Research Paper

JIS

Journal of Information Science
1–16

DOI: 10.1177/0165551518806116
journals.sagepub.com/home/jis
SAGE

Author–Subject–Topic model for reviewer recommendation

Jian Jin
Department of Information Management, School of Government, Beijing Normal University, China

Qian Geng
Department of Information Management, School of Government, Beijing Normal University, China

Haikun Mou
Alibaba Group, China

Chong Chen
Department of Information Management, School of Government, Beijing Normal University, China

Abstract
Interdisciplinary studies are becoming increasingly popular, and research domains of many experts are becoming diverse. This phenomenon brings difficulty in recommending experts to review interdisciplinary submissions. In this study, an Author–Subject–Topic (AST) model is proposed with two versions. In the model, reviewers' subject information is embedded to analyse topic distributions of submissions and reviewers' publications. The major difference between the AST and Author–Topic models lies in the introduction of a 'Subject' layer, which supervises the generation of hierarchical topics and allows sharing of subjects among authors. To evaluate the performance of the AST model, papers in Information System and Management (a typical interdisciplinary domain) in a famous Chinese academic library are investigated. Comparative experiments are conducted, which show the effectiveness of the AST model in topic distribution analysis and reviewer recommendation for interdisciplinary studies.

Keywords
Author–Subject–Topic model; expert finding; expert recommendation; reviewer assignment; reviewer recommendation

1. Introduction

Journal editors play an important role in the process of periodical review. A challenging task of editors is to find proper reviewers whose research interests fit the submissions suitably, which is often referred to as the reviewer assignment problem. Conventionally, the reviewer assignment may be subjective, which potentially leads to a certain degree of unfairness. An increasing number of researchers in the academic field have begun to explore intelligent and automatic recommendation of reviewer candidates.

Some preliminary studies have used various models to retrieve experts in different scenarios, such as the language model [1,2], data fusion methods [3], discriminative models [4] and social network methods [5]. However, these studies ignore the diversities of experts' research interests explicitly. For example, experts may have broad research interests in different fields. Thus, many conventional models potentially fail to capture their expertise comprehensively.

A central task in effective reviewer recommendation is to infer an author's interest on the basis of his or her publications. Then, the generative topic model, which is utilised to model topic distributions of textual data, may help model the

Corresponding author:
Jian Jin, Department of Information Management, School of Government, Beijing Normal University, Beijing 100875, China.
Email: jinjian.jay@bnu.edu.cn

interest and represent a document as mixtures of topical components. Different generative topic models can be applied to analyse experts' research interests, such as the Author–Topic (AT) model [6]. However, topic distributions of each author in the AT model are reckoned depending on his or her own documents only, and the similarity of documents from different authors is ignored. Specifically, author's interest in the AT model is represented as a mixture of topic variables that are associated with authors. In this model, documents are confined to be grouped according to the author–topic distribution only rather than fundamental semantic topics in documents. As a result, topic-similar documents from different authors can be assigned various topic labels. Such limitations affect the overall performance in modelling author's interest, especially when analysing documents in interdisciplinary studies. Considering these limitations, the Author–Interest–Topic (AIT) model was proposed in Kawamae [7]. AIT is a typical hierarchical topic model that assigns a latent variable to each document. The latent variable denotes an author's interest and allows documents with similar topics from different authors to be grouped according to a subject label.

Some unsupervised models, such as AT and AIT, can be utilised to model an author's interest for reviewer recommendation. AT and AIT observe only the appearance of words. However, supervised models are generally expected to perform better than unsupervised models. In some scientific digital libraries, metadata (such as category information about documents) are associated with textual data. These metadata, neglected by unsupervised models, can be used as visual data for improved modelling of the distribution of latent topics in documents and making solid assumptions.

Inspired by the document class layer in AIT, a novel topic model called Author–Subject–Topic (AST) is used to infer an author's interest with external metadata for reviewer recommendation. In AST, a high level of concept is generated and topics are then generated in accordance with the assumed concept. The high level of concept aims to capture the similarity of documents from different authors. Unlike the AIT model that provides each document an unsupervised class label, the advantage of the AST model lies in its exploitation of predefined subject category labels in the academic database. These additional subject category labels in the academic database are utilised to generate each topic within a supervised subject and model authors' interest accurately. To evaluate the proposed AST model, categories of experiments are conducted with different parameter settings on a dataset on a typical interdisciplinary research field, and encouraging results are obtained. The results demonstrate the performance of the AST model and suggest recommending proper experts for academic peer reviews depending on topics in submissions and research interests of experts.

The contributions of this research are threefold. First, a novel probabilistic model is introduced for reviewer recommendation. Experts' interdisciplinary research interests are considered, and a group of reviewers whose research domains are expected to cover the research domains of submissions in interdisciplinary studies are recommended. This critical and practical concern is neglected by many existing models. Second, AST uses not only text information but also other available metadata information. It aims to capture latent connections among authors, subject labels and words. Finally, categories of experiments are conducted using a real Chinese academic dataset to evaluate the performance of the proposed approach. In particular, two evaluation metrics, the topic coverage and the average symmetric Kullback–Leibler (KL) divergence, are utilised to benchmark the matching degree between recommended reviewers and submissions. The results demonstrate the validity of the AST model in expert recommendation for reviewing interdisciplinary studies.

This work extends a previous study that was presented in the Asia Information Retrieval Societies Conference in 2015 [8]. In Mou et al. [8], the AST model was initially introduced by considering predefined subject category labels. However, in this study, two versions of the AST model are introduced. The model presented in Mou et al. [8] is named as AST_1, whereas the new version is named as AST_2. The major differences between the two versions lie in their modelling of the subject information of each document, that is, whether a subject label is assigned according to the category distribution or a subject distribution in the 'Subject' layer. Additional explanations and the corresponding approaches on parameters are explained in section 3. In section 4, categories of experiences are made on two versions of the AST model and other comparative approaches for benchmarking.

The remainder of the research is organised as follows. In section 3, the AST model is proposed with two versions of implementations, and a Gibbs sampling approach is utilised for parameter estimation. Comparative experiments are conducted and the results are presented in section 4. Finally, conclusions are elaborated in section 5.

2. Related work

2.1. Information retrieval for expert recommendation

Expert recommendation in information retrieval finds proper experts from a large number of reviewer candidates to ease information overload. In 1992, latent semantic indexing was initially utilised for analysing the topic distribution of experts' abstracts to recommend reviewers for each submission [9]. In Craswell et al. [10] and Soboroff et al. [11], expert recommendation was involved in the Text REtrieval Conference Enterprise Track, which motivates researchers to explore different models. For example, an author-centred language model and a document-centred language model were proposed for expert recommendation in Balog et al. [1,12].

Hettich and Pazzani [13] modelled expert recommendation as a retrieval problem and utilised the term frequency–inverse document frequency (TF-IDF) weighting to obtain the match score between experts and proposals. Some metadata (such as co-authors, H-index and document quantity), which are considered to be important in modelling experts' research interests, are incorporated for expert recommendation. For example, a logistic regression classifier was built for expert recommendation in Fang et al. [4], and it combines 10 types of metadata, including document and basic association features. Moreira and Wichert [3] used data fusion, in which multiple sensors are built, and each sensor is responsible for a kind of specific information on experts. Han et al. [14] utilised experts' social relationship in recommending programme committee members. Hofmann et al. [15] extracted some contextual factors and combined them with content-based retrieval scores to find similar experts within an organisation. Some heuristic rules and knowledge were considered in reviewer recommendation in Liu et al. [16], in which different aspects are involved, such as reviewer expertise, title and project experience. A multiple objective optimisation problem is formulated to maximise the total expertise level of recommended experts and avoid conflicts between reviewers and authors.

However, the use of metadata alone by these information retrieval-based approaches may be insufficient to cover experts' interests. Thus, much space is available to improve the performance of expert recommendation.

2.2. Optimisation problems for expert recommendation

Tang et al. [17] proposed an expert matching framework, in which different constraints are considered in a convex cost flow problem; they also developed an efficient matching algorithm for online expert recommendation. Xu et al. [18] grouped similar proposals and invited reviewers to provide their expertise on different research fields explicitly. Then, an optimisation problem is formulated to find proper experts in accordance to the matching degree between proposal groups and reviewers' expertise. Silva et al. [19] built an integrated system, in which an optimisation problem is formulated to recommend reviewers for research projects. In this system, several aspects are considered, including relevance degree between reviewers and projects, connectivity between authors and reviewers, and reviewer expertise.

