International Journal of Technical Research and Applications e-ISSN: 2320-8163,
www.ijtra.com Volume 4, Issue 1 (January-February, 2016), PP. 5-16

REPUTATION-BASED TRUST EVALUATION BY EXTRACTING USER'S ASSESSMENT

Neelam Sushma[1], Gajjala Venkata Kondareddy[2]
Computer science and engineering
N.B.K.R.I.S.T-(Neelam Sushma)
N.B.K.R.I.S.T-(Gajjala Venkata Kondareddy)
Vidyanagar, India
iamsushma814@gmail.com

Abstract— **Accurate trust evaluation is crucial for the success of Electronic-commerce systems. Position-based trust representations are mostly used and user's assessment rankings are combined to find out vendor's position. One major problem with this is that identifying whether everything is good with the vendor or not and it is tough for customers to find vendors who are believable, because rankings of vendors in present system is normally high for each vendor.**

In this paper, by observing the remarks of the customers, which were expressed freely in the form of text, Comment-based multi-dimensional trust model, is a fine-grained multi-dimensional trust evaluation model by extracting user's reply statements in electronic-commerce, and a procedure to extract user's remarks for dimension rankings is proposed. With this model comprehensive trust profiles are computed automatically for vendors, including dimension trust scores and weights, as well as overall trust ratings by combining dimension positional rankings.

This model provides an approach that combines dependency relation analysis and lexicon-based opinion mining techniques and further proposes an algorithm based on dependency relation analysis and Latent Dirichlet Allocation (LDA) topic modeling technique to cluster aspect expressions into dimensions and compute combined dimension rankings and weights. This algorithm is named as Lexical-LDA. This model can capably address the problem of identifying the good vendor and rank vendors efficiently.

Index terms- **lexical analysis, topic modeling, multidimensional,**

I. INTRODUCTION

A. What is E-commerce?

Electronic commerce, commonly known as E-commerce or ecommerce, is trading in products or services using computer networks, such as the Internet. Electronic commerce draws on technologies such as mobile commerce, electronic fund transfer, supply chain management, online transaction processing, electronic data interchange (EDI), inventory management systems, and automated data systems. Modern electronic commerce typically uses the World Wide Web for at least one part of the transaction's life cycle, although it may also use other technologies such as e-mail.

E-commerce businesses may employ some or all of the following:

- Online shopping web sites for retail sales direct to consumers
- Providing or participating in online market places, which process third-party business-to-consumer or consumer-to-consumer sales
- Business-to-business buying and selling
- Gathering and using demographic data through web contacts and social media
- Business-to-business electronic data interchange
- Marketing to prospective and established customers by e-mail or fax (for example, with newsletters)
- Engaging in retail for launching new products and services

To do online transactions, the user should trust the third party sellers. Accurate trust evaluation is crucial for the success of e-commerce systems. Reputation reporting systems have been implemented in e-commerce systems such as eBay and Amazon (for third-party sellers), where overall reputation scores for sellers are computed by aggregating feedback ratings. For example on eBay, the reputation score for a seller is the positive percentage score, as the percentage of positive ratings out of the total number of positive ratings and negative ratings in the past 12 months. A well-reported issue with the eBay reputation management system is the "all good reputation" problem, where feedback ratings are over 99% positive on average. Such strong positive bias can hardly guide buyers to select sellers to transact with. At eBay detailed seller ratings for sellers (DSRs) on four aspects of transactions, namely Item as described, communication, postage time, and postage and handling charges, are also reported. DSRs are aggregated rating scores on a 1- to 5-star scale. Still the strong positive bias is present – aspect ratings are mostly 4.8 or 4.9 stars. One possible reason for the lack of negative ratings at e-commerce web sites is that users who leave negative feedback ratings can attract retaliatory negative ratings and thus damage their own reputation.

Although buyers leave positive feedback ratings, they express some disappointment and negativeness in free text feedback comments in some aspects. By analyzing the wealth of information in feedback comments it can uncover buyers' detailed embedded opinions towards different aspects of transactions, and compute comprehensive reputation profiles for sellers.

