where $V = V_p \cup V_a$ is the vertex set, $V_p = \{p_i\}$ ($1 \leq i \leq n$, $n$ is the total number of papers), $V_a = \{a_j\}$ ($1 \leq j \leq m$, $m$ is the total number of authors), $E = E_{pp}, E_{ad}, E_{pt}$ is the edge set, $E_{pp} = \{e_{qi}, p_i, p_j \in V_p\}$, $E_{ad} = \{e_{qi}, a_i, a_j \in V_a\}$ and
Let \( E_{pi} = \{ e_{pi}, p_i \in V_p, a_j \in V_a \} \) correspond to the edges between papers, the edges between authors and the edges between papers and authors, respectively. If \( p_i \) cites \( p_j \), or \( p_i \) is cited by \( p_j \), or \( a_i \) collaborates with \( a_j \), or \( a_i \) is one of the authors of \( p_j \), or \( a_j \) is one of the authors of \( p_i \), then \( e_{ij} = 1 \); otherwise \( e_{ij} = 0 \). \( C = C_p \cup C_a \) is the set of content information, let \( c_{pi} \) denote the text content associated with the paper \( p_i \) and \( C_p = \{ c_{pi} \} (1 \leq i \leq n) \), \( c_{aj} \) denote the text content associated with the papers which are written by author \( a_j \) and \( C_a = \{ c_{aj} \} (1 \leq j \leq m) \). Let \( B \) denote the adjacency matrix for the bibliographic network, and let \( b_k = \{ e_{1,k}, \ldots, e_{n+m,k} \} \) be an adjacency vector.

**Model Description**

The architecture of the proposed GAN-HBNR model is shown in Figure 1. The whole architecture consists of two main modules: the content2vec module and the generative adversarial bibliographic network module. Based on the model, we can learn an effective feature representation vector that preserves both vertex content information and network structure information and thus can be applied to personalized citation recommendation task.

- **Content2vec Module**
  We collect all the text content associated with one paper vertex, we also collect all the text content associated with papers which are written by one author. Thus each vertex in the bibliographic network contains text information. We employ doc2vec approach (Le and Mikolov 2014) as our content2vec module. The content representation of each vertex can be obtained from this module. Therefore, we can maximize the objective as follows:
  \[
  O = \sum_{i=1}^{n+m} \log P(w_{-b} : w_b | v_i) \tag{1}
  \]
  where \( w \) is a word in the text information of \( v_i \), \( b \) is the window size of the word sequence. After optimizing this objective, we can obtain the content representation \( u_i \) of each vertex \( v_i \).

- **Generative Adversarial Bibliographic Network Module**
  This module is the core part of our model, it integrates content information and structure information of the heterogeneous bibliographic network. As \( e_i \) describes link relationships among vertex \( v_i \) and other vertices in the network, the adjacency matrix \( E \) reflects the structure of the network. We extract each adjacency vector \( e_i \) and concatenate it with the corresponding content vector \( u_i \) of \( v_i \) in the GAN-HBNR model. Therefore, the content information and structure information can be learned simultaneously.

  We then employ energy-based generative adversarial network (Zhao et al. 2016) to train the generative bibliographic network module. One difference to Zhao et al. (2016) is that we use a Denoising Autoencoder (DAE) as our energy function, because the DAE has been found to produce superior representations to the standard Autoencoder (Vincent et al. 2010). We define a feed-forward generator network \( G(z) \) that takes a vector \( z \in \mathbb{R}^{h_g} \) as input and produces a generated vector, with \( h_g \) being the number of dimensions in the input noise vector (sampled from \( N(0, I) \)). We also define a discriminator network \( D(x) \), seen as an energy function, that takes vectors \( x \in \mathbb{R}^{n+m} \) and produces an energy estimate \( E \in \mathbb{R} \).
During the encoding phrase of DAE, we use single layer to map the input data to a highly nonlinear latent space, so the encoding process is
\[
y = f(W^c x + b_c)
\]
(2)
where \(W^c\) is a set of learned parameters, \(b_c\) is a learned bias term, \(x^c\) is a corrupted version of \(x\), \(f\) is a nonlinear function, and \(y \in \mathbb{R}^h\) is the hidden representation of \(x\) with size \(h_j\).

