Modeling Sentiment and Aspect Using Syntax: A Topic Model Approach

Rui Xie^a, Meng Xia^b, Chunping Li^{a*}, Raymond Y.K. Lau^c

^aSchool of Software, Tsinghua University, Beijing, China

^bDepartment of Computer Science, Beihang University, Beijing, China

^cDepartment of Information Systems, City University of Hong Kong, Hong Kong

*Chunping Li: cli@tsinghua.edu.cn

Abstract

In this paper, based on Latent Dirichlet Allocation (LDA), we propose a novel probabilistic modeling framework, which aims to reveal the latent aspects and sentiments of reviews simultaneously. Unlike other topic models which only consider the words appearing in online reviews, we consider Part-of-Speech (POS) tags in our model. Since users may use different types of words to express different meanings, we have proposed two Tag Sentiment Aspect models (TSA) to integrate syntactical information into the review mining models. We have applied the proposed models to two datasets, electronic product reviews and movie reviews, and evaluated the results in terms of sentiment aspect extraction and sentiment polarity classification. Our study shows that the proposed models not only achieve promising results on sentiment classification, but also effectively extract different latent sentiment aspects. Moreover, the proposed TSA models are fully unsupervised, and they do not need any manually labeled reviews for training. To incorporate priors, only the lists of positive and negative words are required. Moreover, the proposed TSA models are effective across different domain.

Keywords: Topic Modeling, Sentiment Analysis, Aspect Analysis, User Generated Content.

1. Introduction

Nowadays the development of Web 2.0 makes convenient for customers to express their experience with products. A large amount of product reviews can be easily obtained from website such as Amazon and Epinion. Online reviews are of a useful resource to help customers to make

purchase decisions. By browsing these reviews, potential consumers could be aware of the good or the bad aspect of a specific product. However, facing an overwhelming amount of reviews of products, no one can read the reviews piece by piece. This triggers to explore an approach to obtain useful and hidden information from such a large review corpus and automatically analyze the reviews. The approach tackles two problems: one is *topic identification*, which identifies the aspects of the product the reviews talk about; the other one is *sentiment analysis*, determine the sentiment label (positive, negative) of opinion toward specific aspect. Recently, these problems have drawn great interests.

To analyze reviews, most of previous works use a two-phase way approach: firstly extract the aspects reviews refer to and secondly analyze the opinion expressed in the reviews toward the aspects extracted in the first step. However, in this paper, we exploit models which can tackle the two problems simultaneously. We tackle the two problems in a probabilistic modeling framework based on Latent Dirichlet Allocation (LDA)⁽¹⁾. We choose LDA because it has shown strong power in revealing latent topic in the corpus. We incorporate sentiment to the original LDA by adding a sentiment layer so that our models can reveal the language model hidden in the review corpus associated with aspect and sentiment pair.

In this paper, we propose two probabilistic models: Sentiment Aspect Tag (SAT) model and Tag Sentiment Aspect (TSA) model. Both models can model aspect and sentiment at once. Both models model the generative process of reviews. The only difference lies on how the models incorporate syntactic information. SAT treats POS tags of words as observed variables and models the generative process under different sentiments. The POS tags

are the generating variables in SAT, as well as the words in reviews. For different sentiments, there is a corresponding distribution over POS tags. SAT capture the influence of sentiment label on generating process of different POS tags. On the contrary, TSA capture the opposite one—the influence of different POS tags on generating process of sentiment proportion. In TSA, for each distinct POS tag, there is an associated multinomial distribution over sentiment.

To our best knowledge, not much work can deal with aspect extraction and sentiment analysis simultaneously except^(2,5,6). However, our model has several difference and improvements compared to the existing works: (1) Our models are the first to incorporate the syntax information into the hierarchical probabilistic model. (2) Our models are fully unsupervised while some of the existing works need labeled data to train. (3) Our model is domain independent but by integrating different domain prior information, our model can apply to different domain.

We apply our model to various tasks: document-level sentiment classification, aspects discovery and aspect-specific sentiment words extraction. The experiments show that our models perform pretty well. By applying models on different datasets—electronic product reviews and movie reviews, it shows that our models can easily adjust when domain changes.