Comments are short and therefore co-occurrence of head terms in comments is not very informative. We instead use the co-occurrence of dimension expressions with respect to a same modifier across comments, which potentially can provide more meaningful contexts for dimension expressions. We observe that it is very rare that the same aspect of e-commerce transactions is commented more than once in the same feedback comment. In other words, it is very unlikely that the dimension expressions extracted from the same comment are about the same topic.

We propose Comment-based Multi-dimensional trust (CommTrust), a fine-grained multi dimensional trust evaluation model by mining e-commerce feedback comments. With CommTrust, comprehensive trust profiles are computed for sellers, including dimension reputation scores and weights, as well as overall trust scores by aggregating dimension reputation scores. To the best of our knowledge, CommTrust is the first piece of work that computes fine-grained multidimensional trust profiles automatically by mining feedback comments. In later discussions, we use the terms reputation score and trust score interchangeably. In CommTrust, we propose an approach that combines dependency relation analysis, a tool recently developed in natural language processing (NLP) and lexicon based opinion mining techniques, to extract aspect opinion expressions from feedback comments and identify their opinion orientations. We further propose an algorithm based on dependency relation analysis and Latent Dirichlet Allocation (LDA) topic modeling technique to cluster aspect expressions into dimensions and compute aggregated dimension ratings and weights. We call our algorithm Lexical-LDA.

Rating aggregation algorithms for computing individual reputation scores include simple positive feedback percentage or average of star ratings as in the eBay and Amazon reputation systems, the Beta reputation based on statistical distribution assumption for ratings, as well as more advanced models, which also computes trust score variance and confidence level. More sophisticated reputation models consider factors like time, where recent feedback ratings.

Related work falls into three main areas: 1) computational approaches to trust, especially reputation-based trust evaluation and recent developments in fine-grained trust evaluation; 2) e-commerce feedback comments analysis and 3) aspect opinion extraction and summarization on movie reviews, product reviews and other forms of free text. Similar to that buyers and sellers are referred to as individuals in e-commerce applications, terms like peers and agents are often used to refer to individuals in open systems in various applications in the trust evaluation literature. In a comprehensive overview of trust models is provided. Individual level trust models are aimed to compute the reliability of peers and assist buyers in their decision making whereas system level models are aimed to regulate the behavior of peers, prevent fraudsters and ensure system security.

II. WHAT IS DATA MINING?

A. *Introduction*

Data mining, also called knowledge discovery in data bases, in computer sciences, the process of discovering interesting and useful patterns and relationships in large volumes of data. The field combines tools from statistics and artificial intelligence such as neural networks and machine learning with database management to analyze large digital collections, known as data sets. Data mining is widely used in business (insurance, banking, retail), science research (astronomy, medicine), and government security (detection of criminals and terrorists).

B. *Data Mining Overview*

Data mining is emerging as one of the key features of many homeland security initiatives. Often used as a means for detecting fraud, assessing risk, and product retailing, data mining involves the use of data analysis tools to discover previously unknown. In the context of homeland security, data mining is often viewed as a potential means to identify terrorist activities, such as money transfers and communications, and to identify and track individual terrorists themselves, such as through travel and immigration records. While data mining represents a significant advance in the type of analytical tools currently available, there are limitations to its capability. One limitation is that although data mining can help reveal patterns and relationships, it does not tell the user the value or significance of these patterns. These types of determinations must be made by the user. A second limitation is that while data mining can identify connections between behaviors and/or variables, it does not necessarily identify a causal relationship. To be successful, data mining still requires skilled technical and analytical specialists who can structure the analysis and interpret the output that is created. Data mining is becoming increasingly common in both the private and public sectors. Industries such as banking, insurance, medicine, and retailing commonly use data mining to reduce costs, enhance research, and increase sales. In the public sector, data mining applications initially were used as a means to detect fraud and waste, but have grown to also be used for purposes such as measuring and improving program performance. However, some of the homeland security data mining applications represent a significant expansion in the quantity and scope of data to be analyzed. Two efforts that have attracted a higher level of congressional interest include the Terrorism Information Awareness (TIA) project (now-discontinued) and the Computer-Assisted Passenger Pre-screening System II (CAPPS II) project (now- cancelled and replaced by Secure Flight). As with other aspects of data mining, while technological capabilities are important, there are other implementation and oversight issues that can influence the success of a project's outcome. One issue is data quality, which refers to the accuracy and completeness of the data being analyzed. A second issue is the interoperability of the data mining software and databases being used by different agencies. A third issue is mission creep, or the use of data for purposes other than for which the data were originally collected. A fourth issue is privacy. Questions that may be considered include the degree to which government agencies should use and mix commercial data with government data, whether data sources are being used for purposes other than those for which they were originally designed, and possible application of the Privacy Act to these initiatives. It is anticipated that congressional oversight of data mining projects will grow as data mining efforts continue to evolve.

