Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Study on Feature Selection in Finance Text Categorization

Changqiu Sun, Xiaolong Wang, Jun Xu
Department of Computer Science and Technology
Shenzhen Graduate School,Harbin Institute of Technology
Shenzhen, China
sunchangqiu@gmail.com xlwang@insun.hit.edu.cn hit.xujun@gmail.com

***Abstract*—Document genre information is one of the most distinguishing features in information retrieval, which brings order to the search results. What the genre classification concerned is not the topic but the genre of document. In this paper, two different feature sets were employed: bag of words which are derived by feature selection method and structural features which are selected manually and subjectively. And a comparative study on feature selection in genre classification of Chinese finance text is presented. In empirical results with classifiers on the real world corpora, we find that that manual labeled features can improve the performance clearly.**

***Index Terms*—Text Categorization, Feature Selection, Genre Classification**

I. Introduction

With the developing of the WWW, the available information is becoming abundant than before. However, even with the help of search engine, it's still hard to find the most suitable information due to huge amount of abundant information. In order to help user to acquire relevant and useful information, classification and cluster tools are imported to classify search engine results. All these techniques succeed in topic classifying are topic-centered, which neglect the importance of genre in results ranking, and takes little account of the individual user's needs and preferences.

Our research focus on professional financial search. In finance domain, different genres of information have different authority which is important. Positive or negative news will influence the price of a stock. Depending on circumstances, users are always interested in certain or particular genres of information. For example, the financial analysts are interested in objective news and announcement. But most of common investors pay more attention to the different opinions in financial reviews. So, a further classification of different genres will be very useful for investment decision making.

An important issue of identifying the genres of Chinese finance text, is to find features that discriminate among different genres. Text genre can be indicated by surface features co-occurrence in different genres. The presence or absence of a feature in a document gives a clue to its genre. Since feature set has a decisive influence on genre classification, the focus in this paper is the evaluation of feature extraction techniques and comparison of feature selection methods. We seek answers to the following:

1). To investigate the way of feature extraction using human knowledge and the basis of previous studies. Is that reliable?

2). To investigate the possibility of feature selection methods. Which one can get a good compromise on performance and size of feature set?

The rest of the paper is structured as follows. Section 2 introduces genre classification and sketches out existing work, and describes the genre classes of finance text. In section 3 we analyze our feature space, and in section 4 we describes the feature selection methods. Section 5 introduces the corpus chosen for empirical validation, the experiments results, and the discussion of the results. Section 6 summarizes our research and future work.

II. Genre Classification

A. Genre Classification

There is no agreement on the definition of genre. The term ”genre” is used frequently in connection with music, with literature and arts. Literally, genre is a category of artistic composition, as in music or literature, marked by a distinctive style, form, or content [1]. Outwardly, genre can be viewed through the macro-features, so-called genre facets [2]: amount of graphics, several pictures, or many links, etc. Inside, it can be viewed through the facets: subjectivity, wording and phrasing, etc.

Genre classification is often regarded as orthogonal to the classification based on the document's contents. It groups a set of documents into same sets according to the predefined genre classes. Here we give a problem statement about the genre classification: Let $C = \{c_1, c_2, ...c_m\}$ be a set of categories (genre classes) and $D = \{d_1, d_2, ...d_n\}$ a set of documents. The task of the genre classification consists in assigning class label C_i to each document d_j, if the document d_j belong to C_i, which exactly one category must be assigned to each d_j. It is single-label text categorization.

B. Genre Classes

Previous research on genre classification defined genre classes correspond to individual task. In this paper, We describe the genre of finance text through the subjectivity, the style of writing, the form, the targeted audience and the functional trait. Finance text can be divided into three main types:

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

TABLE I
THE DIFFERENCE BETWEEN ANNOUNCEMENT, NEWS REPORTS AND OPINION&REVIEW

Category	Announcement	News	Opinion
Subjectivity	objective	low subjectivity	subjective
Author	company	reporter	analyst
Authority	authoritative	low authority	no authority
Style	official	formal	informal

1). **Announcement**, which is an official or authoritarian declaration, such as exchange notice, circular in respect of very substantial acquisition, or clarification of announcement from board meeting.