2. Related Works

There are two major directions to discover the hidden aspect and sentiment in the reviews⁽¹⁷⁾. One direction is to apply traditional natural language processing techniques to do text mining of reviews. Approaches of this direction can be viewed as “micro to macro” approaches. These approaches usually analyze individual words, associations between words, sentences structure and linguistic patterns and then form macroscopic topics and sentiments of reviews by sum up individual components. There are usually two steps: first, identify aspects of products and second analyze opinion orientation to corresponding aspects. For aspect identifying, Hu and Liu⁽⁷⁾ use POS tagging to find noun phrase and select frequent nouns as aspect candidates and then filter candidates to obtain actual aspects. In⁽³⁾, frequently co-occurred nouns with seed aspect words are expanded using association rule mining. For sentiment analysis, firstly the opinion words around aspect are extracted and then sophisticated schema to aggregate opinion is applied to the reviews. Often a lexicon is required

to decide the word’s orientation or a lexicon is built by expanding from a seed list. As in⁽⁹⁾, the opinion aggregation function is proposed to aggregate the sentiment of individual opinion words. In another work, Turney⁽¹⁰⁾ computes the polarity of a given opinion word by measuring its Pointwise Mutual Information (PMI) to keywords “excellent” and “poor” and expanding those words which have a higher PMI as seeds.

Another direction to discover the hidden aspect and sentiment in the reviews is to apply probabilistic statistic approach to model the whole corpus. It can seem as a “macro to micro” way because it models the whole generating process of review corpus, and infers the distributions of different sentiments and aspects. Among probabilistic models, LDA is a powerful one in capturing latent topic hidden in corpus. LDA treats document as a mixture of multiply language models, each of which presents a hidden topic/aspect. Lots of works extend LDA model to improve aspects extraction, considering different context in reviews. Brody and Elhadad⁽¹¹⁾ proposed a sophisticated LDA to discover the aspect hidden in the reviews. A connectivity matrix is used to calculate a score, which decides the best hidden aspect number and iteration times. Zhang et al⁽¹²⁾ distinguished general aspect and find-grained aspects. The model can capture ratable aspects and global aspect which makes the result more meaningful.

Usually, topic model are used to capture aspects only, but jointly modeling aspect and sentiment have been tried in some works^(12, 5, 4, 3). We introduce these works and compare them with our models.

Multi-Aspect Sentiment (MAS) model⁽⁶⁾ by Titov and McDonald is an extension of their previously work—MG-LDA, which only extracts topics hidden in the reviews regardless of sentiments. MAS model works in a supervised way because it requires every aspects to be rated by user. In contrast, our models do not need any supervised setting. It only needs some small prior information. Moreover, MAS lacks of the flexibility to adapt to other domains.

Topic Sentiment Mixture (TSM) model⁽⁵⁾ is a model extended from pLSA⁽¹³⁾ by Mei and Zhai. TSM has the well-known deficits of pLSA that it has the problem with inference of new documents and suffers from overfitting. Our models have no such problem because it is based on LDA. In the other hand, TSM does not consider the association of topic and sentiment. The words are drawn from either topic distribution or sentiment distribution. The words are samples of a mixture of sentiment and topic, not a

combination. This makes TSM lack the ability to exact the aspect-specific opinion words. In our models, we sample a word on distribution conditioned on aspect index and sentiment label.

Joint Sentiment/Topic (JST) model⁽⁴⁾ is a fully unsupervised model based on LDA. It can capture topic and sentiment at the same time. Our models are highly related to JST, but we incorporate the POS tag information to obtain more accurate result.

Aspect Sentiment Unification model (ASUM)⁽³⁾ is a model with small adaption of the JST. It introduces the assumption that one sentence in review can only refer to one aspect and one sentiment. This is not true in real situation. Many review sentences have several clauses and usually talks about different aspects and contrast sentiments.

3. Model

The two models we propose extend the basic LDA to model the generative process of review dataset. In LDA, words of reviews are treated as samples that sampled from topics, which is represented as a multinomial distribution over distinct words. We call it topic-word distribution. A review is represented as a multinomial distribution over topics. We call it document-topic distribution. Every word is sampled from a particular topic-word distribution, and the topic is sampled from document-topic distribution. So there is a latent variable for every word to indicate the topic from which it is sampled. In SAT and TSA models, a sentiment layer is added between document and topic. Words are sampled by topic-sentiment-word distribution and there are two corresponding latent variable for each words. Every review is presented by several document-topic distributions with different sentiments. We call it document-sentiment-topic distribution. The difference between SAT and TSA is the way they incorporate syntax information. Both SAT and TSA utilize POS tag information to incorporate syntax, for POS tag can obtain easily with help of POS tagging technique. The difference way to incorporate syntax and original reason are detailed in this section. To infer the latent distributions, Gibbs sampling⁽¹⁴⁾ is used.