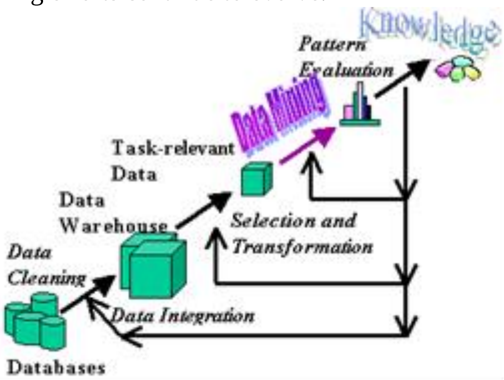


Fig. 1 shows that how knowledge will be extracted, first when we gets data from database, unwanted data and replicated data will be eliminated and it goes to data warehouse, from there relevant data will be selected and transformed and sent to pattern evaluation, after identifying the pattern knowledge has discovered and it is continuous process.

III. DATA MINING APPLICATIONS

Data mining is a process that analyzes the large amount of data to find the new and hidden information that improves business efficiency. Various industries have been adopting data mining to their mission-critical business processes to gain competitive advantages and help business grows. This tutorial illustrates some data mining applications in sale/marketing, banking/finance, health care and insurance, transportation and medicine.

A. Data Mining Applications in Sales/Marketing

Data mining enables the businesses to understand the patterns hidden inside past purchase transactions, thus helping in plan and launch new marketing campaigns in prompt and cost effective way. The following illustrates several data mining applications in sale and marketing.

- Data mining is used for market basket analysis to provide insight information on what product combinations were purchased, when they were bought and in what sequence by customers. This information helps businesses to promote their most profitable products to maximize the profit. In addition, it encourages customers to purchase related products that they may have been missed or overlooked.
- Retails companies uses data mining to identify customer's behavior buying patterns.

B. Data Mining Applications in Banking / Finance

- Several data mining techniques such as distributed data mining has been researched, modeled and developed to help credit card fraud detection.
- Data mining is used to identify customer's loyalty by analyzing the data of customer's purchasing activities such as the data of frequency of purchase in a period of time, total monetary value of all purchases and when was the last purchase. After analyzing those dimensions, the relative measure is generated for each customer. The higher of the score, the more relative loyal the customer is.
- To help bank to retain credit card customers, data mining is used. By analyzing the past data, data mining can help banks to predict customers that likely to change their credit card affiliation so they can plan and launch different special offers to retain those customers.
- Credit card spending by customer groups can be identified by using data mining.
- The hidden correlations between different financial indicators can be discovered by using data mining.
- From historical market data, data mining enable to identify stock trading rules.

C. Data Mining Applications in Health Care and Insurance

The growth of the insurance industry is entirely depends on the ability of converting data into the knowledge, information or intelligence about customers, competitors and its markets. Data mining is applied in insurance industry lately but brought tremendous competitive advantages to the companies who have implemented it successfully. The data mining applications in insurance industry are listed below:

- Data mining is applied in claims analysis such as identifying which medical procedures are claimed together.
- Data mining enables to forecasts which customers will potentially purchase new policies.
- Data mining allows insurance companies to detect risky customers' behavior patterns.
- Data mining helps detect fraudulent behavior.

D. Data Mining Applications in Transportation

- Data mining helps to determine the distribution schedules among warehouses and outlets and analyze loading patterns.

E. Data Mining Applications in Medicine

- Data mining enables to characterize patient activities to see coming office visits.
- Data mining help identify the patterns of successful medical therapies for different illnesses

.

Data mining applications are continuously developing in various industries to provide more hidden knowledge that enable to increase business efficiency and grow businesses.