The decoding phase is a reflection of the encoder, its output \(q_j\) should be close to the input \(x_j\). The decoding process is
\[
q = g(W^d y + b_d)
\]
(3)
where \(W^d\) and \(b_d\) are another learned set of weights and bias terms, \(g\) is a nonlinear function. The final energy value is the mean squared reconstruction error:
\[
E = \frac{1}{n + m} \sum_{i=1}^{n+m} (x_i - y_i)^2
\]
(4)

The energy function is trained to push down on the energy of real vectors \(x\), and to push up on the energy of generated vectors \(G(z)\) (Zhao et al. 2016). Given a positive margin \(m\), a real vector \(x\) and a generated vector \(G(z)\), the discriminator loss \(L_D\) and the generator loss \(L_G\) are formally defined by:
\[
\begin{align*}
L_D(x,z) &= D(x) + [m - D(G(z))]^+ \\
L_G(z) &= D(G(z))
\end{align*}
\]
(5)
(6)
where \([]\)\(^+ = \max(0,)\). Minimizing \(L_G\) with respect to the parameters of \(G\) is similar to maximizing the second term of \(L_D\). It has the same minimum but non-zero gradients when \(D(G(z)) \geq m\). We use the vector \(y\) as the new network representation of the heterogeneous bibliographic network.

One of the advantages of our GAN-HBNR model is that when new vertices enter the network, we do not need to retrain the adversarial bibliographic network module. When we obtain as new adjacency vector \(e_j\), we can feed it into the model and obtain the representation at a complexity of \(O(1)\). If there exist no link between the new vertex and the network, we can exploit its content information.

GAN-HBNR Model based Personalized Citation Recommendation Approach

Given a manuscript \(q\), we propose a heterogeneous bibliographic network representation based personalized citation recommendation approach, which aims to return top ranked scientific papers as reference papers by measuring the similarity scores between the manuscript and all the scientific papers in the dataset. In our work, we formulate the manuscript \(q\) as manuscript author \(q_a\) and manuscript text \(q_t\), i.e., \(q = [q_a, q_t]\). We deem \(q_t\) as a testing paper, \(q_a\) as an author of the manuscript and all the scientific papers in the dataset \(P\) as training papers. Algorithm 1 below summarizes the whole process of the heterogeneous bibliographic network representation based personalized citation recommendation approach.

Algorithm 1 GAN-HBNR Model based Personalized Citation Recommendation Algorithm

**Input:** The heterogeneous bibliographic network \(G = (V, E, C)\) consists of the manuscript text \(q_t\), the manuscript author \(q_a\) (if available), the training papers and all the authors of the training papers, Adjacency matrix \(B\), window size \(w\), the dimension \(h_j\), Number \(Q\).

**Output:** Citation Recommendation list

1: Train a paragraph vector model based on \(C\), obtain the content representations \(U\)
2: X=Merge\([B, U]\)
3: Generate negative samples using \(G(z)\)
4: Repeat:
5: for \(d\)-steps do
6: Update the discriminator by ascending its stochastic gradient:
\[
\nabla_{\theta_d} \frac{1}{m+n} \sum_{i=1}^{m+n} \left[\log(D(\tilde{x}^{(i)})) + \log(1 - D(G(z^{(i)})))\right]
\]
7: end for
8: for \(g\)-steps do
9: Update the generator by descending its stochastic gradient:
\[
\nabla_{\theta_g} \frac{1}{m+n} \sum_{i=1}^{m+n} \log(1 - D(G(z^{(i)})))
\]
10: end for
11: Until converge
12: Obtain the network representations \(Y = \{y_i\}\)
13: Calculate the similarity score \(r_q\) for the manuscript and rank the training papers according to \(r_q\).
14: Select top ranked \(Q\) training papers as citation recommendation list.

The similarity scores is represented as \(r_q = [r_{q_{pp1}}, r_{q_{pp2}}, \cdots, r_{q_{pp}}, r_{q_{qa}}]\), \(p_i \in P\) (i = 1,2,···,n). The input to the recommendation system is the word sequence of training papers and testing papers, all the authors of the training papers and author of the testing papers, as well as the adjacency matrix based on the network. All these papers and authors are mapped into vectors based on our proposed GAN-HBNR model. Thus the similarity scores can be calculated as \(r_q = V_{PR}V_{q_t} + V_{AR}V_{q_a}^T\), where \(V_{PR} = [v_{p_1}; v_{p_2}; \cdots; v_{p_q}]\) is the vector representation of training papers, \(v_{q_t}\) is the vector representation of the manuscript text, \(V_{AR} = [v_{a_1}; v_{a_2}; \cdots; v_{a_q}]\) is the vector representation of authors related to training papers, \(v_{q_a}\) is the vector representation of the manuscript author. Training papers are ranked according to the similarity scores, the
top ranked ones are selected as the final citation recommendation list.

