WEIGHTED MATRIX FACTORIZATION FOR SPOKEN DOCUMENT RETRIEVAL

Kuan-Yu Chen[†#], *Hsin-Min Wang*[†], *Berlin Chen*[*], *Hsin-Hsi Chen*[#]

[†]Institute of Information Science, Academia Sinica, Taipei, Taiwan
[#]National Taiwan University, Taipei, Taiwan
[*]National Taiwan Normal University, Taipei, Taiwan
E-mail: [†]{kychen, whm}@iis.sinica.edu.tw, [*]berlin@ntnu.edu.tw, [#]hhchen@ntu.edu.tw

ABSTRACT

Since more and more multimedia data associated with spoken documents have been made available to the public, spoken document retrieval (SDR) has become an important research subject in the past two decades. Recently, topic models have been successfully used in SDR as well as general information retrieval (IR). These models fall into two categories: probabilistic topic models (PTM) and non-probabilistic topic models (NPTM). One major difference between PTM and NPTM is that the former only takes the words occurring in a document into account, whereas the latter, such as latent semantic analysis (LSA), explicitly models all the words in the vocabulary (including both occurring and non-occurring words). We believe that the non-occurring words can provide additional information that is also useful for SDR. However, to our best knowledge, there is a dearth of work investigating the effectiveness of the non-occurring words for SDR and IR. In order to make effective use of those non-occurring words of documents for semantic analysis, we propose a weighted matrix factorization (WMF) framework, in which the impact of the non-occurring words on the semantic analysis can be modulated properly. The results of SDR experiments conducted on the TDT-2 (Topic Detection and Tracking) collection highlight the performance merits of our proposed framework when compared to several existing topic models.

Index Terms— Spoken document retrieval, topic model, non-probabilistic, non-occurring words

1. INTRODUCTION

Over the past two decades, spoken document retrieval (SDR) [1, 2] has become an interesting research subject in the speech processing community due to large volumes of multimedia associated with spoken documents being made available to the public. A significant amount of research effort has been devoted towards developing robust indexing (or representation) techniques [3-6] so as to extract probable spoken terms or phrases inherent in a spoken document that could match the query words or phrases literally. Instead, SDR research has revolved more around the notion of relevance of a spoken document in response to a query. It is generally agreed that a document is relevant to a query if it can address the stated information need of the query, but not because it happens to contain all the words in the query [7].

In the past, the vector space model (VSM) [7, 8], the Okapi BM25 model [7, 9], and the unigram language model (ULM) [10, 11] are well-representative ones for many information retrieval (IR) applications, including SDR. Their efficient and effective abilities have been proved by many researchers and practitioners for a wide variety of IR-related tasks. Yet, the later effort for further extending these methods to capture context dependence based on *n*-grams of various orders or some grammar structures mostly lead to mild gains or even spoiled results [10, 11]. The reasons are two-fold. On one hand, this is due to the fact that these methods might suffer from the problems of word usage diversity, which sometimes makes the retrieval performance degrade severely as a given query and its relevant documents use quite different sets of words (e.g. synonyms). On the other hand, lots of polysemy words have different meanings in different contexts. As such, merely matching words occurring in the original query and a document may not capture the semantic intent of the query.

To mitigate the above problems, topic models [6, 12-16] attempt to discover a set of latent topics, for which the relevance between a query and a document is not computed directly based on the co-occurrence frequencies of the query words and the document words. These models typically fall into two categories: 1) probabilistic topic models (PTM) [12-15] and 2) non-probabilistic topic models (NPTM) [6, 16]. Latent Dirichlet allocation (LDA) [13, 14] and its precursor, probabilistic latent semantic analysis (PLSA) [12], are two basic formulations of PTM. They both introduce a set of latent topic variables to describe the "*word-document*" co-occurrence characteristics. The relevance between a query and a document is computed based on the frequencies of the query words in the latent topics as well as the likelihood that the document generates the respective topics. On the other hand, latent semantic analysis (LSA) [6, 16] is a well-known representative of NPTM. LSA assumes that the latent topics are orthogonal and can be constructed by decomposing a pre-defined "*word-by-document*" matrix of a training document collection with singular value decomposition (SVD). The role of SVD in LSA is to derive a set of fundamental concepts that represent the document collection. Each document (and query) is subsequently characterized by a vector of weights indicating the strength with respect to each concept. The relevance degree between a query and a document can be estimated by the cosine similarity measure [7] between the query and the document representations (vectors).