IV. ADVANTAGES OF DATA MINING

A. Marking/Retailing

Data mining can aid direct marketers by providing them with useful and accurate trends about their customer's purchasing behavior. Based on these trends, marketers can direct their marketing attentions to their customers with more precision. For example, marketers of a software company may advertise about their new software to consumers who have a lot of software purchasing history. In addition, data mining may also help marketers in predicting which products their customers may be interested in buying. Through this prediction, marketers can surprise their customers and make the customer's shopping experience becomes a pleasant one.

Retail stores can also benefit from data mining in similar ways. For example, through the trends provide by data mining, the store managers can arrange shelves, stock certain items, or provide a certain discount that will attract their customers.

B. Banking/Crediting

Data mining can assist financial institutions in areas such as credit reporting and loan information. For example, by examining previous customers with similar attributes, a bank can estimated the level of risk associated with each given loan. In addition, data mining can also assist credit card issuers in detecting potentially fraudulent credit card transaction. Although the data mining technique is not a 100% accurate in its prediction about fraudulent charges, it does help the credit card issuers reduce their losses.

C. *Law enforcement*

Data mining can aid law enforcers in identifying criminal suspects as well as apprehending these criminals by examining trends in location, crime type, habit, and other patterns of behaviors.

D. Researchers

Data mining can assist researchers by speeding up their data analyzing process; thus, allowing those more time to work on other projects.

V. SURVEY OF THE LITERATURE

The Internet has created vast new opportunities to interact with strangers. The interactions can be fun, informative, and even profitable. But they also involve risks. Is the advice from a self-proclaimed expert at expertcentral.com reliable? Will an unknown dot-com site or eBay seller ship with appropriate packaging, and will the product be as described? Before the Internet, such questions were answered, in part, through reputations. Vendors provided references, Better Business Bureaus tallied complaints, and past personal experience and person-to-person gossip told you whom you could rely upon and whom you could not. And a businessman's standing in the community, e.g., his role at church, served as a valuable hostage. Internet services operate on a vastly larger scale than Main Street and permit virtually anonymous interactions. Nevertheless, reputation systems are playing a major role. Systems are emerging that respect anonymity and operate on the Internet's scale.

A. Literature Survey

One question has raised that to provide trustworthiness among strangers reputation systems are important or not?[1] To answer this question, author examine that how a person will going to trust a stranger naturally, First, when you interact with someone who is unknown to you, you'll check his past history of transactions. Second, the expectation of good product quality and good communication in future interactions gives an incentive for good behavior. Strangers do not have known past histories or the prospect of future interactions, and they are not subject to a network of informed individuals who will punish bad and reward good behavior toward any of them. Reputation systems seek to restore the shadow of the future to each transaction by creating an expectation that other people will look back upon it.

Fame of sellers that are transferred from one to one may not be absolutely true and they may discourages the people [2]. On the Internet, data about history of transactions may be and they may not be reliable, but these are little better than the gossips spread by the people and these are systematic. One of the earliest and best known Internet reputation systems is run by eBay, which gathers comments from buyers and sellers about each other after each transaction. Examination of a large data set from 1999 reveals several interesting features of this system, which facilitates many millions of sales each month. First, despite incentives to free ride, feedback was provided more than half the time. Second, well beyond reasonable expectation, it was almost always positive. Third, reputation profiles were predictive of future performance. However, the net feedback scores that eBay displays encourages Pollyanna assessments of reputations, and is far from the best predictor available. Fourth, although sellers with better reputations were more likely to sell their items, they enjoyed no boost in price, at least for the two sets of items that we examined. Fifth, there was a high correlation between buyer and seller feedback, suggesting that the players reciprocate and retaliate.

Buyers and sellers in online auctions are faced with the task of deciding who to entrust their business based on a very limited amount of information. [3]Current trust ratings on eBay average over 99% positive and are presented as a single number on a user's profile. This paper presents a system capable of extracting valuable negative information from the wealth of feedback comments on eBay, computing personalized and feature-based trust and presenting this information graphically.

A system for extracting typed dependency parses [4] of English sentences from phrase structure parses. In order to capture inherent relations occurring in corpus texts that can be critical in real-world applications, many NP relations are included in the set of grammatical relations used. We provide a comparison of our system with Minipar and the Link parser. The typed dependency extraction facility described here is integrated in the Stanford Parser, available for download.

