Opinion Retrieval

Kam-Fai Wong

Department of Systems Engineering & Engineering Management
The Chinese University of Hong Kong

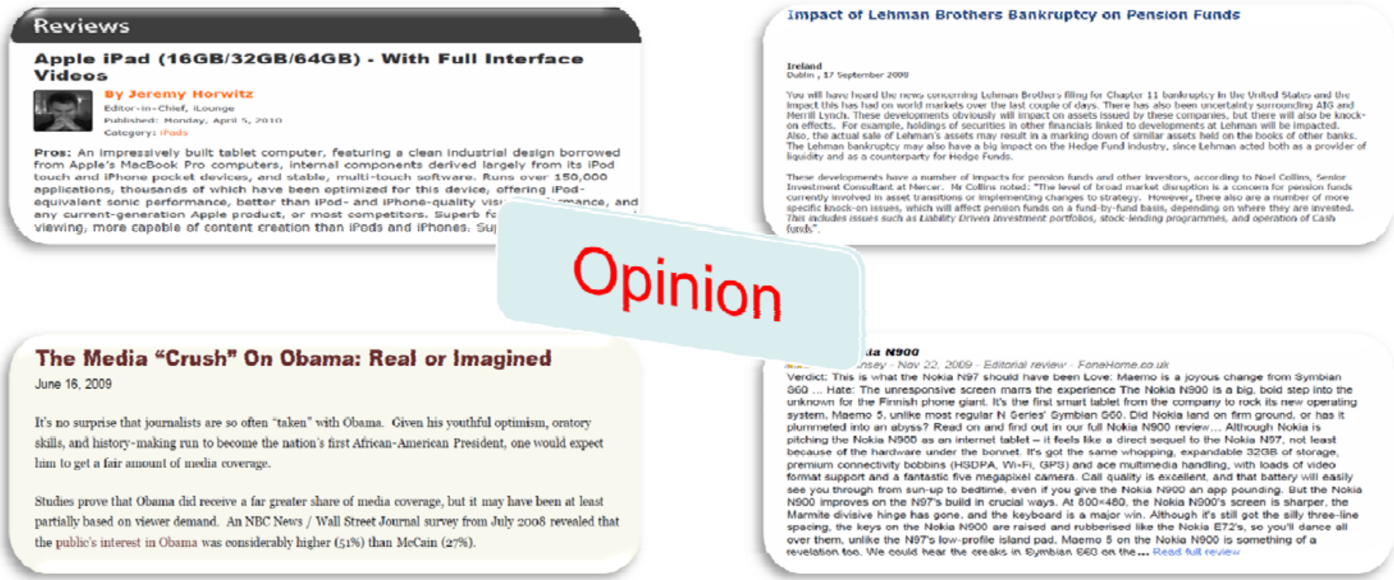
iDB-2010, Japan, August 2-4 2010

Outline

- Introduction
- Motivation
- Our Model
- Evaluation
- Related Work
- Conclusion & Future Work

Introduction

- There is a growing interest in sharing personal opinions on the Web, such as product reviews, economic analysis, political polls, etc.



- Opinion-oriented applications: opinion mining, sentiment classification, opinion summarization, opinion question & answering.

Introduction

- Opinion retrieval was first presented in the TREC 2006 Blog track [Macdonald and Ounis. 2006]. Chinese opinion retrieval was presented in COAE (Chinese Opinion Analysis Evaluation) [Zhao et al., 2008].

- Objective of opinion retrieval:
  - retrieve documents that express an opinion about a given target.