One major difference between PLSA (and other variants of PTM) and LSA (and other variants of NPTM) is that the former only takes the words occurring in a document into account, but the latter explicitly models all the words in the vocabulary (including both words occurring and non-occurring in a document) [17, 18]. A document usually contains only a few distinct words, i.e., most words in the vocabulary do not occur in a document. Although LSA has the advantage of modeling both the occurring and non-occurring words in a given document, treating all the words with equal importance could be a serious disadvantage. In order not to overemphasize the non-occurring words in the original LSA model, we leverage a weighted matrix factorization (WMF) framework to properly modulate the impact of the occurring and non-occurring words on the semantic analysis. We also exploit multi-levels of index features, including word- and syllable-level units, in concert with the proposed WMF framework. The results of SDR experiments on the TDT-2 (Topic Detection and Tracking) collection demonstrate the superior performance of the instantiations of the WMF framework over several existing methods, including LSA. It is worth noting that the WMF framework can be applied to general IR tasks as well.

The remainder of this paper is organized as follows. We briefly review the mathematical formulations of the topic models for SDR in Section 2. In Section 3, we detail our proposed WMF framework. Then, the experimental settings and results are presented in Sections 4 and 5, respectively. Finally, Section 6 gives our conclusion and future work.

2. RELATED WORK

2.1. Probabilistic Topic Models (PTM)

Instead of matching a query and a document in a literal index term space, the relevance between a pair of query and document can be estimated on the grounds of a set of latent topics. For this idea to work, each document d is taken as a document topic model M_d, consisting of a set of K shared latent topics $\{T_1,\ldots,T_k,\ldots,T_K\}$ associated with the document-specific weights $P(T_k|\mathrm{M}_d)$, where each topic T_k in turn offers a unigram distribution $P(w_i|T_k)$ for observing an arbitrary word of the language [12, 14, 19, 20]. For example, in the PLSA model, the probability of a word w_i generated by a document d is expressed by:

$$P_{\mathrm{PLSA}}(w_i|\mathrm{M}_d)=\sum_{k=1}^{K}P(w_i|T_k)P(T_k|\mathrm{M}_d). \quad (1)$$

A document is believed to be more relevant to the query if the query words appear frequently in the topics on which the document has higher weights.

On the other hand, LDA, having a formula analogous to PLSA for document modeling, is thought of as a natural extension to PLSA, and has enjoyed much empirical success for various text IR tasks. LDA differs from PLSA mainly in the inference of model parameters: PLSA assumes that the model parameters are fixed and unknown; while LDA places additional a priori constraints on the model parameters, i.e., thinking of them as random variables that follow some Dirichlet distributions [19]. Since LDA has a more complex form for model optimization, which is hardly to be solved by exact inference, several approximate inference algorithms, such as the variational Bayes approximation [13, 14] and the Gibbs sampling algorithm [21], have been proposed to facilitate the estimation of the parameters of LDA according to different training strategies.

2.2. Non-Probabilistic Topic Models (NPTM)

Unlike probabilistic topic modeling, LSA [4, 6, 16] is an alternative way to describe the "*word-document*" co-occurrence characteristics. It assumes that there is an implicit semantic structure between words and documents, and the semantic structure can be explored by performing SVD on a pre-defined word-by-document matrix. When given N documents, which consist of M distinct words, we have an $M\times N$ matrix $\mathbf{A}$. Each element $\mathbf{A}_{ij}$ of $\mathbf{A}$ is the frequency of word (or term) w_i in document d_j. To eliminate some noisy words (e.g., function words), the inverse document frequency (IDF) can be used to weight the term frequency (TF) count, leading to the well-known TF-IDF [7]. Subsequently, SVD decomposes $\mathbf{A}$ into three sub-matrices:

$$\mathbf{A}_{M\times N}\approx\mathbf{U}_{M\times K}\,\Sigma_{K\times K}\,\mathbf{V}^{T}_{K\times N}=\tilde{\mathbf{A}}_{M\times N}, \quad (2)$$

where $K\le\min(M,N)$; $\mathbf{U}$ and $\mathbf{V}$ are orthonormal matrices, i.e., $\mathbf{U}^T\mathbf{U}=\mathbf{V}^T\mathbf{V}=\mathbf{I}$; and Σ is a diagonal matrix. Each word is uniquely associated with a row vector of matrix $\mathbf{U}$, and each document is uniquely associated with a column vector of matrix $\mathbf{V}^T$. In the retrieval phase, a query is viewed as a new document, and its K-dimensional vector representation is computed by a "*fold-in*" process as follows,

$$\tilde{\mathbf{q}}=\mathbf{q}^T\mathbf{U}\Sigma^{-1}. \quad (3)$$

The relevance degree between a pair of query and document is estimated by the cosine similarity measure between the query and the document representations (vectors).