Latent Dirichlet allocation (LDA) [6], a generative probabilistic model for collections of discrete data such as text corpora. LDA is a three-level hierarchical Bayesian model, in which each item of a collection is modeled as a finite mixture over an underlying set of topics. Each topic is, in turn, modeled as an infinite mixture over an underlying set of topic probabilities. In the context of text modeling, the topic probabilities provide an explicit representation of a document. We present efficient approximate inference techniques based on various methods and an EM algorithm for empirical Bayes parameter estimation. We report results in document modeling, text classification, and collaborative filtering, comparing to a mixture of unigrams model and the probabilistic LSI model.

The first randomized controlled field experiment of an Internet reputation mechanism was explained in [9]. A high-reputation, established eBay dealer sold matched pairs of lots -- batches of vintage postcards -- under his regular identity and under new seller identities (also operated by him) . As predicted, the established identity fared better. The difference in buyers' willingness-to-pay was 8.1% of the selling price. A subsidiary experiment followed the same format, but compared sales by relatively new sellers with and without negative feedback. Surprisingly, one or two negative feedbacks for our new sellers did not affect buyers' willingness-to-pay.

Buyers in online auctions write feedback comments to the sellers from whom they have bought the items. Other bidders read them to determine which item to bid for. In this research, we aim at helping bidders by summarizing the feedback comments. First, we examine feedback comments in online auctions. From the results of the examination, we propose a method called social summarization method. It uses social relationships in online auctions for summarizing feedback comments. This method extracts feedback comments which the buyers seemed to have written from their heart. We implement a system based on our method and evaluate its effectiveness. The results are that our method deleted 80.8% of courteous comments (comments with almost no information). We also found that there are two types of comments in the summaries: comments that are generally infrequent and seem to have been written with real feeling and comments that are generally frequent and seem to have been written with real feeling. Finally, we propose an interactive presentation method of the summaries which identifies the types of the comments. The user experiment indicates that this presentation helps users judge which seller to bid for.

In opinion mining of product reviews [15], one often wants to produce a summary of opinions based on product features/attributes. However, for the same feature, people can express it with different words and phrases. To produce an effective summary, these words and phrases, which are domain synonyms, need to be grouped under the same feature. Topic modeling is a suitable method for the task. However, instead of simply letting topic modeling find groupings freely, we believe it is possible to do better by giving it some pre-existing knowledge in the form of automatically extracted constraints. In this paper, we first extend a popular topic modeling method, called LDA, with the ability to process large scale constraints. Then, two novel methods are proposed to extract two types of constraints automatically. Finally, the resulting constrained-LDA and the extracted constraints are applied to group product features. Experiments show that constrained-LDA outperforms the original LDA and the latest mLSA by a large margin. The problem of topic-sentiment analysis [16] on Weblogs and proposes a novel probabilistic model to capture the mixture of topics and sentiments simultaneously. The proposed Topic-Sentiment Mixture (TSM) model can reveal the latent topical facets in a Weblog collection, the subtopics in the results of an ad hoc query, and their associated sentiments. It could also provide general sentiment models that are applicable to any ad hoc topics. With a specifically designed HMM structure, the sentiment models and topic models estimated with TSM can be utilized to extract topic life cycles and sentiment dynamics. Empirical experiments on different Weblog datasets show that this approach is effective for modeling the topic facets and sentiments and extracting their dynamics from Weblog collections. The TSM model is quite general; it can be applied to any text collections with a mixture of topics and sentiments, thus has many potential applications, such as search result summarization, opinion tracking, and user behavior prediction.

VI. EXISTING SYSTEM

A. Introduction

The Web contains a wealth of opinions about products, politicians, and more, which are expressed in newsgroup posts, review sites, and elsewhere. As a result, the problem of "opinion mining" has seen increasing attention over the last three years from (Turney, 2002; Hu and Liu, 2004) and many others. This paper focuses on product reviews, though our methods apply to a broader range of opinions. Product reviews on Web sites such as amazon.com and elsewhere often associate meta-data with each review indicating how positive (or negative) it is using a 5-star scale, and also rank products by how they fare in the reviews at the site. However, the reader's taste may differ from the reviewers'. For example, the reader may feel strongly about the quality of the gym in a hotel, whereas many reviewers may focus on other aspects of the hotel, such as the decor or the location. Thus, the reader is forced to wade through a large number of reviews looking for information about particular features of interest. We decompose the problem of review mining into the following main subtasks:

I. Identify product features.
II. Identify opinions regarding product features.
III. Determine the polarity of opinions.
IV. Rank opinions based on their strength.