3.1 Sentiment Aspect Tag Model (SAT)

In SAT, POS tag of words is regarded as generated data as good as words in the review. SAT integrates POS tag generating process into the model to incorporate syntax information. The behind assumption is that distribution over

tags varies with sentiment, so for a particular sentiment, different tags may be generated. The graphical presentation of SAT is shown in Figure 1(a). As shown, tag t is generated conditioned on aspect index z and sentiment label l , along with the word w . The tag is considered as the stamp of the word. This is inspired by the work by Wang and McCallum⁽¹⁵⁾, in which the published time of the document is treated as the timestamp of words in the document.

Figure 1. Graphical representation of SAT and TSA.

The generative process of TSA1 is as follows:

1. For each aspect and sentiment pair (z, l) , draw a word distribution $\psi_{z,l} \sim \text{Dir}(\beta_l)$, and draw a tag distribution $\omega_{z,l} \sim \text{Dir}(\mu_l)$.
2. For each review d ,
 - a. Draw a sentiment distribution $\pi_d \sim \text{Dir}(\gamma)$.
 - b. For each sentiment label l , draw an aspect distribution $\theta_{d,l} \sim \text{Dir}(\alpha)$.
 - c. For each word w_i and tag t_i in the review,
 - Choose a sentiment label $j \sim \text{Mul}(\pi_d)$
 - Choose an aspect $k \sim \text{Mul}(\theta_{d,j})$
 - Choose word $w_i \sim \text{Mul}(\psi_{k,j})$ and tag $t_i \sim \text{Mul}(\omega_{k,j})$.

The hyperparameters α , β , γ and μ are parameters that control the prior distribution θ , ψ , π and ω and these hyperparameters are regard as prior information. Notice that for different sentiment label, prior information should be different, so we use asymmetric β_l and μ_l corresponding to sentiment label to exploit such prior information. The notation z and l indicate the index of topic and sentiment respectively. Notation S , A and D denotes the number of sentiment, aspect and review, respectively. Our target is to

calculate posterior distribution of $P(\theta|\mathbf{w})$, $P(\psi|\mathbf{w})$, $P(\pi|\mathbf{w})$ and $P(\omega|\mathbf{w})$ given observation of word vector $\mathbf{w}$. By Gibbs sampling approach, the full conditional probabilities $P(z_i, l_i | \mathbf{z}_{-i}, \mathbf{l}_{-i}, \mathbf{w}, \mathbf{t}, \alpha, \beta, \gamma, \mu)$ should be calculated, where z_i denotes the aspect assignment for w_i , l_i denotes the sentiment assignment for w_i , $\mathbf{z}_{-i}$ denotes the aspect assignments for all words except w_i , $\mathbf{l}_{-i}$ denotes the sentiment assignments for all words except w_i . During Gibbs sampling, we draw aspect and sentiment iteratively for each word w_i according to the following probability distribution:

$$P(z_i, l_i | \mathbf{w}, \mathbf{t}, \beta, \mu, \mathbf{l}_{-i}, \alpha, \gamma,) \propto \frac{N_{k,d}^{-i} + \gamma}{N_d^{-i} + S\gamma} \quad (1)$$

$$\frac{N_{j,k,d}^{-i} + \alpha}{N_{k,d}^{-i} + A\alpha} \frac{N_{w_i,j,k}^{-i} + \beta_{k,w_i}}{\sum_{v=1}^V (N_{v,j,k}^{-i} + \beta_{k,v})} \frac{N_{t_i,j,k}^{-i} + \gamma_{k,t_i}}{\sum_{t=1}^T (N_{t,j,k}^{-i} + \gamma_{k,t})}$$

where $N_{k,d}$ is the number of words assigned to sentiment label k in review d , N_d is the number of words in review d , $N_{j,k,d}$ is the number words assigned to aspect j and sentiment k in review d . $N_{w_i,j,k}$ is the number the word w_i assigned to aspect j and sentiment k , and $N_{t_i,j,k}$ the number the tag t_i assigned to aspect j and sentiment k . The superscript $-i$ denotes the number that excludes the i^{th} position.