**Experiments**

**Datasets**

In order to evaluate the quality of the proposed model, we conduct experiments on two bibliographic datasets: (1) AAN (ACL Anthology Network) dataset\(^1\), which is established by Radev and Muthukrishnan (2009), it consists of conference papers and journal papers in the field of computational linguistics. We remove the papers which have missed titles or abstracts in the dataset, then we use the remaining 12,555 papers published from 1965 to 2013 as experimental dataset. For evaluation purpose, we divide the entire dataset into two disjoint sets, the papers published before 2013 are deemed as training set (11,197 papers) and the remaining papers fall into the testing set (1,358 papers). (2) DBLP dataset\(^2\), which consists of bibliography data in computer science (Tang et al. 2008). Instead of using full dataset, we choose a subset since some samples miss complete references. We also remove the papers which have missed titles or abstracts in the dataset, then we select the papers published before 2013 as training set (33,016 papers) and papers published from 2013 to 2015 as testing set (3172 papers).

In our work, we extract the title and abstract of the papers in the two datasets as document content, and we define a manuscript as the title, abstract and an author of the manuscript.

**Evaluation Methods**

Our ultimate aim is to recommend more relevant reference papers to the given manuscript. We use three common metrics as follows:

**Recall@N**: It is defined as the percentage of the original reference papers that appear in the top N recommended papers. Here we use N=\(\{20,40,60,80,100\}\) to evaluate the proposed approach.

**Mean Average Precision (MAP)**: As Recall@N only considers the top N ranking results, ignoring the exact ranking position. MAP is a precision metric that emphasizes ranking relevant papers higher, which can overcome the above disadvantage. Let \(T_p\) be the set of the testing papers. For a paper \(p_j\) in \(T_p\), the correct reference paper set of \(p_j\) is \(R_j\), and our proposed approach returns a reference paper list \(R_G\). We consider the top 40 recommended papers in the ranking list, so \(|R_G|=40|\). The MAP is defined as:

\[
MAP = \frac{1}{|T_p|} \sum_{p_i \in T_p} \frac{1}{|R_C|} \sum_{r_j \in R_C, \text{rank}(r_j) \leq 20} q(r_j) + \frac{1}{q(r_j)}
\]

where \(r_j \in R_c\) is a correct reference paper, \(\text{rank}(r_j)\) is defined as the position of \(r_j\) in \(R_G\) if \(r_j\) is in \(R_G\), otherwise \(\text{rank}(r_j)\) is defined to be zero. \(q(r_j)\) is set to the number of the correct reference papers which ranks higher than \(r_j\).

**Mean Reciprocal Rank (MRR)**: It measures how far from the top appears the first relevant reference papers. MRR is defined as:

\[
MRR = \frac{1}{|T_p|} \sum_{p_i \in T_p} \left( \frac{1}{\text{rank}(p_{\text{first}})} \right)
\]

where \(\text{rank}(p_{\text{first}})\) is the position of the first relevant reference papers in the reference list \(R_G\).

**Performance on GAN-HBNR Model based Personalized Citation Recommendation**

In this set of experiments, we first focus on the query text only, ignoring the query author information, i.e., \(q_1=\{q_i\}\). \(q_i\) also denotes the non-personalized manuscript. We set the representation size \(h_g\), which is the size of the DAE hidden state, to 100. The generator input noise vector \(h_g\) is set to be the same size. The generator is a three-layer feedforward neural network, the first two layers use ReLU activation function and the output layer uses a sigmoid nonlinear function. The first two layers are both of size 400, with the final output layer being the same size as the vocabulary. Meanwhile, the first two layers use batch normalization (Ioffe and Szegedy 2015). The discriminator encoder consists of a single layer followed by a leaky ReLU nonlinear function (with a leak of 0.03). The decoder is a linear transformation back to the vocabulary size. We optimize both \(G\) and \(D\) using Adam (Kingma and Ba 2015) with an initial learning rate of 0.0001. Our DAE corruption process is to randomly set 40% of the input values as zero, and we use a margin size \(m\) of 5% of the vocabulary size. We follow the same validation procedure as Glover (2016), with a learning rate of 0.01 and using the tanh activation function.