2.3. Some Extensions of NPTM

LSA captures most of the important associations between words and documents, and removes the noise or variability in word usage, which often plague conventional IR models. Intuitively, since the number of dimensions is much smaller than the number of vocabulary words, the words occurring in similar documents will be represented in the nearby vicinity in the K-dimensional space even if they have never co-occurred in the same document. Consequently, LSA has been shown to achieve pretty good performance in IR, and much effort has been paid to improve LSA from different aspects.

Semantic context inference (SCI) [3] is a specially designed model for concept mapping and context expansion of spoken documents in SDR. The major difference between SCI and LSA is that SCI takes the word-word associations into account, while LSA considers the word-document co-occurrence relationships. SCI builds a semantic relation matrix to reflect the word-word associations, and performs SVD on the matrix to remove the noisy factors and capture the most important associations. In addition, a few LSA-based language models [4, 5, 22] attempt to construct a matrix to render the word-ordering information. Regularized latent

semantic indexing (RLSI) [23] formalizes topic modeling as a problem of minimizing a quadratic loss function that is regularized by different norms, and the problem can be decomposed into multiple sub-optimization problems that can be solved in parallel.

3. WEIGHTED MATRIX FACTORIZATION FOR SPOKEN DOCUMENT RETRIEVAL

3.1. Weighted Matrix Factorization (WMF)

As discussed above, LSA has the advantage of modeling both the words occurring and non-occurring in documents. We believe that the non-occurring words can provide additional information that is also useful for SDR. In some sense, the non-occurring words represent the non-relevant concept of a document [18]. However, we also notice that the number of non-occurring words is usually much larger than the number of occurring words in a document. Treating the occurring and non-occurring words with equal importance could be a serious disadvantage of LSA because the non-occurring words might dominate the estimation of model parameters. Therefore, we should not only take all words equally into account, but should also modulate the impact from occurring and non-occurring words properly.

For this idea to go, we propose a weighted matrix factorization (WMF) framework. When given N documents, which consist of M distinct words, we have an $M \times N$ word-by-document matrix $\mathbf{A}$. Our goal is to obtain two precious matrices $\mathbf{X}$ and $\mathbf{Y}$, where $\mathbf{X} \in R^{K \times M}$, $\mathbf{Y} \in R^{K \times N}$, and K is a desired rank. Note that the column vectors of $\mathbf{X}$ and $\mathbf{Y}$ are associated with each unique word and document, respectively. Basically, matrices $\mathbf{X}$ and $\mathbf{Y}$ can be solved by:

$$\min \sum_i \sum_j \left[\mathbf{A}_{ij} - \left(\mathbf{X}^T \mathbf{Y}\right)_{ij}\right]^2. \tag{4}$$

In order to modulate the impact of words occurring and non-occurring in documents, we introduce a weight matrix $\mathbf{W} \in R^{M \times N}$ to define the weight for each element in $\mathbf{A}$. In this paper, we adopt a simple yet effective way as follows,

$$\begin{cases} W_{ij} = 1 & \text{, if } \mathbf{A}_{ij} \neq 0 \\ W_{ij} = \delta & \text{, otherwise,} \end{cases} \tag{5}$$

where δ is a tunable parameter. Finally, matrices $\mathbf{X}$ and $\mathbf{Y}$ can be solved by minimizing the regularized weighted Frobenius distance as follows,

$$\min \sum_i \sum_j \mathbf{W}_{ij} \left[\mathbf{A}_{ij} - \left(\mathbf{X}^T \mathbf{Y}\right)_{ij}\right]^2 + \lambda_X \|\mathbf{X}\|_2^2 + \lambda_Y \|\mathbf{Y}\|_2^2, \tag{6}$$

where $\lambda_X \geq 0$ and $\lambda_Y \geq 0$ are the parameters controlling the regularization on $\mathbf{X}$ and $\mathbf{Y}$, respectively. The minimization problem in Eq. (6) corresponds to a low-rank approximation problem, which can be solved iteratively by [17, 23, 24]:

$$\mathbf{X}_{\cdot i} = \left(\mathbf{Y}\hat{\mathbf{W}}_{i\cdot}\mathbf{Y}^T + \lambda_X \mathbf{I}\right)^{-1} \mathbf{Y}\hat{\mathbf{W}}_{i\cdot}\mathbf{A}_{i\cdot}^T, \tag{7}$$

$$\mathbf{Y}_{\cdot j} = \left(\mathbf{X}\hat{\mathbf{W}}_{\cdot j}\mathbf{X}^T + \lambda_Y \mathbf{I}\right)^{-1} \mathbf{X}\hat{\mathbf{W}}_{\cdot j}\mathbf{A}_{\cdot j}, \tag{8}$$

where $\hat{\mathbf{W}}_{i\cdot} \in R^{N \times N}$ is a diagonal matrix with the weights in the i-th row of $\mathbf{W}$ on the diagonal, and $\hat{\mathbf{W}}_{\cdot j} \in R^{M \times M}$ is a diagonal matrix with the weights in the j-th column of $\mathbf{W}$ on the diagonal.

During the retrieval process, a query can be viewed as a new document. We can first fix $\mathbf{X}$, and then fold-in the new document to calculate the K-dimensional vector representation $\mathbf{Y}_\mathbf{q} \in R^{K \times 1}$. Finally, the relevance degree between the query and a document d is estimated by the cosine measure between $\mathbf{Y}_q$ and $\mathbf{Y}_d$. To take both literal and concept information into consideration, we can further augment the conventional frequency count (or TF-IDF) vector ($\mathbf{q} \in R^{M \times 1}$ and $\mathbf{d} \in R^{M \times 1}$) with the new concept representative vector ($\mathbf{Y}_q \in R^{K \times 1}$ and $\mathbf{Y}_d \in R^{K \times 1}$) to construct a new vector for the query and the document as follows,

$$\left[\mathbf{q}^T, \gamma \mathbf{Y}_q^T\right]^T, \tag{9}$$

$$\left[\mathbf{d}^T, \gamma \mathbf{Y}_d^T\right]^T, \tag{10}$$

where a trade-off parameter γ is used to balance the contribution of the original frequency count vector and the new concept representative vector.

3.2. Using Subword-level Index Units

In this paper, we also integrate subword-level information into topic modeling for SDR. To do this, syllable pairs are taken as the basic units for indexing in addition to words. The recognition transcript of each spoken document, in form of a word stream, was automatically converted into a stream of overlapping syllable pairs. Then, all the distinct syllable pairs occurring in the spoken document collection were then identified to form a vocabulary of syllable pairs for indexing. We can simply use syllable pairs, in replace of words, to represent the spoken documents, and construct the associated topic models accordingly.

4. EXPERIMENTAL SETUP

We used the Topic Detection and Tracking collection (TDT-2) [25] in the experiments. The Mandarin news stories from Voice of America news broadcasts were used as the spoken documents. All news stories were exhaustively tagged with event-based topic labels, which served as the relevance judgments for performance evaluation. The average word error rate obtained for the spoken documents is about 35%. The Chinese news stories from Xinhua News Agency were used as our test queries. More specifically, in the following experiments, we will either use a whole news story as a "long query," or merely extract the tittle field from a news story as a "short query." Table 1 shows some basic statistics of the TDT-2 collection.

4.1. Evaluation Metric

The retrieval performance is evaluated in terms of non-interpolated mean average precision (MAP) following the TREC evaluation [26], which is computed by:

$$\text{MAP} = \frac{1}{|\mathbf{Q}|} \sum_{i=1}^{|\mathbf{Q}|} \frac{1}{N_i} \sum_{j=1}^{N_i} \frac{j}{r_{i,j}}, \tag{11}$$

Table 1. Statistics of the TDT-2 collection.

	TDT-2 (1998, 02~06)			
# Spoken documents	2,265 stories, 46.03 hours of audio			
# Distinct test queries	16 Xinhua text stories (Topics 20001~20096)			
	Min.	Max.	Med.	Mean
Doc. length (in characters)	23	4,841	153	287.1
Length of test short query (in characters)	8	27	13	14
Length of test long query (in characters)	183	2,623	329	532.9
# Relevant documents per test query	2	95	13	29.3

Table 2. Retrieval results (in MAP) with word-level index features for short and long queries.