Some scholars have introduced the academic network to expert recommendation [20–22]. They build a network of reviewer candidates and extract their correlation degree, such as co-authors and citations. All of these extracted relations are also regarded to mirror experts' authority and influence. For example, a convex optimisation framework was formulated on a network-based model in Liu et al. [23], and it incorporates reviewer expertise, authority and diversity. Random Walk with Restart is utilised to involve diverse research backgrounds. Some studies have also investigated the connectivity between reviewers and authors. For example, three aspects of researchers (expertise relevance, individual connectivity and institutional connectivity) were analysed in Yang et al. [24]. These aspects are integrated in a score-based method to recommend experts. Li et al. [25] proposed an intelligent approach for context-aware reviewer assignment. In this approach, reviewer expertise, institution relevance and co-authorship between reviewers and authors are considered. The influence between stringent or lenient reviewers is balanced via a cross-assignment paradigm. Similarly, the matching degree between reviewers and submissions and manuscript and reviewer diversities were consolidated in a multi-objective mixed integer programming problem to assign a group of reviewer candidates for a group of submissions in Wang et al. [26]. A two-phase stochastic-biased greedy algorithm is utilised to analyse this multi-objective, mixed integer programming problem.

Some widely applied approaches in the field of decision science, such as fuzzy optimisation and analytic hierarchy process (AHP), also facilitate reviewer assignment. For example, some researchers have argued that reviewer candidates are expected to be grouped together. Accordingly, a fuzzy model was proposed in Das and Gocken [27], which attempts to maximise the matching degree between submission and reviewer groups and minimise the cost in forming reviewer groups. Similarly, another fuzzy model–based approach was introduced in Tayal et al. [28]. A type 2 fuzzy set is utilised to estimate the score of reviewer's expertise, and a second fuzzy set involving three keywords is adopted to represent submissions. Next, a fuzzy function with different practical constraints is designed to evaluate the matching degree between reviewer candidates and submissions. Li and Watanabe [29] proposed an integrated framework for reviewer recommendation that reckons reviewers' expertise and their relevance to submissions. The expertise and relevance degrees are determined by an AHP-based approach. An optimisation problem is also designed to maximise the matching degree between reviewer candidates and submissions. Topic relevance, expert quality and researcher connectivity were analysed to profile researchers and an AHP-based approach was applied for reviewer recommendation in Sun et al. [30].

However, these studies do not explicitly model the interest of authors according to latent topics in the content of their publications, which negatively affects the overall performance in expert recommendation.

2.3. Topic models for modelling research interest

Wei and Croft [31] utilised the classical latent Dirichlet analysis (LDA) for information retrieval and found that the topic model demonstrates considerable improvements in information retrieval compared with language models. However, author information is neglected in LDA. Considering the limitations of LDA for modelling the interests of authors, different generative topic models are developed.

Some scholars have extended LDA for expert recommendation. The AT model [6], which is a generative model, regards each author as a mixture of topics. Kou et al. [32] considered the topic distribution in different submissions. First, multiple topics of experts and submissions are extracted by the AT model. Next, a group of experts are recommended depending on the different weight topics in submissions. In the system in Kou et al. [33], the topic distribution of each submission and each reviewer was estimated by the expectation–maximisation (EM) algorithm with the AT model. Then, an optimisation problem is formulated for expert recommendations.

Various topic models have extended the AT model by involving different kinds of information. An Author–Topic over Time (AToT) model was proposed in Xu et al. [34], in which topic evolution is considered. This model aims to discover the changing patterns of latent topics and users' interest over time. Tang et al. [35] extended the AT model by complementing the publication venue and proposed an Author–Conference–Topic (ACT) model, in which each author's topic is correlated with words and conference stamps. The Citation–Author–Topic (CAT) model introduces the citation information for the retrieval of experts [36]. The Author–Conference–Topic–Connection (ACTC) model [37] and the Author–Citation–Venue–Topic model [38] extend the ACT and CAT models by adding a subject layer and modelling the citation information, respectively. Most of these models can be categorised as supervised models. In these models, different types of information, such as conference names and citation information, are utilised and integrated to analyse the topic distribution. Notably, the proposed AST model is a supervised topic model. Compared with supervised models, additional subject category labels, which are indicated explicitly in the academic database, are exploited. These subject category labels help AST supervise the generation of topics within a high level of subject and capture the similarity of documents of different authors, which facilitates modelling of experts' interest and topics in their publications accurately.

The Author–Persona–Topic (APT) model introduces a persona layer. This approach allows authors' documents to be divided into clusters, each of which is related to that persona [39]. The AIT model [7] and the Latent Interest–Topic (LIT) model [40] extend the APT model. They can be regarded as hierarchical topic models. In these models, an interest layer and an author class layer are introduced to allow information sharing among authors. Such assumption will reduce noises to a certain degree. All of them are unsupervised models with topics that are generated under different assumptions and observe only the appearance of words. However, metadata associated with textual data, which are helpful in making solid and practical assumptions with explicit information about authors and papers, are neglected by these unsupervised models. Compared with unsupervised models, the proposed AST supervises the topic distribution in accordance with additional subject category labels, and predefined category labels provide a prior about the subject distributions. Accordingly, accurate results regarding the topic distribution can be obtained.

3. The AST

3.1. The probabilistic generative model

Intuitively, when people talk about a discipline system in a certain field, a hierarchical structure may be formed in their minds. The root is a macro research concept, and the leaves are specific pieces of contents that are related to the macro concept. The same system can also be used to describe an expert's expertise. In particular, an expert may interest some 'Subjects', and a topic distribution exists in each of these subjects that helps explain some detailed knowledge.

A supervised 'Class' layer in the AIT model [7] and a supervised 'Persona' layer in the APT model [39] have been reported. Arguably, supervised models are more accurate than unsupervised models. Hence, a novel topic model called the AST model, which is an extension of the AT model, is presented in this research. Two versions of the AST model, namely, AST_1 and AST_2, are implemented. In Figure 1, the AT, AST_1 and AST_2 models are compared.

In the proposed AST models, a topic is generated according to a subject–topic distribution and then a word is generated according to a topic–word distribution. First, an author a picks a subject from his or her own subject distribution. Next, similar to Ramage et al. [41], the author-specific label projection matrix $L^{(a)}$ is introduced, each row of which denotes one subject label of the author a. For example, four subjects are considered regarding a, and the indicator vector of a is $\Lambda^{(a)} = \{0,1,1,0\}$. Therefore, a is interested on subjects 2 and 3. Then, the corresponding $L^{(a)}$ can be denoted as