We are interested in studying whether personalized citation recommendation can provide more appropriate and individualized recommendation results to the users than non-personalized citation recommendation. We denote a personalized manuscript by \(q_2=\{q_i,q_j\}\). When a user who inputs the manuscript for the personalized citation recommendation has not yet published any papers, the proposed approach will be reduced to non-personalized citation rec-

\(^1\)http://clair.eecs.umich.edu/aan/index.php
\(^2\)http://arxiv.org/citation
ommendation for the user, because the query information contains $q_i$ only. Table 1 reports the experimental results.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Approach</th>
<th>MAP</th>
<th>MRR</th>
<th>Recall@20</th>
<th>Recall@40</th>
<th>Recall@60</th>
<th>Recall@80</th>
<th>Recall@100</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAN</td>
<td>GAN-HBNR, $q_2$</td>
<td>0.293</td>
<td>0.312</td>
<td>0.586</td>
<td>0.678</td>
<td>0.737</td>
<td>0.749</td>
<td>0.787</td>
</tr>
<tr>
<td></td>
<td>GAN-HBNR, $q_1$</td>
<td>0.280</td>
<td>0.299</td>
<td>0.563</td>
<td>0.665</td>
<td>0.712</td>
<td>0.726</td>
<td>0.769</td>
</tr>
<tr>
<td></td>
<td>GAN-HBNR, $q_2$</td>
<td>0.289</td>
<td>0.309</td>
<td>0.557</td>
<td>0.649</td>
<td>0.695</td>
<td>0.711</td>
<td>0.758</td>
</tr>
<tr>
<td></td>
<td>GAN-HBNR, $q_1$</td>
<td>0.273</td>
<td>0.296</td>
<td>0.541</td>
<td>0.632</td>
<td>0.673</td>
<td>0.695</td>
<td>0.734</td>
</tr>
<tr>
<td>DBLP</td>
<td>GAN-HBNR, $q_2$</td>
<td>0.280</td>
<td>0.299</td>
<td>0.563</td>
<td>0.665</td>
<td>0.712</td>
<td>0.726</td>
<td>0.769</td>
</tr>
<tr>
<td></td>
<td>GAN-HBNR, $q_1$</td>
<td>0.273</td>
<td>0.296</td>
<td>0.541</td>
<td>0.632</td>
<td>0.673</td>
<td>0.695</td>
<td>0.734</td>
</tr>
</tbody>
</table>
As shown in Table 4, the results returned by the GAN-HBNR with $q_2$ approach have four records that match the ground truth citation list of the manuscript, whereas the results returned by the GAN-HBNR with $q_1$ and TriDNR approaches have three and two matching records, respectively. This observation demonstrates that the GAN-HBNR with $q_2$ approach obtained a better result in this case study since the manuscript author, manuscript text, as well as text information and author information of training papers are utilized in this approach. There are two same citations in the top-5 results returned by the GAN-HBNR with $q_2$ approach and GAN-HBNR with $q_1$ approach, but the number of the corrected papers in the top-5 results returned by the GAN-HBNR with $q_2$ approach is more than that returned by the GAN-HBNR with $q_1$ approach. We attribute it to the GAN-HBNR with $q_2$ approach incorporates manuscript author information, while the GAN-HBNR with $q_1$ does not do it. The top 5 recommended results returned by TriDNR only contain two corrected papers, this is due to TriDNR only consider paper information, ignoring author information and author-paper information in bibliographic network representation.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Approach</th>
<th>MAP</th>
<th>MRR</th>
<th>Recall@20</th>
<th>Recall@40</th>
<th>Recall@60</th>
<th>Recall@80</th>
<th>Recall@100</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAN</td>
<td>GAN-HBNR, $q_1$</td>
<td>0.280</td>
<td>0.299</td>
<td>0.563</td>
<td>0.665</td>
<td>0.712</td>
<td>0.726</td>
<td>0.769</td>
</tr>
<tr>
<td></td>
<td>TriDNR</td>
<td>0.259</td>
<td>0.273</td>
<td>0.551</td>
<td>0.648</td>
<td>0.689</td>
<td>0.702</td>
<td>0.753</td>
</tr>
<tr>
<td></td>
<td>Doc2Vec</td>
<td>0.246</td>
<td>0.265</td>
<td>0.540</td>
<td>0.631</td>
<td>0.675</td>
<td>0.689</td>
<td>0.741</td>
</tr>
<tr>
<td></td>
<td>Node2Vec</td>
<td>0.237</td>
<td>0.251</td>
<td>0.533</td>
<td>0.623</td>
<td>0.662</td>
<td>0.673</td>
<td>0.729</td>
</tr>
<tr>
<td></td>
<td>LINE</td>
<td>0.225</td>
<td>0.240</td>
<td>0.521</td>
<td>0.617</td>
<td>0.651</td>
<td>0.662</td>
<td>0.713</td>
</tr>
<tr>
<td></td>
<td>DeepWalk</td>
<td>0.214</td>
<td>0.228</td>
<td>0.509</td>
<td>0.603</td>
<td>0.640</td>
<td>0.651</td>
<td>0.700</td>
</tr>
<tr>
<td>DBLP</td>
<td>GAN-HBNR, $q_1$</td>
<td>0.273</td>
<td>0.296</td>
<td>0.541</td>
<td>0.632</td>
<td>0.673</td>
<td>0.695</td>
<td>0.734</td>
</tr>
<tr>
<td></td>
<td>TriDNR</td>
<td>0.253</td>
<td>0.271</td>
<td>0.518</td>
<td>0.621</td>
<td>0.650</td>
<td>0.671</td>
<td>0.715</td>
</tr>
<tr>
<td></td>
<td>Doc2Vec</td>
<td>0.245</td>
<td>0.263</td>
<td>0.509</td>
<td>0.613</td>
<td>0.641</td>
<td>0.659</td>
<td>0.703</td>
</tr>
<tr>
<td></td>
<td>Node2Vec</td>
<td>0.234</td>
<td>0.250</td>
<td>0.497</td>
<td>0.601</td>
<td>0.632</td>
<td>0.647</td>
<td>0.690</td>
</tr>
<tr>
<td></td>
<td>LINE</td>
<td>0.223</td>
<td>0.239</td>
<td>0.483</td>
<td>0.587</td>
<td>0.621</td>
<td>0.635</td>
<td>0.679</td>
</tr>
<tr>
<td></td>
<td>DeepWalk</td>
<td>0.214</td>
<td>0.226</td>
<td>0.471</td>
<td>0.573</td>
<td>0.610</td>
<td>0.622</td>
<td>0.667</td>
</tr>
</tbody>
</table>

