Yong-ping Du, Chang-qing Yao, Nan Li, 2015. Using heterogeneous patent network features to rank and discover influential inventors. *Frontiers of Information Technology & Electronic Engineering*, **16**(7):568-578. [doi:10.1631/FITEE.1400394]

Using heterogeneous patent network features to rank and discover influential inventors

Key words: Heterogeneous patent network, Influence, Rule-based ranking

Contact: Yong-ping Du
E-mail: ypdu@bjut.edu.cn

Introduction

- Patent data contains rich information and denotes the innovative technology which is being protected. It also represents the competitive picture of the enterprise.
- The topic model is a type of statistical model for discovering the topics that occur in a collection of documents.
- We propose a new rule-based iteration approach to rank the inventors in a patent heterogeneous network.
- The influence of the inventor is measured and we obtain a ranked inventor list. In addition, we use the topic distribution to discover the interest of the top ranked inventors by the LDA model.

Framework of our method (I)

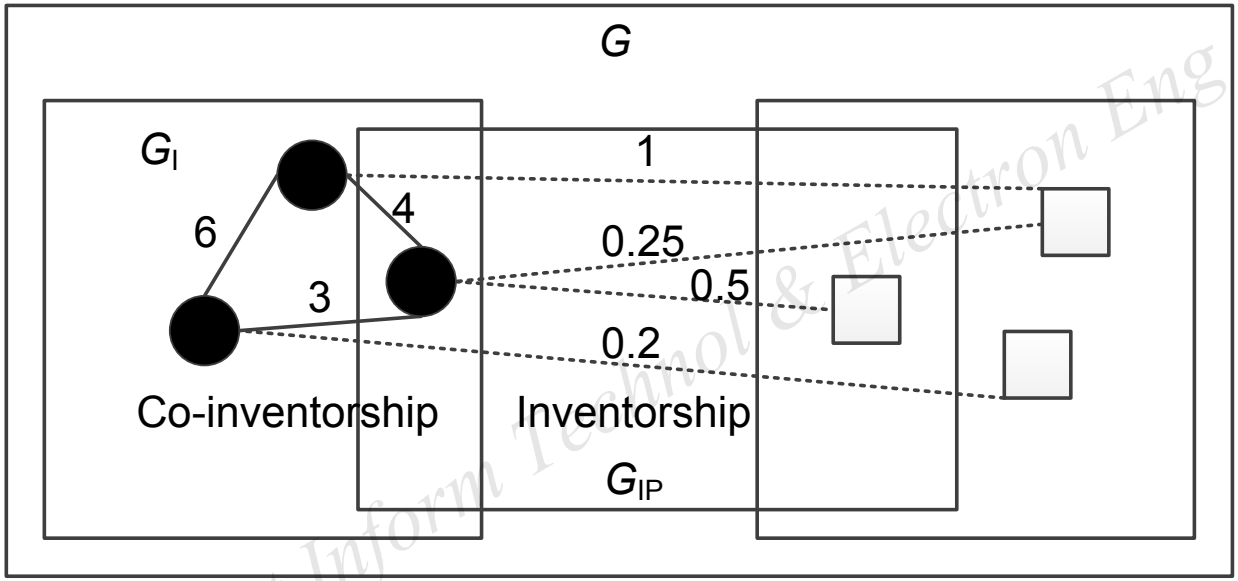

Fig. 1 Data model of the inventor-patent heterogeneous network

The inventor is represented by a black circle and the patent by a square. The network G is constituted by two subnetworks: co-inventor network G_I and inventor network G_{IP}

Framework of our method (II)

- **Rule 1** Highly ranked inventors tend to co-invent with other highly ranked inventors. We obtain the rank value of inventor *k* by

$$\mathrm{RankInventor}_i(k) = \alpha\left[\sum_{r=1}^{n} M_{\mathrm{II}}(k,r)\mathrm{RankInventor}(r)\cdot w_{\mathrm{II}} + \mathrm{RankInventor}_{i-1}(k)\cdot(1-w_{\mathrm{II}})\right]. \quad (2)$$

- **Rule 2** Highly ranked inventors generally invent highly ranked patents. We obtain the rank value of patent *j* by

$$\mathrm{RankPatent}(j) = \frac{\beta\sum_{k=1}^{n} M_{\mathrm{PI}}(j,k)\mathrm{RankInventor}(k)}{\mathrm{RankPatent}_{\max}(M)}, \quad (3)$$

- **Rule 3** Highly ranked patents are invented by highly ranked inventors. We update the rank value of inventor *r* by

$$\mathrm{RankInventor}_i(r) = \gamma\left[\sum_{j=1}^{m} M_{\mathrm{IP}}(r,j)\mathrm{RankPatent}(j)\cdot w_{\mathrm{IP}} + \mathrm{RankInventor}_{i-1}(r)\cdot(1-w_{\mathrm{IP}})\right], \quad (4)$$

Framework of our method (III)