	VSM	**LSA**	**SCI**	**WMF**	**Hybrid**
short	0.273	0.379	0.270	0.438	0.448
long	0.484	0.512	0.413	0.561	0.561

Table 3. Retrieval results (in MAP) with subword-level index features for short and long queries.

	VSM	**LSA**	**SCI**	**WMF**	**Hybrid**
short	0.257	0.330	0.270	0.328	0.338
long	0.499	0.466	0.349	0.475	0.513

where $|\mathbf{Q}|$ is the number of test queries, N_i is the total number of documents that are relevant to the *i*-th test query, and $r_{i,j}$ is the position (rank) of the *j*-th document that is relevant to the *i*-th query, counting down from the top of the ranked list.

5. EXPERIMENTAL RESULTS

First, our proposed WMF model is compared with two instances of non-probabilistic topic models, namely LSA and SCI. The results when using different types of queries (i.e., long or short queries) and different kinds of index features (i.e., word- or subword-level index features) are shown in Tables 2 and 3. The number of latent topics of all the topic models is set to 128, the regularization parameters (λ_X and λ_Y) are set to 1, and the frequency count of words is weighted by using the standard IDF method. From the results, at first glance, it seems that the high word error rate (WER) for the spoken document collection (about 35%) does not lead to catastrophic failures probably due to the reason that recognition errors are overshadowed by a large number of spoken words correctly recognized in the documents. The experimental results seem to reveal that the proposed WMF framework outperforms both LSA and SCI in most cases. It is worth noting that, when using word-level indexing features for SDR, WMF yields significant improvements over LSA and SCI. We have also found that WMF achieved better performance when δ in Eq. (5) was set around 0.08. The results confirm our idea that the words non-occurring in a document should not be considered as important as the occurring ones, although the former might provide additional information.

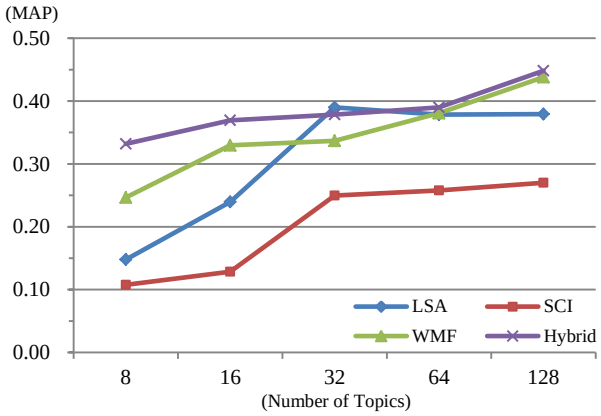

Figure 1. Retrieval results (in MAP) for short queries with word-level index features with respect to the number of latent topics.

From Tables 2 and 3, it can also be observed that the conventional term matching strategy (denoted by VSM) can also achieve a certain level of performance. Although merely matching terms in the original query and document may not always capture the semantic intent of a query, term matching still provides an important clue for retrieval, which is complementary to the concept matching. To use both literal and concept information, we concatenate the VSM and the WMF features to construct a new index vector for both queries and documents (*cf.* Eqs. (9) and (10) in Section 3.1), and the retrieval results are also shown in Tables 2 and 3 (denoted by "Hybrid"). As expected, the hybrid method outperforms VSM and WMF in all cases. In fact, it performs better than all the other models in all cases.

Finally, Figure 1 reports the retrieval results for short queries with respect to the number of latent topics when using word-level index features. It is known that the way to systemically determine the optimal number of latent topics for topic models is still an open issue and needs further investigation. As can be seen from Figure 1, the performance of most non-probabilistic topic models is apt to be improved as the topic number increases, except that LSA seems to saturate when the topic number is larger than 32. Due to space limitations, we only report on the results for one setting, but similar tendencies are observed for other settings (e.g., different types of queries and indexing mechanisms).

6. CONCLUSION & FUTURE WORK

This paper has proposed a weighted matrix factorization framework for spoken document retrieval, which suggests a promising way to improve the latent semantic analysis model by directly modulating the impact of words occurring and non-occurring in documents on the document (or query) representations. The utility of the proposed framework has been validated by extensive comparisons with several existing information retrieval models. Our future work includes the development of supervised training, incorporation of some prior knowledge, and extension to probabilistic topic models.

7. REFERENCES

[1] C. Chelba, T. J. Hazen, and Murat Saraclar, "Retrieval and browsing of spoken content," *IEEE Signal Processing Magazine*, 25(3), pp. 39-49, 2008.