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

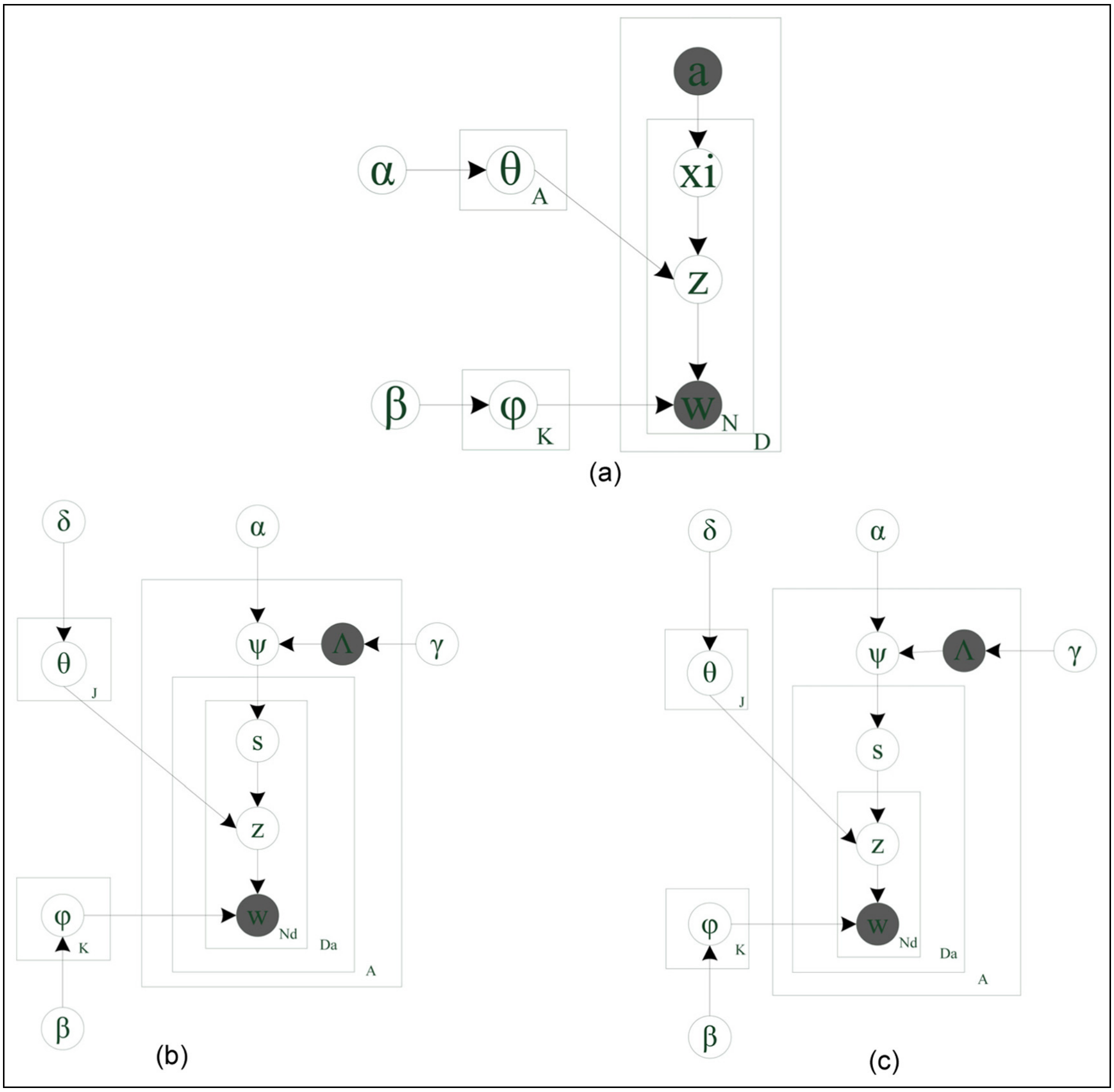

Figure 1. The Author–Topic model and the Author–Subject–Topic model(s): (a) the AT model, (b) the AST_1 model and (c) the AST_2 model.

where $\psi^{(a)}$ denotes the subject distribution of a's research in the 'Subject' layer. Each topic is selected from the topic distribution that is related to a chosen subject, and each word is selected from a word distribution that is related to the chosen topic. Finally, one submission can be generated iteratively. For clarity, notations of the AST_1 and AST_2 models are listed in Table 1.

The formal generative process of the AST_1 model can be described as follows:

1. For each topic $k \in \{1, ..., K\}$:
2. Generate $\varphi_k \sim Dir(\bullet|\beta)$
3. For each subject $j \in \{1, ..., J\}$:
4. Generate $\theta_j \sim Dir(\bullet|\delta)$
5. For each expert $a \in \{1, ..., A\}$:
6. For each subject $j \in \{1, ..., J\}$:
7. Generate $\Lambda_j^{(a)} \in \{0, 1\} \sim Bernoulli(\bullet|\gamma_j)$
8. Let $\alpha^{(a)} = L^{(a)} * \alpha$
9. Generate $\psi^{(a)} \sim Dir(\bullet|\alpha^{(a)})$
10. For each document $d \in \{1, ..., D\}$:

Table 1. Notations in the proposed topic model.

Symbol	Description
α, β, δ	Hyperparameters of Dirichlet distributions
γ	Hyperparameters of Bernoulli distribution
Da	The total number of expert *A*'s documents
Nd	The total number of words in document *D*
s, z, w	*s* for subject, *z* for topic, *w* for word
J	The total number of subjects
K	The total number of topics
A	The total number of authors
θ	A $J \times K$ matrix that indicates subject–topic distribution
φ	A $K \times V$ matrix that indicates topic–word distribution
ψ	An $A \times J$ matrix that indicates expert–subject distribution
Λ	The indicator if the subject belongs to expert *A*
$L^{(a)}$	The matrix to project α
x_i	Author *x* associated with the *i*th word in document *D*

11. For each word i in $\{1, ..., N\}$:
12. Generate $s_i \sim Mult(\bullet|\psi^{(a)})$
13. Generate $z_i \sim Mult(\bullet|\theta_{si})$
14. Generate $w_i \sim Mult(\bullet|\varphi_{zi})$

Similarly, the formal generative process of the AST_2 model is described.

1. For each topic $k \in \{1, ..., K\}$:
2. Generate $\varphi_k \sim Dir(\bullet|\beta)$
3. For each subject $j \in \{1, ..., J\}$:
4. Generate $\theta_j \sim Dir(\bullet|\delta)$
5. For each expert $a \in \{1, ..., A\}$:
6. For each subject $j \in \{1, ..., J\}$:
7. Generate $\Lambda_j^{(a)} \in \{0, 1\} \sim Bernoulli(\bullet|\gamma_j)$
8. Let $\alpha^{(a)} = L^{(a)} * \alpha$
9. Generate $\psi^{(a)} \sim Dir(\bullet|\alpha^{(a)})$
10. For each document $d \in \{1, ..., D\}$:
11. Generate $s_i \sim Mult(\bullet|\psi^{(a)})$
12. For each word i in $\{1, ..., N\}$:
13. Generate $z_i \sim Mult(\bullet|\theta_{si})$
14. Generate $w_i \sim Mult(\bullet|\varphi_{zi})$

3.2. Parameter estimation

In this research, Gibbs sampling is utilised to infer parameters in the AST model(s). Gibbs sampling is a widely utilised approach for parameter estimation in many topic model–based approaches [42,43].

The joint probability of the AST model(s) is presented in equation (1). Three sets of parameters, ψ, θ and φ, should be estimated. ψ denotes the probability that a subject is chosen by an expert. θ denotes the probability that a topic is selected by a subject. φ denotes the probability that a word is generated by a topic. Notably, Λ is an observed value. Given a particular Λ, γ becomes '*D*-separated' with the rest of the model, and the hyperparameter α is the only one restrained to the labelled subjects

$$p(w, s, z|\alpha, \beta, \delta, \Lambda, \gamma) = \int p(s, \psi|\alpha, \Lambda, \gamma) d\psi p(z, \theta|s, \delta) d\theta p(w, \varphi|z, \beta) \tag{1}$$

Considering that w is observed, the conditional probability $p(s_{adi}|s_{-adi}, \alpha, \Lambda, \gamma, z)$ and $p(z_{adi}|z_{-adi}, w, s, \beta, \delta)$ can be inferred. s_{adi} and z_{adi} represent the subject and topic of token i in author a's dth paper, respectively. s_{-adi} and z_{-adi} represent the vector of subject assignments and vector of topic assignments in all corpora except for the ith token in author

a's dth paper, respectively. The latent 'Subject' layer is different in AST_1 and AST_2. Thus, the sampling equations for $p(s_{adi}|s_{-adi}, \alpha, \Lambda, \gamma, z)$ should be derived individually. For simplicity, hyperparameters are omitted in the following parts. If $s_{adi}=j$ in AST_1, then the following condition can be obtained

$$\begin{aligned} p(s_{adi}=j|s_{-adi}, z) &\propto \frac{p(s, z)}{p(s_{-adi}, z_{-adi})} \\ &= \frac{p(s)p(z|s)}{p(s_{-adi})p(z_{-adi}|s_{-adi})} \\ &= \frac{n^j_{a,-adi}+\alpha}{\sum_j^J \left(n^j_{a,-adi}+\alpha\right)} \frac{n^k_{j,-adi}+\delta}{\sum_k^K \left(n^k_{j,-adi}+\delta\right)} \end{aligned} \tag{2}$$

where $n^j_{a,-adi}$ represents the number of subject j that a is assigned to, except adi, the ith token in author a's dth paper, and $n^k_{j,-adi}$ represents the number of topic k that subject j is assigned to, except adi. In addition, the hyperparameter α is restrained with each author's Λ.