2). **News report**, which present factual information objectively. News reports was published by newspapers or distributed electronically on financial news sites, and written by reporter. News reports are supposed to be objective in the reporting of the latest stock prices and various events that are likely to influence the stock price of a particular company. Pure objectivity in factual news reporting could be too ideal in reality. Most of news reports have certain subjectivity. It is not only the objective description to the affairs, but also joins author's personal judgment and the standpoint.

3). **Review&opinion** that describe authors' views, estimation and judgments about stock or a company, give the implications of the events of the day for future stock prices and advise on investment. Reviews appear in columns such as ”Opinion”, ”Forum”, ”Blog”, are written by a single judge like analyst or blogger. Since it is a personal viewpoint or belief, opinions vary widely on a given subject by different people. Of course, opinions are subjective.

C. Related Work

There is a lot of work on genre identification. Most research are three-step approach: $i)$ Definition of particular genre classes of respective domain. $ii)$ Selection of features. $iii)$ quantification of the classification methods. The methods and features of automatic genre identification are relatively independent tasks, so in the following we give a brief of the existing work in respect of methods and features.

Traditional text classification techniques perform well on genre classification. For corpus-specific genre classification, they have either used term frequency analysis(sometimes called bag-of-word, BOW) [3] [4] or a linguistic approach involves POS tagging [5] [6]. In previous research [7], the linguistic approaches do achieve better accuracy than term frequency approaches. Finn et al. [5] investigate different features: Bag-of-Words, Part-of-Speech statistics, and text statistics for building genre classifiers and their ability to transfer across multiple topic domains. All three techniques proved to be effective in single domain tests in English. Working across domains and across languages is a more challenging and complex task. Peng et al. [8] present a simple method based on character-level n-gram language models to classify genre. The method view documents merely as a sequence of characters, without requiring feature selection or extensive pre-processing which makes it easy to implement. Eissen et al. [9] gives an overview of features for genre classification.

Genre classification of web pages is a popular topic now. The presentation-related features are commonly used for this kind of task [9] [10]. This type of features relate to the appearance of a document, such as hyperlinks, amount of graphics and tables, HTML tags, URL specifications, etc. Lim et al. [11] used URL depth, presence of a filename at the end of the URL, document type (HTML, ASP, PHP etc.), top-level domain and genre-specific words in the URL (faq, news, board, detail etc.) as features. Eissen et al. [9] show that with a small set of features, which captures linguistic and presentation related aspects, text statistics, and word frequency classes, acceptable classification results can be achieved. Their analysis reveals that about 70% of the documents are assigned correctly.

III. STRUCTURAL FEATURES

Each document hit by search engine have three parts: page title, URL and content. We define text info between HTML tag <title> and </title> as the page title. The page title is different from content title for it may contains additional column or source info. Such as “ <title>社保基金去年经营收益率高达43%$_{(1)}$_要闻公告$_{(2)}$ _新浪财经$_{(3)}$_新浪网$_{(4)}$(China's National Pension Fund Reports 43 pct Return in Last Year$_{(1)}$_News&Announcement$_{(2)}$_Sina Finance$_{(3)}$_Sina$_{(4)}$)</title>”. Part 1 is content title, part 2 gives genre column info, part 3 gives topic info and part 4 gives source info. More attention of our work is given to the feature selection of each part. We combine domain knowledge and empirical data to select distinguishing structural features. Our approach fall into three types: context features, phrases and text patterns.

Context features are the features pertaining to the text after cleaning but have no relevant to the content. Page title and URL sometimes are useful to indicate genre. For example, “新闻(news)” is used in page title of news or announcement text, while “点评(comment)” is only used in title of some opinion text. Appearance of some commonly used words like “news”, “blog” etc in URL also can help the classifier to identify the genre. Just grasp such context features is not enough to identify genres for the following reasons: (1)Most page titles do not include genre column info. (2)Some genre column info has ambiguity with content.