[2] L. S. Lee and B. Chen, "Spoken document understanding and organization," *IEEE Signal Processing Magazine*, 22(5), pp. 42-60, 2005.

[3] C. L. Huang, B. Ma, H. Li, and C. H. Wu, "Speech indexing using semantic context inference," in *Proc. INTERSPEECH*, pp.717-720, 2011.

[4] W. Naptali, M. Tsuchiya, and S. Nakagawa, "Context dependent class language model based on word co-occurrence matrix in LSA framework for speech recognition," in *Proc. ACS*, pp.275-280, 2008.

[5] W. Naptali, M. Tsuchiya, and S. Nakagawa, "Word co-occurrence matrix and context dependent class in LSA based language model for speech recognition", *International Journal of Computers*, pp.85-95, 2009.

[6] S. Deerwester, S. Dumais, G. Furnas, T. Landauer, and R. Harshman, "Indexing by latent semantic analysis", *Journal of the American Society of Information Science*, 41(6), pp. 391-407, 1990.

[7] C. D. Manning, P. Raghavan and H. Schtze, *Introduction to Information Retrieval*, New York: Cambridge University Press, 2008.

[8] G. Salton , A. Wong , and C. S. Yang, "A vector space model for automatic indexing," *Communications of the ACM*, 18(11), pp.613-620, Nov. 1975

[9] K. S. Jones, S. Walker, and S. E. Robertson. "A probabilistic model of information retrieval: development and comparative experiments (parts 1 and 2)," *Information Processing and Management*, 36(6), pp. 779-840, 2000.

[10] J. M. Ponte and W. B. Croft, "A language modeling approach to information retrieval," in *Proc. SIGIR*, pp. 275-281, 1998.

[11] W. B. Croft and J. Lafferty (eds.), "Language modeling for information retrieval," Kluwer International Series on Information Retrieval, Volume 13, Kluwer Academic Publishers, 2003.

[12] T. Hoffmann, "Unsupervised learning by probabilistic latent semantic analysis," *Machine Learning*, 42, pp. 177-196, 2001.

[13] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *Journal of Machine Learning Research*, pp. 993-1022, 2003.

[14] D. M. Blei and J. Lafferty, "Topic models," in A. Srivastava and M. Sahami, (eds.), *Text Mining: Theory and Applications*, Taylor and Francis, 2009.

[15] K. Y. Chen, H. M. Wang, and B. Chen, "Spoken document retrieval leveraging unsupervised and supervised topic modeling techniques," *IEICE Transactions on Information and Systems*, pp. 1195-1205, 2012.

[16] M. Berry, S. Dumais, and G. O'Brien, "Using linear algebra for intelligent information retrieval," *SIAM Review*, pp.573-595, 1995.

[17] W. Guo and M. Diab, "Modeling sentences in the latent space," in *Proc. ACL*, pp.864-872, 2012.

[18] Y. Lv, and C. X. Zhai, "Query likelihood with negative query generation", in *Proc. CIKM*, pp.1799-1803, 2012.

[19] K. Y. Chen, H. S. Chiu, and B. Chen, "Latent topic modeling of word vicinity information for speech recognition," in *Proc. ICASSP*, pp.5394-5397, 2010.

[20] X. Wei and W. B. Croft, "LDA-based document models for ad-hoc retrieval," in *Proc. SIGIR*, pp. 178-185, 2006.

[21] T. L. Griffiths and M. Steyvers, "Finding scientific topics," *Proc. of the National Academy of Sciences*, pp. 5228-5235, 2004.

[22] J. T. Chien, M. S. Wu and H. J. Peng, "Latent semantic language modeling and smoothing," *International Journal of Computational Linguistics and Chinese Language Processing*, 9(2), pp. 29-44, 2004.

[23] Q. Wang, J. Xu, H. Li, and N. Craswell, "Regularized latent semantic indexing," in *Proc. SIGIR*, pp.685-694, 2011.

[24] N. Srebro and T. Jaakkola, "Weighted low-rank approximations," in *Proc. ICML*, pp.720-727, 2003.

[25] LDC, "Project topic detection and tracking," *Linguistic Data Consortium*, 2000.

[26] J. Garofolo, G. Auzanne, and E. Voorhees, "The TREC spoken document retrieval track: A success story," in *Proc. TREC*, pp. 107-129, 2000.