Having the conditional probability, the approximate posterior probability of θ , π , ψ and ω is able to estimate as follows:

$$\theta_{j,k,d} = \frac{N_{j,k,d} + \alpha}{N_{k,d} + A\alpha} \quad (2)$$

$$\pi_{k,d} = \frac{N_{k,d} + \gamma}{N_d + S\gamma} \quad (3)$$

$$\psi_{j,k,w} = \frac{N_{w,j,k} + \beta_{k,w}}{\sum_{v=1}^V (N_{v,j,k} + \beta_{k,v})} \quad (4)$$

$$\omega_{j,k,t} = \frac{N_{t,j,k} + \mu_{k,t}}{\sum_{t=1}^T (N_{t,j,k} + \mu_{k,t})} \quad (5)$$

3.2 Tag Sentiment Aspect Model (TSA)

TSA model differs from SAT in interpreting relations between POS tags and sentiments. TSA, shown in Figure 1(b), can be regarded as an inverse version of SAT that inverts the dependency between POS tags and sentiments, which is obviously observed from graphical representation in Figure 1. In SAT, sentiment variable l points to tag variable t . That means tags are conditioned by sentiments. However, in TSA, the relation is opposite. It is hard to say which interpretation is more suitable, it is alike the “chicken

and egg” puzzle. However, the interpretation in TSA is more natural to understand, because in language, syntax e.g. grammar regulate the expression which implies that syntax dominate the content one expresses. Another advantage is that it gets a simpler model, shown in Figure 1, with one less hidden distribution to infer, compared to SAT.

The generative process of TSA is as follows:

1. For every aspect and sentiment pair (z, l) , draw a discrete distribution over words $\psi_{z,l} \sim \text{Dir}(\beta_l)$
2. For each review d ,
 - a. For each type of tag t , draw a sentiment distribution $\pi_{d,t} \sim \text{Dir}(\gamma_t)$.
 - b. For each sentiment label l , draw an aspect distribution $\theta_{d,l} \sim \text{Dir}(\alpha)$.
 - c. For each word w_i with tag t_i in the review,
 - Choose a sentiment label $j \sim \text{Mul}(\pi_{d,t_i})$
 - Choose an aspect $k \sim \text{Mul}(\theta_{d,l})$
 - Choose word $w_i \sim \text{Mul}(\psi_{k,j})$.

Noticed that in TSA, syntax information is incorporated in sentiment distributions π for every review, for different π correspond to different types of tag. The hyperparameter γ is set asymmetric to encode syntactic information. Other parameters and variables are the same as SAT, except there is no hyperparameter μ and distribution ω . Like SAT, we the Gibbs sampling to infer hidden distribution according to full conditional probability is as following:

$$P(z_i, l_i | \mathbf{w}, \mathbf{t}, \beta, \mu, \mathbf{l}_{-i}, \alpha, \gamma,) \propto \frac{N_{k,t,d}^{-i} + \gamma_{k,t}}{\sum_{l=1}^S (N_{l,t,d}^{-i} + \gamma_{l,t})} \quad (6)$$

$$\frac{N_{j,k,d}^{-i} + \alpha}{N_{k,d}^{-i} + A\alpha} \frac{N_{w_i,j,k}^{-i} + \beta_{k,w_i}}{\sum_{v=1}^V (N_{v,j,k}^{-i} + \beta_{k,v})}$$

where $N_{k,t,d}$ is the number of words assign to sentiment k with a tag t in review d . Other notations are the same with SAT. The approximate probability distribution of θ and ψ can be calculated using a same form in Equation (2) and (4), and distribution of π is obtained as following:

$$\pi_{k,t,d} = \frac{N_{k,t,d} + \gamma_{k,t}}{\sum_{l=1}^S (N_{l,t,d} + \gamma_{l,t})} \quad (7)$$

4. Experiment and Analysis

We use the dataset of electronic device reviews used in ⁽³⁾ and movie reviews used in ⁽¹⁶⁾ to verify our models. The two review corpus are from different domain to show the flexibility of our unsupervised model when domain adjust. We analyze the power of SAT and TSA in latent aspect and

sentiment discovery and then evaluate power of sentiment classification of the two models. We also give a comparison between our models and the existing models, such as JST⁽⁴⁾, and ASUM⁽³⁾.

The electronic device dataset contains about 25,000 reviews, each rated 1 to 5, from 7 product categories such as MP3, laptop, etc. The movie dataset contains 1,000 positive and 1000 negative reviews. We use NLTK¹ for the preprocessing. First, the sentences are segmented. Then, we tokenize every sentence and perform a POS tagging. After that, we remove the punctuation, numbers, and non-alphabet tokens, apply WordNet's² stemming and retain the words with 'NN', 'JJ', 'VB' and 'RB' tags. Table 1 shows summary information of datasets after preprocessing.