Phrases are surface, and are represented by lexical items. For example, the appearance of word ”记者(reporter)” and phrase ”接受采访(accept visiting)” gives hint that the document belongs to news text, while the phrase ”本公司(our company)” is commonly used in announcement text. It is also useful to detect the presence of interrogative or exclamatory sentence. All these linguistic features can be extracted without complicated parsing.

Text patterns are the features which repeat in a predictable manner. These patterns include sentence, clause and phrase patterns that give hint about the genre of a text. For example, in announcement, a title often follows the “<公司全称>+ 关于...公告(<FULL COMPANY NAME>+announcement about...)” pattern. In opinion text, it is popular to give a conclusion in“<维持/给予...>...<某级别>+评级 (<keep/give...>...<RANK SCORE>)” pattern，or express their opinion by sentence pattern like “<第一人称代词><ADV(OPTIONAL)><认为/确信..>...(<FIRST PERSON PRONOUN><ADV(OPTIONAL)><think/be sure..> ...)”. These text patterns are typically described by hand-crafted rules like regular expression and grasped by a parser.

We designed a feature extractor to extract the above features. All of the features are represented by the rules like regular expression. The extractor generate a feature if and only if the rules of this feature is matched. About 100 above features were obtained and employed in our system. In our research, the classifiers also employ bag-of-words.

IV. Feature selection methods

Eight methods are included in our study. Each of them uses a threshold *T* to achieve a desired degree of term elimination from the full vocabulary in a document corpus. Namely, eliminate the features whose result values are smaller than *T*. They are document frequency (DF), information gains (IG), mutual information (MI), a χ^2-test (CHI), term strength (TS), Expected cross entropy, Weight of evidence for text, Odds ratio. All of them have advantages and disadvantages, and some advantages outweigh their drawbacks, but some ones do not. We would like to explain them concretely as follows:

A. Document Frequency (DF)

The document frequency for a word is the number of documents in which the word occurs. It computes the document frequency for each word in the training corpus and removes those words whose document frequency is less than some predetermined threshold *T*.

B. Information Gain (IG)

Information gain was included as the well known measure successfully used in some text-learning experiment. It is frequently employed as a term goodness criterion in the field of machine learning. It measures the number of bits of information obtained for category prediction by knowing the presence or absence of a word in a document. Namely, for unique words of the training set, we compute their information gain, whose information gain is less than some predetermined threshold are removed. Concretely, let $\{c_i\}_{i=1}^{m}$ denotes the set of categories in the target space. The information gain of word *t* is defined to be:

$$\begin{aligned} IG(t) = &-\sum_{i=1}^{m} P\left(C_i\right) \log P\left(C_i\right) \qquad (1) \\ &+ P(t) \sum_{i=1}^{m} P\left(C_i/t\right) \log P\left(C_i/t\right) \\ &+ P\left(\bar{t}\right) \sum_{i=1}^{m} P\left(C_i/\bar{t}\right) \log P\left(C_i/\bar{t}\right) \end{aligned}$$

This definition includes the estimation of the conditional probabilities of a category given a term, and the entropy computations in the definition. The probability estimation has a time complexity of O(N) and a space complexity of O(VN) where N is the number of training documents, and V is the vocabulary size. The entropy computation has a time complexity of O(Vm) [12].

C. Mutual Information (MI)

For feature term *t*, the way to duel with it is:

$$MI(t) = \sum_{i=1}^{m} P\left(C_i\right) \log {}^{P\left(t/C_i\right)}/_{P(t)} \qquad (2)$$

The MI computation has a time complexity of $O(Vm)$, similar to the IG computation.

D. Statistic (CHI)

For feature term *t*, the way to handle it like:

$$X^2\left(t, C_i\right) = \frac{N \times (AD - CB)^2}{(A + C)(B + D)(A + B)(C + D)} \qquad (3)$$

$$X_{avg}^2 = \sum_{i=1}^{m} X^2\left(t, C_i\right) \qquad (4)$$

where *A* is the number of times *t* and C_i co-occur, *B* is the number of times *t* occurs without C_i, *C* is the number of times C_i occurs without *t*, *D* is the number of times neither C_i nor *t* occurs, and *N* is total number of documents. The χ^2 statistic has a natural value of zero if *t* and C_i are independent. The computation of CHI scores has a quadratic complexity, similar to MI and IG.