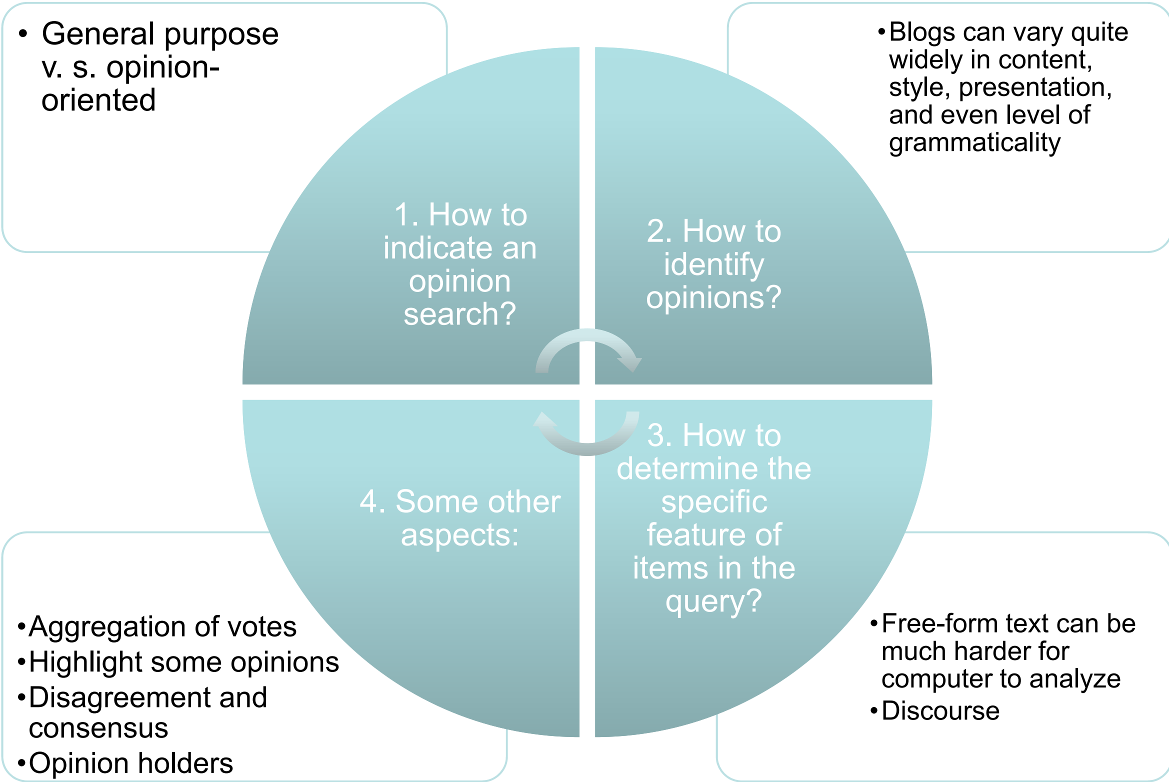
- The topic of the document is not required to be the same as the target, but an opinion about the target has to be presented in the document.

Introduction

- Comparison between information retrieval and opinion retrieval
  - Information need
    - Fact vs. opinion
  - Measurement
    - Similarity vs. ?
  - Granularity
    - Document vs. sentence
  - Top-*k*
    - Documents on the first page vs. top-*k* documents

Introduction

Pang Bo [Pang 2008] suggested that A complete opinion retrieval application might involve attacking each of the following problems.



Introduction

- We realize our objective for opinion retrieval in the following three phases:

1. Re-rank the documents retrieved by general purpose search engine (2, 3)
2. Integrate opinion features into current retrieval processing (4)
3. Develop an opinion-oriented search engine, i.e. opinion searcher (1)

Outline

- Introduction
- Motivation
- Our Model
- Evaluation
- Related Work
- Conclusion & Future Work

Motivation

- A 2-stage approach was proposed in TREC
  - Detect the relevance of the document, $\text{Score}_{\text{rel}}$
  - Identify the opinion of the document, Score_{op}
- An example of opinion retrieval, e.g. Q='Avatar'

A. 阿凡达明日将在中国上映。
Tomorrow, Avatar will be shown in China.
B. 我预订到了IMAX影院中最舒服的位子。
I've reserved a comfortable seat in IMAX.
C. 阿凡达是我最喜欢的一部3D电影。
Avatar is my favorite 3D movie.