TABLE 1. SUMMARY OF DATASETS

	<i>Electronic</i>	<i>Movie</i>
Number of reviews	24260	2000
Number of sentences	292338	65720
Number of tokens	2424985	756261
Vocabulary size	68451	42962

Before modeling, we set the parameters. The sentiment number is set to be 3 for positive, negative and objective. The hyperparameters can be regarded as priors representing bias with prior knowledge. For aspect there should be no bias, α is set symmetric. After trying several value and find the value does not real affect. We use symmetric α to be 0.1. For SAT, γ is set symmetric to be 1 because there should not be bias between reviews. However, in TSA, γ is associated with tag, which is bias on sentiment considering different tags. For tags "NN", the objective element of parameter γ_t is set to 2, twice as the sentimental elements; for tags "VB", "RB" and "JJ", the sentimental elements of parameter γ_t is set to 2, twice as the objective elements. In SAT, this prior information is encoded in parameter μ . For objective sentiment label, element corresponding to "NN" of μ_1 is set twice as elements corresponding to "RB", "VB" and "JJ" with value 1; for positive and negative sentiment label, elements corresponding to "RB", "VB" and "JJ" of μ_1 is set twice as element corresponding to "NN" with value 1. The β is also set asymmetric according to different sentiment in both SAT and TSA. We use a positive word list and a negative word list as prior information for word distribution under different sentiment. For positive aspects, we set the

elements of β to be 0 for words in negative list and 0.001 for all the other words. Similarly, for negative aspects, we set the elements of β to be 0 for the positive seed words and 0.001 for all the other words. For objective aspects, β is set symmetric to be 0.001. For aspect number, we try several different values and find the suitable for electronic device reviews is 50 and 25 is suitable for movie reviews. The iteration time is set to be 1000 for Gibbs sampling. After modeling, we examine top 10 words for every aspect with different sentiments to analyze the ability of our model in aspect and sentiment words extraction. Then we use the hidden distribution π to classify the review into positive and negative to analyze the ability of our model in sentiment classification.

Before we verify our models, we first emphasize the ability models should achieve. Basic ability is to extract hidden aspect. The criterion is that the aspect should be coherent but specific. This is the same criterion when evaluating clustering. In our models, we combining consider aspect and sentiment, so the additional criterion for our models is that the models should distinguish aspect words and opinion words. The final criterion is that our model can model opinion words' polarity correctly. As this criterion is difficult to quantified, we performance an additional task—sentiment classification to quantify the abilities of our models.

We analyze word distribution ψ for different aspect and sentiment. Each sentimental aspect is represented by top 5 frequent words. Due to space limitation, we only show 3 aspects with different sentiment for electronic device and give detailed analysis to illustrate our models' capacity.

TABLE 2. SENTIMENT CLASSIFICATION OF REVIEWS

	Electronic Device Reviews			Movie Reviews		
Model	<i>Pos</i>	<i>neg</i>	<i>overall</i>	<i>pos</i>	<i>neg</i>	<i>overall</i>
SAT	82.8%	80.4%	83.1%	82.6%	81.4%	82%
TSA	87.8%	84.4%	86.2%	86.8%	84.4%	85.6%
JST	74.4%	68.6%	71.5%	70.8%	76.6%	74.2%
ASUM	80.6%	74.4%	79.8%	76.4%	72.8%	74.6%

We perform review sentiment classification on two dataset of different domain for our models, JST and ASUM to do a comparison. We use distribution π for sentiment classification on review level. Distribution $\pi_{k,d}$ presents the probability of sentiment k in review d , we compare the probability of sentiment positive and negative, and assign larger label to the review. Special case is in TSA, where $\pi_{k,t,d}$

¹ <http://www.nltk.org>

² <http://wordnet.princeton.edu/>

has another dimension of tag t . We simply sum up over t to get unified distribution. For electronic dataset, we discard reviews rated 3-4 and treat reviews rated 5 as positive ones and reviews rated 1-2 as negative one.

From Table 2, comparison between our own models shown that TSA perform better than SAT. This indicates that the assumption behind TSA model is more suitable for reality. We also compare our models with JST⁽⁴⁾ and ASUM⁽³⁾. Both JST and ASUM are unable to beat our models. The result claims that incorporate syntax information indeed improve sentiment extraction. An interesting observation is that accuracy of the negative reviews is lower than positive ones. One possible reason can be that the negation of sentence is not captured by current model. Negation is complicated when considering context and our model use “bag of words” assumption are not good to capture that.