E. Expected Cross Entropy

Expected Cross Entropy is proposed and evaluated by Mlademnic and Grobelnik [13], on which is based information theoretic ideas. For feature term *t*, we can duel with it as:

$$CE(t) = P(t) \sum_{i=1}^{m} P\left(C_i|t\right) \log \frac{P\left(C_i|t\right)}{P(C_i)} \qquad (5)$$

F. Term Strength

This method estimates term importance based on how commonly a term is likely to appear in related documents closely. It tries to derive document pairs whose similarity is above a threshold using the cosine value of the two document vectors. Term Strength then is computed based on the estimated conditional probability that a term occurs in the second half of a pair of related documents given that it occurs in the first half. Suppose *x* and *y* be an arbitrary pair of distinct but related documents, and t be a term, then the strength of the term is defined as: $s(t) = P_r(t \in y/t \in x)$.

G. Weight of Evidence for Text

For feature term *t*, the way to duel with it is:

$$\begin{aligned} WET(t) = P(t)\sum_{i=1}^{m} P(C_i) \\ \times \left|\log\frac{P(t|C_{pos})(1-P(t|C_{neg}))}{(1-P(t|C_{pos}))P(t|C_{neg})}\right| \end{aligned} \tag{6}$$

where $P(C_i)$ is the probability of the *i*-th document in corpus; $P(t)$ is the probability of term *t*, $P(\bar{t}) = 1 - P(t)$ is the probability of term *t* no-occur, $P\left(C_i|t\right)$ is the probability of *t* and C_i co-occur, $P\left(C_i/\bar{t}\right)$ is the probability of C_i occurs without *t*.

H. Odds Ratio

Odds ratio is commonly used in information retrieval, where the problem is to rank out documents according to their relevance for the positive class value using occurrence of different words as features. Our experiments show that this measure is especially suitable to be used in a combination with the Naive Bayesian classifier for our kind of problem. Odds ratio has always been used in two category case. A formula is defined as follow:

$$OR(t) = \log\frac{P(t|C_{pos})(1-P(t|C_{neg}))}{(1-P(t|C_{pos}))P(t|C_{neg})} \tag{7}$$

where C_{pos} represents the positive corpus case and C_{neg} is opposite case. If odds ratio is requested to suit for multi-class, we must change the formulas as:

$$\begin{aligned} MC-OR(t) &= \sum_{i=1}^{m} P(C_i) \times |OR(t, C_i)| \\ &= \sum_{i=1}^{m} P(C_i)\left|\log\frac{P(t|C_i)(1-P(t|C_{else}))}{P(t|C_{else})(1-P(t|C_i))}\right| \end{aligned} \tag{8}$$

where C_{else} represent all the classes accept *i*-th class. Naturally, the *i*-th class was seen as positive corpus and others were negative corpus together. So we have $P(t|C_{else}) = P(t) - P(t,C_i)\big/_{1-P(C_i)}$.

V. Classify Methods

A. Naive Bayes Classifier

A basic premise of Naive Bayes is the assumption of independence between characteristics. So we assume features are independent reciprocally. Though calculating the posterior probability of document, Classifier assigns it to the class which gets the biggest value. Namely:

$$\begin{aligned} C &= \max_i P(d|C_i)P(C_i) \\ &= \max_i P(C_i)\prod_k (t_k|C_i)^{N(t_k,d)} \end{aligned} \tag{9}$$

where

$$P(t_k|C_i) = \frac{1 + \sum_{l=1}^{d_i} tf(t_{kl})}{|V| + \sum_{j=1}^{|V|}\sum_{l=1}^{d_i} tf(t_{jl})} \tag{10}$$

represent the number of times t_j occurs in the *l*-th document labeled with C_i, $|V|$ is the total number of all of terms, d_i is the total number of documents labeled with C_i and $N(t_k, d)$ is the number of times term t_k occurs in document *d*.