The sampling equation of AST_2 can be derived as

$$\begin{aligned} p(s_{ad}=j|s_{-ad}, z) &\propto \frac{p(z_{ad}|s_{ad})p(s_{ad})}{p(z_{-ad}|s_{-ad})p(s_{-ad})} \\ &= \frac{\Delta(n_a+\alpha)\Delta(n_j+\delta)}{\Delta(n_{a,-ad}+\alpha)\Delta(n_{j,-ad}+\delta)}, \quad n_a=\left\{n^j_\alpha\right\}_{j=1}^J, n_{a,-ad}=\left\{n^k_{j,-ad}\right\}_{j=1}^J \\ &= \frac{n^j_{\alpha,-ad}+\alpha}{\sum_j^J \left(n^j_{\alpha,-ad}+\alpha\right)} \frac{\prod_k^K \Gamma\left(n^k_j+\delta\right)\Gamma\left(\sum_k^K n^k_{j,-ad}+\delta\right)}{\prod_k^K \Gamma\left(n^k_{j,-ad}+\delta\right)\Gamma\left(\sum_k^K n^k_j+\delta\right)}, \quad n_j=\left\{n^k_j\right\}_{k=1}^K, n_{j,-ad}=\left\{n^k_{j,-ad}\right\}_{k=1}^K \end{aligned} \tag{3}$$

Similarly, s_{ad} represents the subject in author a's dth paper; s_{ad} represents the vector of subject assignments in all corpora except ad, author a's dth paper; z_{ad} represents the vector of topic assignments in author a's dth paper, whereas z_{-ad} represents the vector of subject assignments in all corpora except for author a's dth paper; $n^j_{a,-ad}$ represents the number of subject j that author a is assigned to, except ad; $n^k_{j,-ad}$ represents the number of topic k that subject j is assigned to, except ad; n^k_j represents the number of topic k that subject j is assigned to; n_a denotes the number vector of subjects that author a is assigned to; $n_{a,-ad}$ denotes the number vector of subjects that author a is assigned to, except ad; n_j denotes the number vector of topics that subject j is assigned to; $n_{j,-ad}$ denotes the number vector of topics that subject j is assigned to, except ad. Moreover, the hyperparameter α is restrained with each author's Λ.

The conditional probability, $p(z_{adi}|z_{-adi}, w, s, \beta, \delta)$, in AST_1 and AST_2, as shown in Figure 1, is expected to have the same form in terms of sampling equations. If $z_{adi}=k$, then the following condition can be obtained

$$\begin{aligned} p(z_{adi}=k|z_{-adi}, w, s=j) &\propto \frac{p(w, z)}{p(w_{-adi}, z_{-adi})} \\ &= \frac{p(z)(w|z)}{p(z_{-adi})p(w_{-adi}|z_{-adi})} \\ &= \frac{n^v_{k,-adi}+\beta}{\sum_v^V \left(n^v_{k,-adi}+\beta\right)} \frac{n^k_{j,-adi}+\delta}{\sum_k^K \left(n^k_{j,-adi}+\delta\right)} \end{aligned} \tag{4}$$

where $n^v_{k,-adi}$ denotes the number of word v that topic k is assigned to, except adi, and $n^k_{j,-adi}$ records the number of topic k that subject j is assigned to, except adi. In accordance with Dirichlet multinomial conjugation, the parameters of two models can be estimated as

$$\psi_{a,s} = \frac{n_a^{(s)}+\alpha}{\sum_s^S n_a^{(s)}+S\alpha}, \ \theta_{s,k} = \frac{n_s^{(k)}+\delta}{\sum_k^K n_s^{(k)}+K\delta}, \ \phi_{k,v} = \frac{n_k^{(v)}+\beta}{\sum_v^V n_k^{(v)}+V\beta} \tag{5}$$

3.3. Model comparisons

The difference between AST_1 and AST_2 is that each document in AST_2 is given a subject label in accordance with the category distribution, which can be regarded as hard clustering. Conversely, each document in AST_1 is given a subject distribution in the 'Subject' layer, which can be regarded as soft clustering. Such soft clustering is practically important because many research papers may cover several subjects, especially in an interdisciplinary field. For example, a paper entitled 'Study of Information Science Development and Education' falls into a category entitled 'Intelligence, Information Science' although it also discusses Education.

In AT, documents are grouped by a given author under a single author–topic distribution. Hence, this model fails to consider the similarity among different documents. Compared with AT, visual predefined category labels regarding an author's interest are introduced as a supervised layer in AST, which aims to describe the knowledge of different authors. This approach allows documents with similar topics of different authors to be grouped according to subjects. Such difference helps AST model the wide range of experts' research interests and facilitate clustering of authors' documents precisely.

Similar to AST, the hierarchical latent Dirichlet analysis (hLDA) models topics in different levels [42]. In hLDA, a document path is utilised to organise a hierarchical concept of topics, which helps capture the different layers of topics and mirror the underlying topic generality. Indeed, hLDA is applicable for analysing a hierarchical structure for a discipline system. However, hLDA ignores author information, especially for publications with multiple authors, and exploits the value of predefined information in scientific library. Such difference helps AST be distinct from other hierarchical topic models. Specifically, indicator variables in AST are utilised to represent author's interests and describe which mixture of topics each document belongs to. Hence, the AST model can be a better fit than hLDA.

AST_2 and AIT can be used to assign a class label to a single document. The two models allow latent variables to be associated with author's interest and help documents with similar topics from different authors to be grouped according to a subject label. However, the difference between AST_2 and AIT lies in the visual predefined category labels of authors. Compared with AIT, predefined category labels in AST_2 provide a prior about the subject distribution in accordance with the visual information in scientific library.

Therefore, the AST model describes a wide variety of authors' interests, in which visual information helps boost the overall performance in modelling a broad range of topics, which is essentially the core of interdisciplinary studies.

4. Experimental study and discussion

4.1. Dataset

A dataset in Information Systems and Management was built to evaluate the effectiveness of the proposed model. Information Systems and Management is a typical interdisciplinary domain that characterises a modern trend in the research field. To obtain the corresponding dataset, first, 256 scholars holding a project funded by the National Natural Science Foundation of China from 2004 to 2014 in this area were selected as seed experts. Then, research papers and co-authors of these seed experts were crawled from the WANFANG DATA, which is a comprehensive academic database in China. A total of 5519 scholars and 75,880 abstracts were obtained. All of these scholars were regarded as reviewer candidates in the subsequent experiments.

Careful examination of the dataset shows some interesting phenomena. Scholars in Information Systems and Management are found to have a very broad range of research interests, which are indicated by labels regarding authors' research subjects in WANFANG DATA. It can be seen from Figure 2 that experts e1 and e2 present a diversity of research interests. Similar cases can be easily observed in this interdisciplinary research area.

Table 2 presents the top 10 research hotspots and top 10 subject category labels regarding all publications within this dataset. In WANFANG DATA, research hotspots of an author are listed in his or her profile page. The profile page is initially provided by WANFANG DATA depending on an author's publication list, and the author can revise his or her profile page manually after claiming. In the webpage of each publication in WANFANG DATA, a subject category label is provided in accordance with the main scope of the corresponding journal or conference. In this study, hotspots of all 5519 scholars and subject category labels of all 75,880 papers were collected, and top 10 hotspots and top 10 category labels were listed. The column of research hotspots in Table 2 lists some popular keywords in the area of Information Systems and Management, which are provided by the WANFANG DATA. The column of category labels lists the major subareas in this specific field. A wide range of research interests can be found in this field, such as industrial technology, medical science, economics and computer science. The research area of Information Systems and Management is indeed a typical interdisciplinary field in the viewpoint of either macro- or microscopic statistics.