- **Inventor model based on LDA**

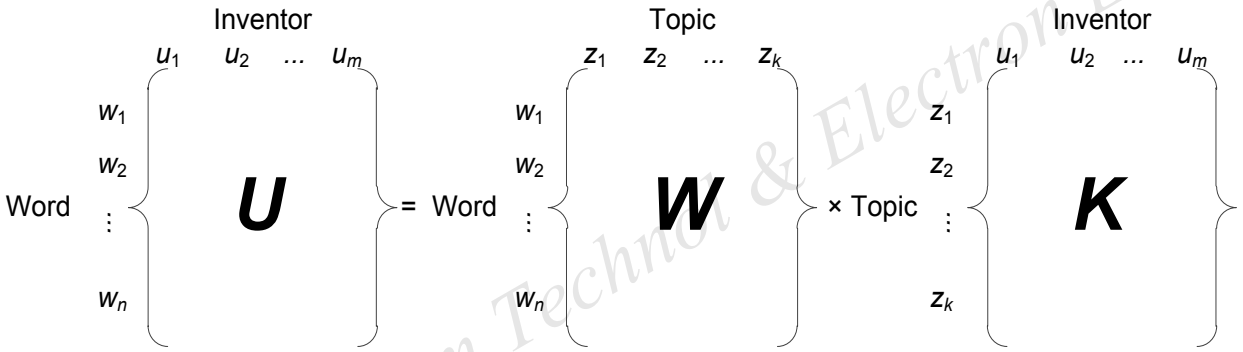

Fig. 6 Updated three-layer inventor-topic-word model of latent Dirichlet allocation (LDA)

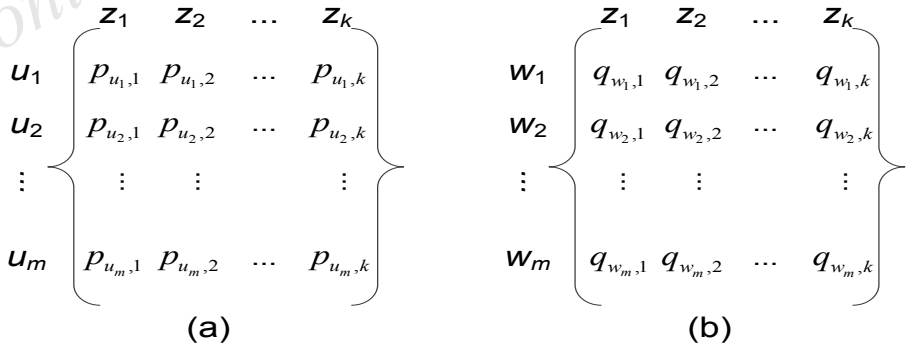

Fig. 7 Matrix representation in the LDA model: (a) inventor-topic; (b) word-topic

Major results

Table 2 The top 10 inventors ranked based on our inventor-ranking method and the PageRank algorithm

Rank	Inventor	
	Inventor-ranking	PageRank
1	Wei Zhu	Wei Zhu
2	Lijuan Wang	Lijun Yan
3	Junbao He	Shuren Guo
4	Lijun Yan	Yong Wang
5	Rongling Wu	Binwen Liang
6	Binwen Liang	Chao Wang
7	Junyong An	Marcus Dirk
8	Chao Wang	Wei Li
9	Zhenwen Duan	Lijuan Wang
10	Shuren Guo	Gang Zhao

Table 4 Evaluation results by different parameter values

N	S			SP		
	$\alpha=\beta=\gamma=1.0$	$\alpha=0.8$, $\beta=\gamma=0.1$	$\alpha=\beta=0.1$, $\gamma=0.8$	$\alpha=\beta=\gamma=1.0$	$\alpha=0.8$, $\beta=\gamma=0.1$	$\alpha=\beta=0.1$, $\gamma=0.8$
10	0.42	**0.68**	0.38	0.37	**0.64**	0.36
20	0.45	**0.74**	0.44	0.41	**0.71**	0.40
50	0.54	**0.81**	0.53	0.50	**0.77**	0.50
80	0.67	**0.78**	0.62	0.60	**0.72**	0.57
100	0.65	**0.79**	0.63	0.59	**0.75**	0.60

Bold values are the highest for each N

Major results (Con'd)

Table 5 Topic distribution probability of the top 10 inventors

Inventor	Probability			
	Topic 1	Topic 2	Topic 3	Topic 4
Wei Zhu	0.316	0.207	0.239	0.237
Lijuan Wang	0.277	0.213	0.248	0.261
Junbao He	0.248	0.258	0.253	0.241
Lijun Yan	0.278	0.23	0.22	0.271
Rongling Wu	0.245	0.259	0.264	0.231
Binwen Liang	0.252	0.305	0.231	0.212
Junyong An	0.175	0.296	0.256	0.272
Chao Wang	0.280	0.230	0.256	0.233
Zhenwen Duan	0.273	0.303	0.179	0.244
Shuren Guo	0.321	0.229	0.217	0.232

Conclusions

- We propose a new inventor-ranking method based on the heterogeneous patent network. The rank value of the inventor is calculated iteratively by the rules.
- The topic distributions of the inventor are also analyzed by the LDA model, and they can denote the interest of the inventor, especially for the top ranked inventors.
- Experiments on the real data set show that our method achieves good performance and the inventor-ranking method converges faster than the PageRank algorithm.