B. SVM Classifier

Support Vector Machine is a popular technique for classification, and has been shown to be highly effective at traditional text categorization. For our research, we chose to use libSVM [14], an open source SVM package. The libSVM package has several kernel functions available, and we chose to use the radial basis functions (RBF) for training and testing, with all parameters set to their default values.

C. VSM Based on Cosine Similarity

According to the cosine similarity between documents and the center vector of each class, VSM labels the test document with the class which has the biggest value of similarity. Namely:

$$\begin{aligned} C &= \max_j Cos(d_i, V_j) \\ &= \frac{d_i \times V_j}{|d_i|\,|V_j|} \\ &= \sum_{l=1}^{n} w(t_{il})w(t_{jl}) \Big/ \sum_{l=1}^{n} w(t_{il})^2 \sum_{l=1}^{n} w(t_{jl})^2 \end{aligned} \tag{11}$$

VI. Experiment and Analysis

A. Experimental Settings and Data collections

For Chinese lexical analysis, we select the ELUS [15] system to segment and tag the corpus for our research. It is one of the best lexical analyzers in the Chinese natural language with high segmentation and tag accuracy.

In order to evaluate the eight feature selection methods, we design experiment I. In experiment I, all employed classifiers only employ bag-of-words. We designed experiment II to investigate whether the structural features is reliable to the task.

Using the definitions and genre class descriptions developed, the corpus we used for this paper was collected from the Internet. There are six thousand documents with human assigned genre labels in the full collection. The ratio of training set and testing set is 2-to-1.

B. Performance Measures

The effectiveness of a feature selection method is evaluated by the performance of classifier. For making a reasonable consideration about precision and recall, a comprehensive analysis about system's efficiency, we used both *Macro-F1* and *Micro-F1* as evaluation criterion. *Macro-F1* and *Micro-F1* were defined as follow.

$$Macro-F1=\sum_{i=1}^{m}\frac{N_i}{N}\times\frac{2\times precision_i\times recall_i}{precision_i+recall_i} \quad (12)$$

$$Micro-F1=\frac{1}{m}\times\sum_{i=1}^{m}\frac{2\times precision_i\times recall_i}{precision_i+recall_i} \quad (13)$$

where N_i is the total number of test documents which belong to the *i*-th class, and *N* is the total number of test documents. $precision_i$ and $recall_i$ are the *i*-th class's precision and recall respectively and there are *m* categories. Micro-averaging gives equal weight to every document, while macro-averaging gives equal weight to each category.

C. Primary Results

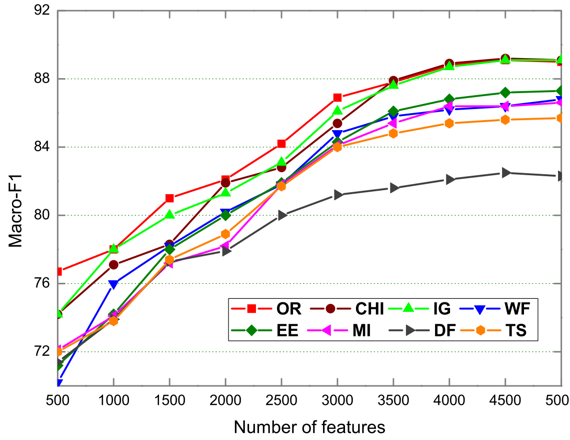

Fig. 1. Macro-F1 vs Number of Features(Feature set: BOW)

All of these three classifiers get good performance, and their curves conform to each other approximately. We only show Naive Bayes classifier's results as representation.

Figure 1 and Figure 2 give the result of experiment I, show the performance curves after feature selection with above methods. Experiments with OR outperformed others. Figure 3 and 4 demonstrate experiment II's results, display the performance curves with feature sets selected by above methods and additional structural features. Experiments show that structural feature set can significantly improve system performance.