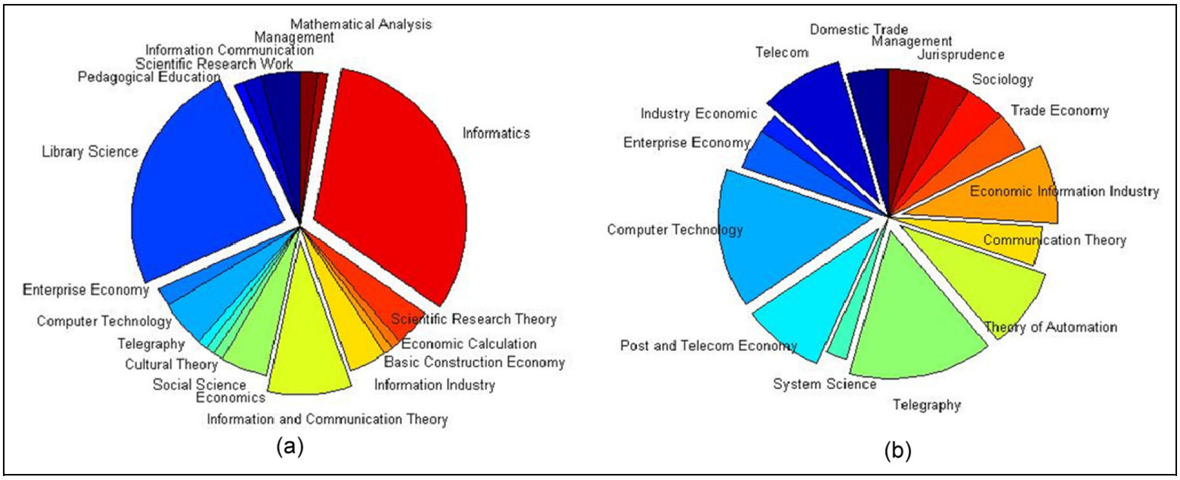

Figure 2. Category distribution of two experts: (a) expert e1 and (b) expert e2.

Table 2. Top 10 research hotspots and category labels.

Research hotspot (top 10)	Category labels (top 10)
Data mining	Industrial technology
Genetic algorithm	Medical science
E-commerce	Economics
Supply chain	Mathematical sciences
Knowledge management	Astronomy
AHP	Environmental science
Emulation	Agricultural sciences
Ontology	Computer science
Data warehouse	Clinical medicine
Web service	Military

AHP: analytic hierarchy process.

To evaluate the performance of the AST model(s) and the AT model, 500 experts with more than five category labels were randomly selected from the dataset for benchmarking. Experts with more than five category labels can be generally referred to have diversified research interests. If too few category labels are selected (such as one or two), then the AST model is argued to be similar to the AT model. For effective comparisons with the AT model, those with more than five category labels were selected. The Natural Language Processing and Information Retrieval (NLPIR) was then utilised for Chinese word segmentation and part-of-speech (POS) tagging on paper abstracts. Only nouns, adjectives and noun verbs with appearance of greater than or equal to five were used. Finally, 601 subjects and 15,533 unique words were left in the 10,843 papers.

In this study, 10% of words in each document were selected randomly as the testing dataset and the rest were used as the training dataset. Notably, the experiments aimed not to examine subject distributions of authors in the testing set but to evaluate the generality of the proposed model. Subject distributions in the testing set must be consistent with the training set. Specifically, the testing set is suggested to be sampled from the original dataset with the same distribution. Hence, for each document, most words are utilised as the training set and others are utilised as the testing set to estimate the perplexity. Similar approaches for the perplexity estimation, such as AToT [34], ACTC [37] and LIT [40], have also been reported.

For simplicity, the three hyperparameters α, β and δ in the AST model(s) were set to 50/K, 0.01 and 0.01, respectively. The hyperparameters α and β in the AT model were also set to 50/K and 0.01, respectively. Each of the authors' 'Subject' layers in the AST_1 and AST_2 models was supervised and was therefore set in accordance with the actual subject labels. The 'Topic' layers of the AST_1, AST_2 and AT models were set in the range from 50 to 500. A total of 1000 iterations were conducted in the Gibbs sampling on a machine with Core i5 3.20 GHz processor.

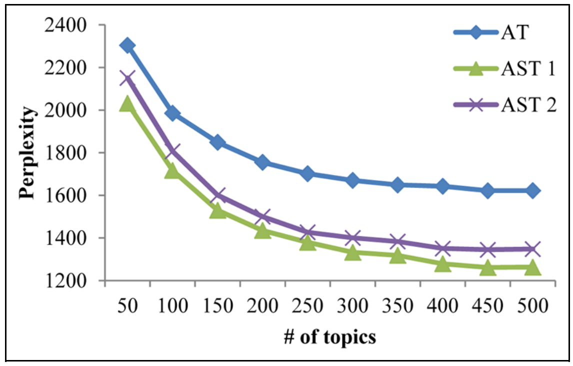

Figure 3. The perplexity of the AST model(s) and AT model.

4.2. Evaluating the AST model

4.2.1. Results of perplexity. Perplexity, which is a popular evaluation metric to assess the predictive power of a topic model, was applied to estimate performance. This metric is defined as the exponential of the negative normalised predictive geometric mean per word likelihood. The model with low perplexity is better than that with high perplexity. Therefore, the model fits the data better. Mathematically, the model is represented as

$$Perplexity = \exp\left\{-\frac{\sum_d \log(p(w_d|a_d))}{\sum_d N_d}\right\} \tag{6}$$

A critical parameter in different generative topic models, such as LDA and AT, is the number of topics. In this study, the number of topics was evaluated. Figure 3 shows the comparison of perplexity values of the three models. As shown in this figure, the perplexity of the three models decreases first as the number of topics increases and then converges to a relative stationary state when the topic is set to 450. Accordingly, 450 topics were applied in the subsequent experiments. The AST model(s) present a considerably lower perplexity value than AT, which shows that AST is better. The reason is that the 'Subject' layer of the AST model(s) has a high degree of confidence and is designed to capture the inter-similarity of documents to reduce noises to a certain degree. AST_1 performs better than AST_2. Therefore, the subject distribution assumption in AST_1 fits the data better than a solitary subject label in AST_2.

4.2.2. Reviewer recommendation using the AST model. A typical application of the AST model(s) is to recommend experts in a specific field for periodical review. In the subsequent experiments, 10 abstracts were randomly selected from the dataset. They were regarded as submissions in evaluating the effectiveness of AST on reviewer recommendation.

The NLPIR was first used to pre-process these submissions. Then, the posterior probability $p(a|txt)$ was utilised to estimate the ability of an expert a to get through the given submission txt. In accordance with the Bayesian formula, the following condition can be derived

$$p(a|txt) = \frac{p(a)p(txt|a)}{p(txt)} \propto p(a)p(txt|a) \tag{7}$$

where $p(txt|a)$ estimates the probability that the paper txt is generated given an expert a. $p(a)$ and $p(txt)$ indicate the prior probability of an expert a and a given submission txt, respectively. $p(txt)$ is the same for different experts and can therefore be ignored. $p(a)$ reflects the authority of experts and it can be assumed to follow a uniform distribution following Balog et al. [12]. Hence, the focus is on $p(txt|a)$ and the recommendation function can be

$$p(a|txt) = \prod_{w_i}\sum_s\sum_t p(a|s)p(s|t)p(t|w_i) \tag{8}$$

In this function, the given submission txt contains a list of words w_i. The probability of $p(a|s)$, $p(s|t)$ and $p(t|w_i)$ can be derived from a trained AST model. Accordingly, top@5, top@10 and top@20 authors were recommended as review candidates for a particular paper txt.

In consideration of a practical scenario for the reviewer recommendation, the difference between this research and previous studies [35,37,39] is that the research interests of recommended experts in this study aim to match and cover

Table 3. The KL divergence in three models.

		Top@5	Top@10	Top@20
AT	Min	0.183	1.024	3.834
	Mean	0.319	1.392	5.307
	Max	0.488	1.786	6.860
AST_2	Min	0.193	1.155	4.102
	Mean	0.359	1.724	6.069
	Max	0.556	2.494	8.777
AST_1	Min	0.199	1.000	3.949
	Mean	0.372	1.658	6.246
	Max	0.579	2.581	9.032

KL: Kullback–Leibler; AT: Author–Topic; AST: Author–Subject–Topic.

Table 4. Topic coverage in the AST_1 model.