Table 3. Illustration of Top-5 Reference papers generated by the Three Network Representation based Citation Recommendation Approaches on DBLP dataset

<table>
<thead>
<tr>
<th>Title of the Manuscript</th>
<th>Approaches</th>
<th>Top-5 System Generated Reference Papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequential Summarization: A Full View of Twitter Trending Topics</td>
<td>GAN-HBNR with $q_2$</td>
<td>1) TweetMotif: Exploratory search and topic summarization for Twitter (✓) 2) Experiments in microblog summarization (✓) 3) Event summarization using tweets (✓) 4) Summarizing sporting events using Twitter (✓) 5) Twitter topic summarization by ranking tweets using social influence and content quality</td>
</tr>
<tr>
<td></td>
<td>GAN-HBNR with $q_1$</td>
<td>1) Event summarization using tweets (✓) 2) Towards real-time summarization of scheduled events from twitter streams 3) TweetMotif: Exploratory search and topic summarization for Twitter (✓) 4) Summarizing sporting events using Twitter (✓) 5) Multi-Tweet Summarization for Flu Outbreak Detection</td>
</tr>
<tr>
<td></td>
<td>TriDNR</td>
<td>1) TweetMotif: Exploratory search and topic summarization for Twitter (✓) 2) Event summarization using tweets (✓) 3) Towards real-time summarization of scheduled events from twitter streams 4) Twitter topic summarization by ranking tweets using social influence and content quality 5) Joint topic modeling for event summarization across news and social media streams</td>
</tr>
</tbody>
</table>

**Conclusion**

In this paper, we propose a generative adversarial network based heterogeneous bibliographic network representation model, which integrates network structure and the vertex content information into a unified framework and represents different types of vertices in the heterogeneous network in a continuous and common vector space. A personalized citation recommendation approach is developed based on the obtained network representation. We evaluate our proposed approach on the AAN and DBLP datasets, the results reveals that the combination of network-based structure and content-based analysis can improve bibliographic network representation for personalized citation recommendation. In the future, we plan to explore venue information in the process of bibliographic network representation.

**Acknowledgements**

The work described in this paper was supported by Nationa
References