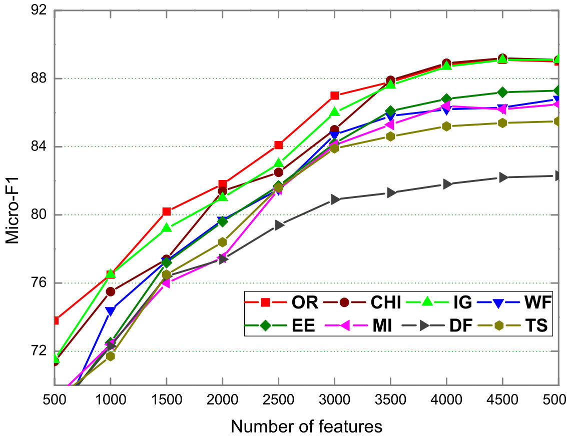

Fig. 2. Micro-F1 vs Number of Features(Feature set: BOW)

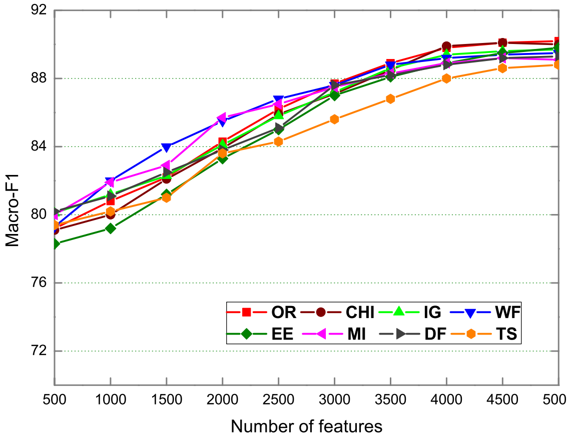

Fig. 3. Macro-F1 vs Number of Features(Feature set: BOW + Structural Features)

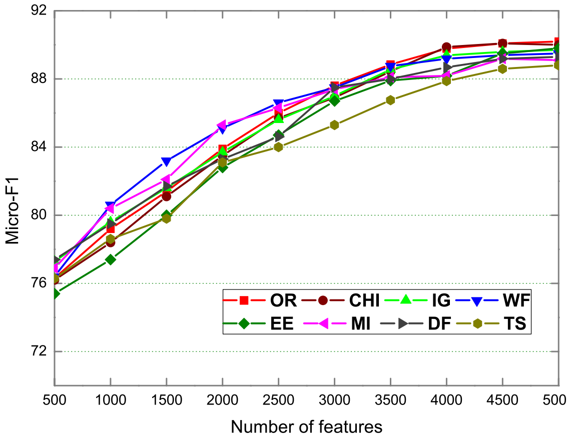

Fig. 4. Micro-F1 vs Number of Features(Feature set: BOW + Structural Features)

D. Discussion

In our experiments, IG and CHI are the most effective in aggressive term removal with up to 90% without losing categorization accuracy; while TS is comparable with up to 48% term removal on the finance corpus. The best performing feature scoring measures are Odd ratio, IG and CHI, the worst is TS. Others perform comparable or worse than Odds ratio and better than TS. Besides, when we chose to VSM classifier, the differences among feature selection methods were slight and the margins between curves were narrow. But when Naive Bayes classifier was adopted, odds ratio had a significant advantage than others. Although DF did not achieve a good result in our experiments, DF, the simplest technique for vocabulary reduction with the lowest cost in computation, can be reliably used instead of IG or CHI when the computations of these measures are too expensive. For Expected cross entropy, it is similar to Information gain. We can get this point form nearly synchronous trend. The difference is that instead of calculating average over all possible feature values, expected cross entropy only calculates the value denoting that word occurred in a document. For TS, it can reach to around 50% vocabulary reduction but is not first-rank at higher vocabulary reduction levels. Because this method does not use information about term category associations. In this sense, it is similar to the DF criterion, but different from the IG, MI and χ^2 statistic. The χ^2 statistic measures the lack of independence between word *w* and class c_j. But it cannot always get a normalized value. Hence, the χ^2 statistic is considered not to suit for unevenly distributed data collection [16]. On all domains odds ratio is among the best performance even only a small number of features used. Figures show that Odds ratio is one of the two best performing measures.