	Top@5	Top@10	Top@20
Min	0.466	0.527	0.676
Mean	0.506	0.658	0.784
Max	0.550	0.745	0.869

AST: Author–Subject–Topic.

the entire research domain of the submission. This condition leads to a dilemma in evaluating the performance of the recommend reviewers because the reviewer invitation of different journals is generally regarded as privileged information. Although these data are publicly available, the manual reviewer assignment can be barely guaranteed as an ideal gold standard. To evaluate the performance of expert recommendation, different approaches, such as pooled relevance judgements [44], are used. However, they cannot be applied in this research due to the following two reasons. First, pooled relevance judgements require manual evaluation. However, the dataset involves intensive interdisciplinary studies, which need a time-consuming process for annotators who are familiar to different research fields and research interests of many reviewer candidates to identify the 'ground-truth'. Second, the pooled relevance judgements only evaluate precision. In this study, a group of experts whose research domains are expected to cover the research domains of the submission as completely as possible were recommended.

Hence, the average symmetric KL divergence and topic coverage were applied to evaluate the performance of reviewer recommendation. The average symmetric KL divergence was utilised to measure the information diversity among different reviewer candidates. In this research, it can be described as the distance between reviewer *i* and reviewer *j* with their topic distributions. A large value of KL divergence indicates considerable difference between two reviewers. The average symmetric KL divergence is defined as

$$sKL(i,j) = \sum_{t}^{T}\left[\theta_{it}\log\frac{\theta_{it}}{\theta_{jt}} + \theta_{jt}\log\frac{\theta_{jt}}{\theta_{it}}\right] \tag{9}$$

where θ_{it} represents the probability of reviewer *i*'s topic *t* and θ_{jt} represents that of reviewer *j*'s topic *t*. Table 3 shows the comparative results of the symmetric KL divergence for AT and AST. The 'Min', 'Mean' and 'Max' rows present the minimal, average and maximal values of symmetric KL divergence between all pairs of recommend reviewers for all 10 submissions. For top@5, the corresponding 'Min', 'Mean' and 'Max' rows present the minimal, average and maximal values of symmetric KL divergence among all top five recommend reviewers.

As shown in Table 3, recommended reviewer candidates' average symmetric KL divergence of the AST_1 and AST_2 models at top@5, top@10 and top@20 are larger than those of AT. Therefore, the topics of each reviewer in AST are more distinguished than those in AT. The objective of this experiment is to recommend a group of reviewers whose interest covers the submission topics as completely as possible. Therefore, AST is better than AT for reviewer recommendation that involves interdisciplinary studies.

The larger average symmetric KL divergence of the AST model(s) than that of the AT model is because the latter model only infers topics in the document layer regarding each independent author, which leads to an unavoidable

Table 5. Top six topics with top eight words in SUBJECT #358.

Topic #389	Probability	Topic #229	Probability
模型 (Model)	0.0495	系统 (System)	0.1006
验证 (Verification)	0.0182	实现 (Realisation)	0.0613
计算 (Calculation)	0.0172	应用 (Application)	0.0558
数学模型 (Mathematical model)	0.0048	开发 (Development)	0.0217
计算方法 (Computational method)	0.0043	体系结构 (Architecture)	0.0093
约束 (Constraint)	0.0037	数据库 (Database)	0.0059
预测 (Forecast)	0.0037	用户 (User)	0.0049
矩阵 (Matrix)	0.0025	面向对象 (Object-oriented)	0.0024
Topic #90	**Probability**	**Topic #222**	**Probability**
仿真 (Simulation)	0.2133	信息系统 (Information system)	0.0054
算法 (Algorithm)	0.0366	应用前景 (Application prospect)	0.0019
网络 (Network)	0.0230	预测方法 (Prediction technique)	0.0017
优化 (Optimisation)	0.0111	评价指标 (Evaluation index)	0.0016
模拟 (Simulation)	0.0044	约束条件 (Constraint condition)	0.0013
原型系统 (Prototype system)	0.0031	语义 (Semantics)	0.0011
可扩展性 (Extendibility)	0.0028	有效性 (Effectiveness)	0.0011
启发式算法 (Heuristic algorithm)	0.0018	结构 (Structure)	0.0010
Topic #234	**Probability**	**Topic #365**	**Probability**
设计 (Design)	0.1080	知识 (Knowledge)	0.0248
软件 (Software)	0.0280	描述 (Description)	0.0219
程序 (Programme)	0.0142	流程 (Technological process)	0.0155
核心 (Core)	0.0111	架构 (Framework)	0.0147
编程 (Programming)	0.0102	互联网 (Internet)	0.0076
图形 (Graphical)	0.0061	决策支持系统 (DDS)	0.0065
系统设计 (System design)	0.0052	数据仓库 (Data warehouse)	0.0065
自动化 (Automation)	0.0034	解决方案 (Solution proposal)	0.0044

decrease in the KL divergence over topics. However, the AST model(s) allow the supervised 'Subject' layer to be shared across different authors and documents to be clustered within the 'Subject' layer, which help reduce some noises and increase the KL divergence. For most indicators in Table 3, AST_1 performs better than AST_2. Therefore, the proposed approach for soft clustering of experts' documents performs better than the hard clustering in an interdisciplinary research field. In consideration of perplexity and the average symmetric KL divergence, only the AST_1 model was evaluated in the subsequent experiments.

The topic coverage aimed to measure the correlation between recommended reviewers and one submission, and it can be defined as the percentage of topic aspects that are covered by reviewers [45]

$$Coverage = \frac{N_a \cap N_{txt}}{N_{txt}} \tag{10}$$

where N_a represents the number of topics that N reviewer candidates cover and N_{txt} refers to the number of topics that the abstract covers. A total of 450 topics were selected in consideration of the previous experiments regarding perplexity. In this experiment, top 100 topics of each expert were selected to create the entire set of topics. Similarly, top 100 topics of each abstract were utilised to constitute the topic set of the abstract. Top 100 topics of each expert and top 100 topics of one submission were adopted to denote the expertise of expert a and the central concerns of one submission txt. Comparisons with different topic models in terms of topic coverage would be unfair because topics with different words were extracted from different approaches and the recommended number of topics in each model differed. The topic coverage was reported only on AST_1, as shown in Table 4.

The table shows encouraging results. For the recommend reviewers for each of the 10 submissions, the coverage value reaches 0.55 if five reviewers are recommended and the value increases to 0.87 if 20 reviewers are invited.

Table 6. Top six topics with top eight words in SUBJECT #353.

Topic #147	Probability	Topic #267	Probability
供应链 (Supply chain)	0.0258	客户 (Customer)	0.0558
企业 (Enterprise)	0.0195	电子商务 (E-commerce)	0.0482
效率 (Efficiency)	0.0127	信息系统 (Information system)	0.0449
供应链管理 (Supply chain management)	0.0093	业务流程 (Operation flow)	0.0160
供应商选择 (Vendor selection)	0.0092	客户关系管理 (CRM)	0.0140
据点扩张 (Base expansion)	0.0084	信息资源 (Information resource)	0.0112
顾客 (Customer)	0.0067	柔性 (Flexibility)	0.0108
知识重用 (Knowledge reuse)	0.0064	管理信息系统 (Management information system)	0.0088
Topic #365	**Probability**	**Topic #145**	**Probability**
知识 (Knowledge)	0.0248	探讨 (Discussion)	0.0337
描述 (Description)	0.0219	创新 (Innovation)	0.0264
流程 (Technological process)	0.0155	理论 (Theory)	0.0247
架构 (Framework)	0.0147	知识管理 (Knowledge management)	0.0150
互联网 (Internet)	0.0076	知识经济 (Knowledge economy)	0.0036
决策支持系统 (DDS)	0.0065	实证研究 (Empirical study)	0.0032
数据仓库 (Data warehouse)	0.0065	理论分析 (Theoretical analysis)	0.0032
解决方案 (Solution proposal)	0.0044	范式 (Paradigm)	0.0029
Topic #222	**Probability**	**Topic #307**	**Probability**
信息系统 (Information system)	0.0054	目标 (Target)	0.0274
应用前景 (Application prospect)	0.0019	环境 (Environment)	0.0249
预测方法 (Prediction technique)	0.0017	需求 (Demand)	0.0234
评价指标 (Evaluation index)	0.0016	优化 (Optimisation)	0.0168
约束条件 (Constraint condition)	0.0013	方案 (Scheme)	0.0162
语义 (Semantics)	0.0011	指标体系 (Index system)	0.0138
有效性 (Effectiveness)	0.0011	因素 (Factor)	0.0114
结构 (Structure)	0.0010	层次分析法 (AHP)	0.0063

CRM: customer relationship management; AHP: analytic hierarchy process.