5. Conclusions

The “Bag of words” assumption is more suitable for traditional text classification but it when comes to opinion mining, it may not a good one since opinion is expressed in a more complex way. This triggers the need to explore more information and patterns hidden opinion text. In this paper, we have explored the possibility of integrating POS information and sentiment aspects in the opinion modeling by proposing SAT and TSA. The experimental results reveal the merits of our approach.

Two directions can be further pursued in our future works. One is to exploit other language information, for example, the dependency relations among terms. A synonym thesaurus can also be used to explore the relations between sentiment and aspect words. The other direction is to seek a better presentation for prior information and adjust the TSA model to fit the corpus. In our models, the prior probabilities of different tags are fixed. However, it would be interesting to explore a semi-supervised learning approach to suggest these probabilities, with which only a small amount of labeled data are required.

Acknowledgment

This work is supported in part by China NSFC under Grant 61672309 and China MOST Fundamental Research Project under Grant 2017FY201407.

References

- (1) D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3:993-1022, 2003.
- (2) Raymond Y. K. Lau, Chunping Li, Stephen S. Y. Liao: Social analytics: Learning fuzzy product ontologies for aspect-oriented sentiment analysis. *Decision Support Systems* 65: 80-94 (2014).
- (3) Y. Jo and A. Oh. Aspect and sentiment unication model for online review analysis. *Proceedings of the 4th ACM international conference on Web search and data mining*, February 09-12, 2011, Hong Kong, China.
- (4) C. Lin and Y. He. Joint sentiment/topic model for sentiment analysis. In *Proceeding of the 18th ACM conference on Information and knowledge management*, pages 375 - 384, New York, NY, USA, 2009. ACM.
- (5) Q. Mei, X. Ling, M. Wondra, H. Su, and C. Zhai. Topic sentiment mixture: modeling facets and opinions in weblogs. In *Proceedings of the 16th international conference on World Wide Web*, pages 171-180, New York, NY, USA, 2007. ACM
- (6) I. Titov and R. Mcdonald. A joint model of text and aspect ratings for sentiment summarization. In *Proc. ACL-08: HLT*, pages 308-316, 2008.
- (7) M. Hu and B. Liu. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168 - 177, New York, NY, USA, 2004. ACM.
- (8) A. M. Popescu and O. Etzioni. Extracting product features and opinions from reviews. In *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, pages 339 - 346, Morristown, NJ, USA, 2005.
- (9) X. Ding , B. Liu , P. S. and Yu, A holistic lexicon-based approach to opinion mining, *Proceedings of the international conference on Web search and web data mining*, February 11-12, 2008, Palo Alto, California, USA
- (10) P. D. Turney. Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews. In *ACL '02: Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, pages 417-424, Morristown, NJ, USA, 2001. Association for Computational Linguistics.
- (11) S. Brody and N. Elhadad. An unsupervised aspect-sentiment model for online reviews. In *Proceedings of the 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pages 804 - 812, Los Angeles, California, June 2010. Association for Computational Linguistics.
- (12) Wenping Zhang, Raymond Y. K. Lau, Chunping Li: Adaptive Big Data Analytics for Deceptive Review Detection in Online Social Media. *ICIS 2014*
- (13) T. Hofmann. Probabilistic latent semantic analysis. In *Proceedings of the 15th Conference on Uncertainty in AI* (1999), San Fransisco, California: Morgan Kaufmann.
- (14) T. Griffiths and M. Steyvers. Finding scientific topics. *Proceedings of the National Academy of Sciences*, 101(90001):5228-5235, 2004.
- (15) X. Wang , A. McCallum, Topics over time: a non-Markov continuous-time model of topical trends, *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2006, Philadelphia, PA, USA
- (16) B. Pang, L. Lee, and S. Vaithyanathan. Thumbs up?: sentiment classification using machine learning techniques. In *EMNLP '02: Proceedings of the ACL-02 conference on Empirical methods in natural language processing*, pages 79 - 86, Morristown, NJ, USA, 2002. Association for Computational Linguistics.
- (17) Yi Cai, Raymond Y. K. Lau, Stephen S. Y. Liao, Chunping Li, Ho-fung Leung, Louis C. K. Ma: Object typicality for effective Web of Things recommendations. *Decision Support Systems* 63: 52-63 (2014).