- The overall score for ranking is computed as

$$Score_{\mathrm{doc}} = Score_{\mathrm{op}} + Score_{\mathrm{rel}}$$

where $Score_{\mathrm{rel}} = tf_{\mathrm{Q}} \times idf_{\mathrm{Q}}$, $Score_{\mathrm{op}} = weight_{\mathrm{comfortable}} + weight_{\mathrm{favorite}}$

Motivation

- Limitations:
  - Relevance of the document ≠ relevance of the opinion
  - Degree of the sentiment word ≠ importance of the opinion
- Our Method:
  - We proposed to handle opinion retrieval in the granularity of sentence.
  - Word pair was proposed to maintain both intra-sentence and inter-sentence contextual information.
  - Contextual information is integrated into our graph-based opinion retrieval model.

Outline

- Introduction
- Motivation
- Our Model
- Evaluation
- Related Work
- Conclusion & Future Work

Formal Definition

- Given a document set $D=\{d_1,d_2,d_3,\ldots,d_n\}$, and a specific query $Q=\{q_1,q_2,q_3,\ldots,q_z\}$, where $q_1,q_2,q_3,\ldots,q_z$ are query keywords. Opinion retrieval aims at retrieving documents from D with relevant opinion about the query *Q.*

- In addition, we construct a sentiment word lexicon V_o and a topic term lexicon V_t .

- *Definition*: *topic-sentiment word pair* p_{ij} *consists* of two elements, one is from V_t, and the other one is from V_o.

$$p_{ij} = \{< t_i, o_j > | t_i \in V_t , o_j \in V_o)\}$$

Intra-sentence Information

- Intra-sentence contextual information
  - The association between an opinion and its corresponding target can be expressed in a word pair.
  - Practically, a word pair represents a relevant opinion.

- There may be more than one opinion in one sentence. We split each sentence into a set of word pairs:

$$s_l \rightarrow \{< t_i, o_j > | t_i = \min Dist(t_i, o_j) \text{ for each } o_j\}$$

- The more relevant opinions the sentence includes, the higher weight it carries.

Inter-sentence Information

- Inter-sentence contextual information
  - The relationship among the opinions on the same topic
  - The contribution of a word pair is determined by the inter-sentence information.

- We assume that the more sentences contain the same opinion, the more contribution the opinion makes to those sentences, and hence the document.

HITS Model

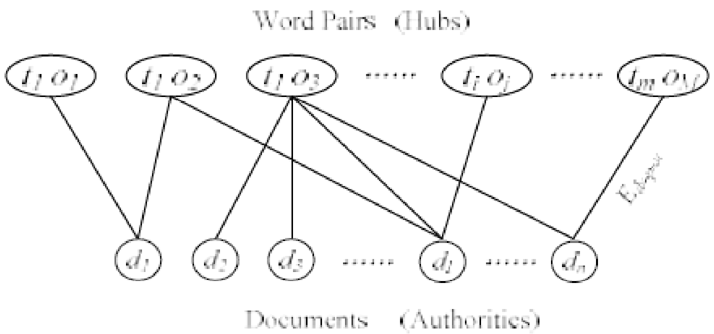
- Graph-based ranking algorithms, such as HITS or PageRank, have been traditionally and successfully used in citation analysis, social networks [Wan et al., 2008; Li et al., 2009; Erkan and Radev, 2004, Li et al., 2009].

- Graph-based ranking algorithm is a way of deciding on the importance of a vertex within a graph, by taking into account global information recursively computed from the entire graph, rather than relying only on local information.

- Because contributions vary a lot from word pair to word pair, we apply HITS model to opinion retrieval.

HITS Model

- Our proposed opinion retrieval model is based on HITS model and it contains two layers.
  - The word pairs layer is considered as hubs and the documents layer authorities.



A word pair that has links to many documents denotes a strong associative degree between the two items, i.e. <t_i,o_j>.

A document that has links to many word pairs is with many *relevant opinions*, and it will result in high ranking.

HITS Model

- We compute the contribution by the weight of the edge connecting between the word pairs and the documents.

$$w_{ij}^{k} = \frac{1}{|d_k|} \sum_{p_{ij} \in s_l \in d_k} \left[\lambda \cdot rel(t_i, s_l) + (1-\lambda) opn\left(o_j, s_l\right)\right]$$

λ is introduced as the trade-off parameter to balance the $rel(t_i, s_l)$ and $opn(o_j, s_l)$;

$rel(t_i, s_l)$ is computed to judge the relevance of t_i in s_l which belongs to d_k;

$$rel(t_i, s_l) = tf_{t_i, s_l} \times isf_{t_i}$$

$opn(o_j, s_l)$ is the contribution of o_j in s_l which belongs to d_k.

$$opn(o_j, s_l) = \frac{tf_{o_j, s_l}}{tf_{o_j, s_l} + 0.5 + (1.5 \times \frac{len(s_l)}{asl})}$$

HITS Model

- All word pairs are initialized equally. In each iteration T+1, the scores of Hubs and Authorities are updated according to the scores in iteration T.