Submissions are barely completely trained in AST because category labels of authors are not always available in the WANFANG DATA. In addition, no prior information about category labels of a particular expert is available for prediction on the interest of this author, and the only data contain relevant papers of this author. Without category labels of authors, relevant papers cannot be utilised as training data and a new round of training about the model cannot be executed for the prediction about the interest of that author. Therefore, word topics were assumed as an approximation to the submission topics. For each particular word w, the topic distribution $p(t|w)$ can be inferred. Accordingly, the topic distribution of a given abstract can be approximated as the average value of probabilities of all words in the abstract. A similar approximate approach was also applied in Tang et al. [35].

4.2.3. Example of the topic discovered. Tables 5–7 show three exemplary subjects discovered by AST_1. The three subjects are computer science, management science and clinical medicine. Each subject is shown with top six topics and top eight words. In Table 5, six topics are highly relevant, each of which describes a particular research topic in the subject of computer science, such as models and algorithms in topic #389, applications and development in topic #229 and computer programming design in topic #234. Similar phenomena can also be found in three other subjects.

Some topics in the above-mentioned tables are shared across different subjects, and these topics can be referred as an interdisciplinary study. In Tables 5 and 6, topics #222 and #365 have a high probability in the subjects of computer science and management science. Topic #222 concerns applications of information system, which is a typical interdisciplinary research field and is involved in the subjects of computer science and management science. Topic #365, which mainly concerns solutions of decision support systems, is involved in the subjects of computer science and management science.

Furthermore, subjects may consist of specific and interdisciplinary topics. Table 7 shows that topic #99 in the subject of clinical medicine, which may concern some related treatment of leukaemia, is a specific topic. Conversely, topics

Table 7. Top six topics with top eight words in SUBJECT #12.

Topic #180	Probability	Topic #377	Probability
鉴别诊断 (Differential diagnosis)	0.0335	患者 (Patient)	0.1983
信号 (Signal)	0.0261	差异 (Difference)	0.0594
扫描 (Scan)	0.0180	年龄 (Ages)	0.0311
电刺激 (Electric stimulation)	0.0172	统计学 (Statistics)	0.0209
志愿者 (Volunteer)	0.0170	临床 (Clinical)	0.0208
痛阈 (Threshold of pain)	0.0146	分析 (Analysis)	0.0163
表现 (Performance)	0.0143	评估 (Assessment)	0.0083
诊断 (Diagnosis)	0.0112	疾病 (Disease)	0.0080
Topic #99	Probability	Topic #132	Probability
回顾性分析 (Retrospective analysis)	0.0636	方法 (Method)	0.1199
不良反应 (Adverse reactions)	0.0518	患者 (Patient)	0.0548
复发 (Recrudesce)	0.0410	有效 (Effective)	0.0251
疗效 (Curative effect)	0.0287	显著性 (Significance)	0.0087
预后 (Prognosis)	0.0091	感染 (Infection)	0.0039
临床特点 (Clinical characteristics)	0.0086	静脉 (Vein)	0.0037
综合征 (Syndrome)	0.0080	影响因素 (Influential factors)	0.0016
白血病 (Leukaemia)	0.0067	临床效果 (Clinical result)	9.97E-04
Topic #223	Probability	Topic #165	Probability
比较 (Comparison)	0.1184	感染 (Infected)	0.1897
统计学 (Statistics)	0.0539	血清 (Serum)	0.0470
检测 (Testing)	0.0378	转录 (Transcription)	0.0393
药物 (Drug)	0.0219	中药 (Traditional Chinese medicine)	0.0253
实验 (Experiment)	0.0203	特异性 (Specificity)	0.0245
剂量 (Dose)	0.0193	阳性率 (Positive rate)	0.0229
志愿者 (Volunteer)	0.0086	地区 (Area)	0.0218
自动化 (Automation)	0.0034	解决方案 (Solution proposal)	0.0044

#222 and #365 are interdisciplinary topics. Their contents are general and observed in different subjects other than specific topics.

From the results, AST_1 is found to capture the correlation between topics and cluster words with specific or interdisciplinary topics effectively.

5. Conclusion

In this study, a novel graph model called the AST model is proposed with two versions for reviewer recommendation. A supervised 'Subject' layer is introduced in the AST model. This approach allows sharing of subjects among different authors, which helps cluster words and documents with few noises. Using a large volume of real data in the WANFANG DATA, Information Systems and Management (a typical interdisciplinary research field) is selected and comparative experiments are conducted. The results demonstrate that the AST model outperforms the classical AT model in modelling experts' research interests for reviewer recommendation. Compared with the AST_2 model, the AST_1 model performs better, which shows that soft clustering is more suitable than hard clustering for reviewer recommendation in an interdisciplinary research field.

Given the availability of an increasing amount of metadata, potential studies are suggested to extend the AST model in consideration of different factors, such as time stamp and link structure. In this study, top reviewers are recommended in accordance with the content of a submission. However, other aspects can also be considered in reviewer recommendation, such as the authority of a reviewer, the relations between a reviewer and authors of a submission, and the recent review willingness of a reviewer. In practical applications, a list of reviewers should be recommended to a group of submissions rather than for a single submission. These future studies are believed to ease the difficulty in reviewer recommendation. Future works should compare different recommendation algorithms and use additional persuading evaluation metrics to evaluate the effectiveness of the AST model.

Acknowledgements

We thank the anonymous reviewers for their comments that have contributed to important improvements of the paper.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

This work was supported by the Humanities and Social Sciences Research Youth Fund Project of Ministry of Education (No. 16YJC870006).

Note

1. http://ictclas.nlpir.org/

References

[1] Balog K, Azzopardi L and De Rijke M. A language modeling framework for expert finding. *Inform Process Manag* 2009; 45(1): 1–19.

[2] Balog K and de Rijke M. Finding similar experts. In: *SIGIR'07 Proceedings of the 30th annual international ACM SIGIR conference on research and development in information retrieval*, Amsterdam, 23–27 July 2007, pp. 821–822. New York: ACM.

[3] Moreira C and Wichert A. Finding academic experts on a multisensor approach using Shannon's entropy. *Expert Syst Appl* 2013; 40(14): 5740–5754.

[4] Fang Y, Si L and Mathur AP. Discriminative models of integrating document evidence and document-candidate associations for expert search. In: *SIGIR'10 Proceedings of the 33rd international ACM SIGIR conference on research and development in information retrieval*, Geneva, 19–23 July 2010, pp. 683–690. New York: ACM.

[5] Dai AM and Storkey AJ. The grouped author-topic model for unsupervised entity resolution. In: Honkela T, Duch W, Girolami M, et al. (eds) *Artificial neural networks and machine learning – ICANN 2011*. Berlin: Springer, 2011, pp. 241–249.

[6] Rosen-Zvi M, Chemudugunta C, Griffiths T et al. Learning author-topic models from text corpora. *ACM T Inform Syst* 2010; 28(1): 1–38.

[7] Kawamae N. Author interest topic model. In: *SIGIR'10 proceeding of the 33rd international ACM SIGIR conference on research and development in information retrieval*, Geneva, 19–23 July 2010, pp. 887–888. New York: ACM.

[8] Mou H, Geng Q, Jin J, et al. An author subject topic model for expert recommendation. In: Zuccon G, Geva S, Joho H, et al. (eds) *Information retrieval technology. AIRS 2015*. Cham: Springer, 2015, pp. 83–95.

[9] Dumais ST and Nielsen J. Automating the assignment of submitted manuscripts to reviewers. In: *SIGIR'92 proceedings of the 15th annual international ACM SIGIR conference on research and development in information retrieval*, Copenhagen, 21–24 June 1992, pp. 233–244. New York: ACM.

[10] Craswell N, de Vries AP and Soboroff I. Overview of the TREC 2005 enterprise track. In: *TREC'05 proceedings of the fourteenth text retrieval conference*, Gaithersburg, MD, 15–18 November 2005, pp. 1–7. NIST.