Figures can not reveal structural features' weakness, but it does exist. To sum up, the features used in the proposed methods are selected manually and subjectively, not derived by a statistical procedure. So this way is not suit for the case that people cannot understand categorization or features directly, because they can not select out appropriate words well and truly. Fortunately, we always make it collaborate with feature selection methods which can make up its defect. Generally, people can participate in feature selection artificially in most case.

VII. Conclusions

Feature selection attempts to remove non-informative words from documents in order to improve categorization effectiveness and reduce computational complexity. In this paper, eight feature selection methods are evaluated and given a comparative study. Different algorithms exhibit variously, among which odds ratio and χ^2 statistic show a little better effect than others. In empirical results on the real world corpora, structural features, which can improve the performance, can be applied successfully to identify the genres of Chinese finance text.

In the next stage, other classifiers and data collections will be exploited to make better distinctions and further discussions between feature selection methods.

Acknowledgment

This investigation was supported in part by the National Natural Science Foundation of China (No. 60435020 and No. 90612005) and the National 863 Program of China (No.2006AA01Z197).

References

[1] *The American Heritage Dictionary of the English Language*. Houghton Mifflin Company, 2000.

[2] B. Kessler, G. Nunberg, and H. Schutze, "Automatic detection of text genre," in *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics*, Madrid, Spain, 1997.

[3] E. Stamatatos, N. Fakotakis, and G. Kokkinakis, "Text genre detection using common word frequencies," in *Proceedings of the 18th Int. Conference on Computational Linguistics*, Saarbrucken, Germany, 2000.

[4] J. Xu, Y. Ding, X. Wang, and Y. Wu, "Genre identification of chinese finance text using machine learning method," in *Proceedings of the 2008 IEEE International Conference on Systems, Man, and Cybernetics*, Singapore, October 2008 To be appeared.

[5] E. Finn and N. Kushmerick, "Learning to classify documents according to genre," in *IJCAI-03 Workshop on Computational Approaches to Style Analysis and Synthesis*, 2003.

[6] M. Santini, "A shallow approach to syntactic feature extraction for genre classification." in *Proceedings of the 7th Annual Colloquium for the UK Special Interest Group for Computational Linguistics*, Birmingham, UK, 2004.

[7] E. Stamatatos, G. Kokkinakis, and N. Fakotakis, "Automatic text categorization in terms of genre and author," *Computational Linguistics*, vol. 26, no. 4, pp. 471–495, 2000.

[8] F. Peng, D. Schuurmans, and S. Wang, "Language and task independent text categorization with simple language models," in *Proceedings of Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology*, Edmonton, Canada, 2003.

[9] S. M. zu Eissen and B. Stein, *Genre Classification of Web Pages: User Study and Feasibility Analysis*. Berlin: Springer, 2004.

[10] E. S. Boese and A. E. Howe, "Effects of web document evolution on genre classification," in *Proceedings of the 14th Conference on Information and Knowledge Management*, Bremen, Germany, 2005.

[11] C. S. Lim, K. J. Lee, and G. C. Kim, "Multiple sets of features for automatic genre classification of web documents," *Information Processing and Management*, vol. 41, no. 5, pp. 1263–1276, September 2005.

[12] Y. Yang and J. O. Pedersen, "A comparative study on feature selection in text categorization." Morgan Kaufmann Publishers, 1997, pp. 412–420.

[13] D. Mladenic and M. Grobelnik, "Feature selection for unbalanced class distribution and naive bayes," in *In Proceedings of the 16th International Conference on Machine Learning (ICML*. Morgan Kaufmann Publishers, 1999, pp. 258–267.

[14] C.-C. Chang and C.-J. Lin, *LIBSVM:a library for support vector machines*, http://www.csie.ntu.edu.tw/ cjlin/libsvm/, 2001.

[15] W. JIANG, X.-L. WANG, Y. GUAN, and J. ZHAO, "Research on chinese lexical analysis system by fusing multiple knowledge sources," *Chinese Journal of Computers*, vol. 30, no. 1, pp. 137–145, 2007.

[16] T. Dunning, "Accurate methods for the statistics of surprise and coincidence," *COMPUTATIONAL LINGUISTICS*, vol. 19, pp. 61–74, 1993.