- The convergence of the iteration is achieved when the difference between the scores computed at two successive iterations falls below a given threshold.

- The documents are ranked by the Authorities scores.

Outline

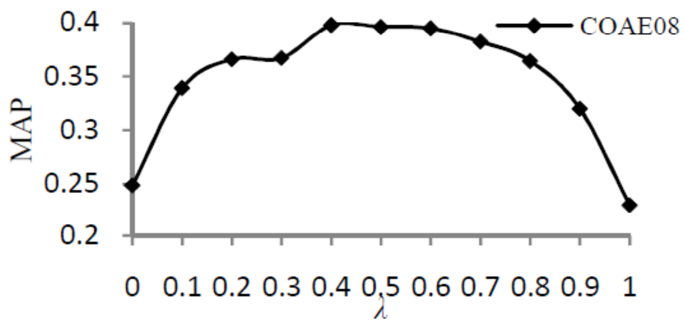
- Introduction
- Motivation
- Our Model
- Evaluation
- Related Work
- Conclusion & Future Work

Experiment Setting

- Dataset:
  - COAE08 dataset, which consists of 40000 blogs and reviews. 20 queries are provided in COAE08.
- Sentiment Lexicon:
  - *The Lexicon of Chinese Positive Words*
  - *Lexicon of Chinese Negative Words*
  - The opinion word lexicon provided by National Taiwan University
  - Sentiment word lexicon and comment word lexicon from *Hownet*
- Topic Term Collection:
  - The dictionary-based method
  - The web-based method
- Baseline Approach:
  - ROCC [Zhang and Yu, 2007]

Experimental Parameter & Metrics

- Experimental parameter tuning (λ in Equation 1)



  - λ =0.4

- Experimental Metrics
  - MAP: Mean Average Precision
  - Rpre: R-precision
  - bPref: binary Preference
  - P@10: Precision at 10 documents

Experimental Result 1

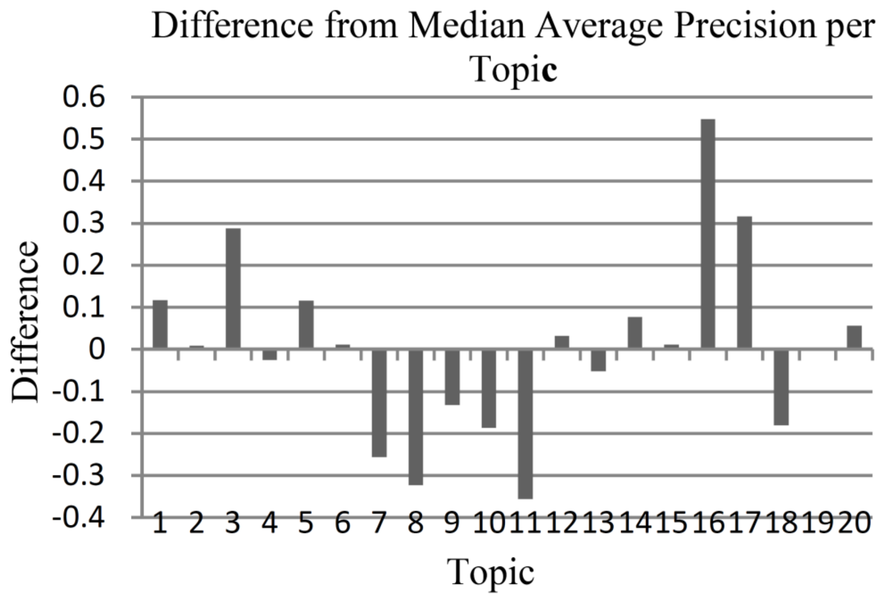
- Comparison of different approaches on COAE08 dataset, and the best is highlighted