[11] Soboroff I, de Vries AP and Craswell N. Overview of the TREC 2006 enterprise track. In: *TREC'06 proceedings of the fifteenth text retrieval conference*, Gaithersburg, MD, 14–17 November 2006, pp. 1–20. NIST.

[12] Balog K, Azzopardi L and De Rijke M. Formal models for expert finding in enterprise corpora. In: *SIGIR'06 proceedings of the 29th annual international ACM SIGIR conference on research and development in information retrieval*, Seattle, WA, 6–11 August 2006, pp. 43–50. New York: ACM.

[13] Hettich S and Pazzani MJ. Mining for proposal reviewers: lessons learned at the National Science Foundation. In: *KDD'06 proceedings of the 12th ACM SIGKDD international conference on knowledge discovery and data mining*, 20–23 August 2006, Philadelphia, PA, pp. 862–871. New York: ACM.

[14] Han S, Jiang J, Yue Z, et al. Recommending program committee candidates for academic conferences. In: *Proceedings of the 2013 workshop on computational scientometrics: theory & applications*, San Francisco, CA, 28 October 2013, pp. 1–6. New York: ACM.

[15] Hofmann K, Balog K, Bogers T et al. Contextual factors for finding similar experts. *J Am Soc Inf Sci Tec* 2010; 61(5): 994–1014.

[16] Liu O, Wang J, Ma J et al. An intelligent decision support approach for reviewer assignment in R&D project selection. *Comput Ind* 2016; 76: 1–10.

[17] Tang W, Tang J, Lei T et al. On optimization of expertise matching with various constraints. *Neurocomputing* 2012; 76(1): 71–83.

[18] Xu Y, Ma J, Sun Y et al. A decision support approach for assigning reviewers to proposals. *Expert Syst Appl* 2010; 37(10): 6948–6956.

[19] Silva T, Jian M and Chen Y. Process analytics approach for R&D project selection. *ACM T Inform Syst* 2014; 5(4): 21.

[20] Datta A, Yong JTT and Ventresque A. T-RecS: team recommendation system through expertise and cohesiveness. In: *WWW'11 proceedings of the 20th international conference on World Wide Web*, Hyderabad, India, 28 March–1 April 2011, pp. 201–204. New York: ACM.

[21] Gollapalli SD, Mitra P and Giles CL. Ranking authors in digital libraries. In: *JCDL'11 proceedings of the 11th annual international ACM/IEEE joint conference on digital libraries*, Ottawa, ON, Canada, 13–17 June 2011, pp. 251–254. New York: ACM.
[22] Wang GA, Jiao J, Abrahams AS et al. ExpertRank: a topic-aware expert finding algorithm for online knowledge communities. *Decis Support Syst* 2013; 54(3): 1442–1451.
[23] Liu X, Suel T and Memon N. A robust model for paper reviewer assignment. In: *RecSys'14 proceedings of the 8th ACM conference on recommender systems*, Foster City, CA, 6–10 October 2014, pp. 25–32. New York: ACM.
[24] Yang C, Ma J, Silva T et al. A multilevel information mining approach for expert recommendation in online scientific communities. *Comput J* 2015; 58(9): 1921–1936.
[25] Li L, Wang Y, Liu G et al. Context-aware reviewer assignment for trust enhanced peer review. *PLoS ONE* 2015; 10(6): e0130493.
[26] Wang F, Zhou S and Shi N. Group-to-group reviewer assignment problem. *Comput Oper Res* 2013; 40(5): 1351–1362.
[27] Das GS and Gocken T. A fuzzy approach for the reviewer assignment problem. *Comput Ind Eng* 2014; 72: 50–57.
[28] Tayal DK, Saxena PC, Sharma A et al. New method for solving reviewer assignment problem using type-2 fuzzy sets and fuzzy functions. *Appl Intell* 2013; 40(1): 54–73.
[29] Li X and Watanabe T. Automatic paper-to-reviewer assignment, based on the matching degree of the reviewers. *Procedia Comput Sci* 2013; 22: 633–642.
[30] Sun J, Xu W, Ma J et al. Leverage RAF to find domain experts on research social network services: a big data analytics methodology with MapReduce framework. *Int J Prod Econ* 2015; 165: 185–193.
[31] Wei X and Croft WB. LDA-based document models for ad-hoc retrieval. In: *SIGIR'06 proceedings of the 29th annual international ACM SIGIR conference on research and development in information retrieval*, Seattle, WA, 6–11 August 2006, pp. 178–185. New York: ACM.
[32] Kou NM, Leong HU, Mamoulis N et al. Weighted coverage based reviewer assignment. In: *SIGMOD'15 proceedings of the 2015 ACM SIGMOD international conference on management of data*, Melbourne, VIC, Australia, 31 May–4 June 2015, pp. 2031–2046. New York: ACM.
[33] Kou NM, Leong Hou U, Mamoulis N et al. A topic-based reviewer assignment system. In: *Proceedings of the VLDB endowment – proceedings of the 41st international conference on very large data bases*, Kohala Coast, HI, August 2015, vol. 8, pp. 1852–1855. New York: ACM.
[34] Xu S, Shi Q, Qiao X, et al. Author-topic over time (AToT): a dynamic users interest model. In:Park J, Adeli H, Park N et al. (eds) *Mobile, ubiquitous, and intelligent computing* (Lecture Notes in Electrical Engineering), vol. 274. Berlin: Springer, 2014, pp. 239–245.
[35] Tang J, Zhang J, Yao L, et al. ArnetMiner: extraction and mining of academic social networks. In: *KDD'08 proceedings of the 14th ACM SIGKDD international conference on knowledge discovery and data mining*, Las Vegas, NV, 24–27 August 2008, pp. 990–998. New York: ACM.
[36] Tu Y, Johri N, Roth D, et al. Citation author topic model in expert search. In: *COLING'10 proceedings of the 23rd international conference on computational linguistics: posters*, Beijing, China, 23–27 August 2010, pp. 1265–1273. Stroudsburg, PA: Association for Computational Linguistics.
[37] Wang J, Hu X, Tu X, et al. Author-conference topic-connection model for academic network search. In: *CIKM'12 proceedings of the 21st ACM international conference on information and knowledge management*, Maui, HI, 29 October–2 November 2012, pp. 2197–2183. New York: ACM.
[38] Yang Z, Hong L and Davison BD. Academic network analysis: a joint topic modeling approach. In: *ASONAM'13 IEEE/ACM international conference on advances in social networks analysis and mining*, Niagara Falls, ON, Canada, 25–28 August 2013, pp. 324–333. New York: ACM.
[39] Mimno D and McCallum A. Expertise modeling for matching papers with reviewers. In: *KDD'07 proceedings of the 13th ACM SIGKDD international conference on knowledge discovery and data mining*, San Jose, CA, 12–15 August 2007, pp. 500–509. New York: ACM.
[40] Kawamae N. Latent interest-topic model: finding the causal relationships behind dyadic data. In: *CIKM'10 proceedings of the 19th ACM international conference on information and knowledge management*, Toronto, ON, Canada, 26–30 October 2010, pp. 649–658. New York: ACM.
[41] Ramage D, Hall D, Nallapati R, et al. Labeled LDA: a supervised topic model for credit attribution in multi-labeled corpora. In: *EMNLP'09 proceedings of the 2009 conference on empirical methods in natural language processing*, Singapore, 6–7 August 2009, pp. 248–256. Stroudsburg, PA: Association for Computational Linguistics.
[42] Teh YW, Jordan MI, Beal MJ et al. Hierarchical Dirichlet processes. *J Am Stat Assoc* 2006; 101(476): 1566–1581.
[43] Griffiths TL and Steyvers M. Finding scientific topics. *Proc Natl Acad Sci U S A* 2004; 101(suppl. 1): 5228–5235.
[44] Buckley C and Voorhees EM. Retrieval evaluation with incomplete information. In: *SIGIR'04 Proceedings of the 27th annual international ACM SIGIR conference on research and development in information retrieval*, Sheffield, 25–29 July 2004, pp. 25–32. New York: ACM.
[45] Karimzadehgan M, Zhai C and Belford G. Multi-aspect expertise matching for review assignment. In: *CIKM'08 proceedings of the 17th ACM conference on information and knowledge management*, San Francisco, CA, 26–30 October 2008, pp. 1113–1122. New York: ACM.