Approach	COAE08			
	Evaluation metrics			
Run id	MAP	R-pre	bPref	P@10
IR	0.2797	0.3545	0.2474	0.4868
Doc	0.3316	0.3690	0.3030	0.6696
ROSC	0.3762	0.4321	0.4162	0.7089
Baseline	0.3774	0.4411	0.4198	0.6931
GORM	**0.3978**	**0.4835**	**0.4265**	**0.7309**

  - IR: A classical information retrieval model
  - Doc: The 2-stage document-based opinion retrieval model
  - ROSC: This was the model which achieved the best run in TREC Blog 07
  - GORM: our proposed graph-based opinion retrieval model

Experimental Result 2

- Difference from Median on COAE08 dataset



  – The Median Precision is 0.3724.

Experimental Result 3

- Top-5 highest weight word pairs for 5 queries in COAE08 dataset

Top-5 MAP				
陈凯歌 Chen Kaige	国六条 Six States	宏观调控 Macro-regulation	周星驰 Stephen Chow	Vista Vista
<陈凯歌 支持> Chen Kaige Support	<房价 上涨> Room rate Rise	<经济 平稳> Economics Steady	<电影 喜欢> Movie Like	<价格 贵> Price Expensive
<陈凯歌 最佳> Chen Kaige Best	<调控 加强> Regulate Strengthen	<价格 上涨> Price Rise	<周星驰 喜欢> Stephen Chow Like	<微软 喜欢> Microsoft Like
<《无极》 骂> Limitless Revile	<中央 加强> CCP Strengthen	<发展 平稳> Development Steady	<主角 最佳> Protagonist Best	<Vista 推荐> Vista Recommend
<影片 优秀> Movie Excellent	<房价 平稳> Room rate Steady	<消费 上涨> Consume Rise	<喜剧 好> Comedy Good	<问题 重要> Problem Vital
<阵容 强大的> Cast Strong	<住房 保障> Housing Security	<社会 保障> Social Security	<作品 精彩> Works Splendid	<性能 不> Performance No

Discussion

- Result 1 showed that GORM outperformed the other approaches in all metrics.
  – About 20% improvement of MAP was achieved by sentence-based approach.

- Result 2 showed that GORM performed well in most of the queries. Except for:
  – Topic 11, i.e. '指环王' (Lord of the King): there were only 8 relevant documents without any opinion and 14 documents with relevant opinions.

  – Topic 8, i.e. '成龙' (Jackie Chan) & topic 7, i.e. '李连杰' (Jet Lee): there were a number of similar relevant targets for the two topics.

- Result 3 showed that high-weighted word pairs could represent the relevant opinions about the corresponding queries.

Outline

- Introduction
- Motivation
- Our Model
- Evaluation
- Related Work
- Conclusion & Future Work

Related Work

- Sentiment lexicon-based approaches
  - Hannah et al proposed a lightweight lexicon-based statistical approach [Hannah et al., 2007].
  - Amati et al generated a weighted dictionary from previous TREC relevance data [Amati et al., 2007].
  - Na et al. created a pseudo opinionated word composed of all opinion words, which was shown to be very effective in TREC 2008 [Na et al., 2009].
  - Huang and Croft proposed an effective relevance model by considering both query-independent and query-dependent sentiment [Huang and Croft, 2009].

Related Work

- Unified models for opinion retrieval:
  - Eguchi and Lavrenko proposed an opinion retrieval model in the framework of generative language modeling [Eguchi and Lavrenko, 2006].
  - Mei et al. tried to build a fine-grained opinion retrieval system for consumer products [Mei et al., 2007].
  - Zhang and Ye proposed a generative model to unify topic relevance and opinion generation [Zhang and Ye, 2008].
  - Huang and Croft proposed a unified opinion retrieval model according to the K-L divergence between the two probability distributions of opinion relevance model and document mode [Huang and Croft, 2009].

Outline

- Introduction
- Motivation
- Our Model
- Evaluation
- Related Work
- Conclusion & Future Work

Conclusion

- The information need for opinion retrieval has been proposed.
- Both intra-sentence and inter-sentence contextual information are well represent by word pairs.
- A sentence-based opinion retrieval approach is unified through the graph-based model which performs well on COAE08 dataset.

Future work

- It is worth further study on how they could be applied to other opinion oriented applications, e.g. opinion summarization, opinion prediction, etc.
- The characteristics of blogs will be taken into consideration, i.e., the post time, which could be helpful to create a more time sensitivity graph to filter out fake opinions.
- Opinion holder is another important role of an opinion, and the identification of opinion holder is a main task in NTCIR. It would be interesting to study opinion holders, e.g. its seniority, for opinion retrieval.

References

- Giambattista Amati, Edgardo Ambrosi, Marco Bianchi, Carlo Gaibisso, and Giorgio Gambosi. 2007. FUB, IASI-CNR and University of Tor Vergata at TREC 2007 Blog Track. In *Proceedings of the 15th Text Retrieval Conference*.
- Koji Eguchi and Victor Lavrenko. Sentiment retrieval using generative models. 2006. In *EMNLP '06, Proceedings of 2006 Conference on Empirical Methods in Natural Language Processing*, page 345-354.
- Gunes Erkan and Dragomir R. Radev. 2004. Lexpagerank: Prestige in multi-document text summarization. In *EMNLP '04, Proceedings of 2004 Conference on Empirical Methods in Natural Language Processing*.
- David Hannah, Craig Macdonald, Jie Peng, Ben He, and Iadh Ounis. 2007. University of Glasgow at TREC 2007: Experiments in Blog and Enterprise Tracks with Terrier. In *Proceedings of the 15th Text Retrieval Conference*.
- Xuanjing Huang, William Bruce Croft. 2009. A Unified Relevance Model for Opinion Retrieval. In *Proceedings of CIKM*.
- Fangtao Li, Yang Tang, Minlie Huang, and Xiaoyan Zhu. 2009. Answering Opinion Questions with Random Walks on Graphs. In *ACL '09, Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics.*
- Bing Liu, Minqing Hu, and Junsheng Cheng. 2005. Opinion observer: Analyzing and comparing opinion s on the web. In *WWW '05: Proceedings of the 14th International Conference on World Wide Web*.
- Craig Macdonald and Iadh Ounis. 2006. Overview of the TREC-2006 Blog Track. In *Proceedings of the 14th Text Retrieval Conference*.
- Qiaozhu Mei, Xu Ling, Matthew Wondra, Hang Su, and Chengxiang Zhai. 2007. Topic sentiment mixture: Modeling facets and opinions in weblogs. In *WWW '07: Proceedings of the 16 International Conference on World Wide Web*.

References

- Seung-Hoon Na, Yeha Lee, Sang-Hyob Nam, and Jong-Hyeok Lee. 2009. Improving opinion retrieval based on query-specific sentiment lexicon. In *ECIR '09: Proceedings of the 31st annual European Conference on Information Retrieval*, pages 734-738.
- Bo Pang and Lillian Lee. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2): 1-135.
- Xiaojun Wan and Jianwu Yang. 2008. Multi-document summarization using cluster-based link analysis. In *SIGIR '08: Proceedings of the 31th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 299-306. ACM.
- Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. 2005. Recognizing contextual polarity in phrase-level sentiment analysis. In *EMNLP '05, Proceedings of 2005 Conference on Empirical Methods in Natural Language Processing*.
- Min Zhang and Xingyao Ye. 2008. A generation model to unify topic relevance and lexicon-based sentiment for opinion retrieval. In *SIGIR '08: Proceedings of the 31st Annual International ACM SIGIR conference on Research and Development in Information Retrieval*, pages 411-418. ACM.
- Wei Zhang and Clement Yu. 2007. UIC at TREC 2007 Blog Track. In *Proceedings of the 15th Text Retrieval Conference*.
- Jun Zhao, Hongbo Xu, Xuanjing Huang, Songbo Tan, Kang Liu, and Qi Zhang. 2008. Overview of Chinese Opinion Analysis Evaluation 2008. In *Proceedings of the First Chinese Opinion Analysis Evaluation*.

Thank you!