B. Existing System

This system introduces OPINE, an unsupervised information extraction system that embodies a solution to each of the above subtasks. OPINE is built on top of the Know- It All Web information-extraction system (Etzioni et al., 2005). Given a particular product and a corresponding set of reviews, OPINE solves the opinion mining tasks outlined above and outputs a set of *product features*, each accompanied by a list of *associated opinions* which are ranked based on strength (*e.g.*, "abominable" is stronger than "bad). This output information can then be used to generate various types of opinion summaries. This paper focuses on the first 3 review mining subtasks and our contributions are as follows:

1. We introduce OPINE, a review-mining system whose novel components include the use of *relaxation labeling* to find the semantic orientation of words in the context of given product features and sentences.
2. We compare OPINE with the most relevant previous review-mining system (Hu and Liu, 2004) and find that OPINE's precision on the *feature extraction* task is 22% better though its recall is 3% lower on Hu's data sets. We show that 1/3 of this increase in precision comes from= using *assessment* mechanism on review data while the rest is due to Web PMI statistics.
3. While many other systems have used extracted opinion phrases in order to determine the polarity of sentences or documents, OPINE is the first to report its precision and recall on the tasks of *opinion phrase extraction* and *opinion phrase polarity determination* in the context of known product features and sentences.

The key components of OPINE described in this paper are the PMI feature assessment which leads to high-precision feature extraction and the use of relaxation-labeling in order to find the semantic orientation of potential opinion words. The review-mining work most relevant to our research is that of (Hu and Liu, 2004) and (Kobayashi et al., 2004). Both identify product features from reviews, but OPINE significantly improves on both. (Hu and Liu, 2004) doesn't assess candidate features, so its precisions lower than OPINE's. (Kobayashi et al., 2004) employs an iterative semi-automatic approach which requires human input at every iteration. Neither model explicitly addresses *composite* (feature of feature) or *implicit* features. Other systems (Morinaga et al., 2002; Kushal et al., 2003)also look at Web product reviews but they do not extract opinions about particular product features. OPINE's use of metonymy lexico-syntactic patterns is similar to that of many others, from (Berland and Charniak, 1999) to(Almuhareb and Poesio, 2004).Recognizing the subjective character and polarity of words, phrases or sentences has been addressed by many authors, including (Turney, 2003; Riloff et al., 2003;Wiebe, 2000; Hatzivassiloglou and McKeown, 1997). Most recently, (Takamura et al., 2005) reports on the use of spin models to infer the semantic orientation of words. The paper's global optimization approach and use of multiple sources of constraints on a word's semantic orientation is similar to ours, but the mechanism differs and they currently omit the use of syntactic information. Subjective phrases are used by (Turney, 2002; Pang and Vaithyanathan, 2002; Kushal et al., 2003; Kim and Hovy,2004) and others in order to classify reviews or sentences as positive or negative. So far, OPINE's focus has been on extracting and analyzing opinion phrases corresponding to specific features in specific sentences, rather than on determining sentence or review polarity.

A well-reported issue with the eBay reputation management system is the "all good reputation" problem where feedback ratings are over 99% positive on average. Such strong positive bias can hardly guide buyers to select sellers to transact with. At eBay detailed seller ratings for sellers (DSRs) on four aspects of transactions, namely *item as described*, *communication*, *postage time*, and *postage and handling charges*, are also reported. DSRs are aggregated rating scores on a 1- to 5-star scale. Still the strong positive bias is present – aspect ratings are mostly 4.8 or 4.9 stars. One possible reason for the lack of negative ratings at e-commerce web sites is that users who leave negative feedback ratings can attract retaliatory negative ratings and thus damage their own reputation.

Disadvantages of Existing System

- All good reputation problem
- Buyers are feel like complex to select trustworthiness seller
- Lack of negative ratings

VII. PROPOSED SYSTEM

A. Overview

We propose Comment-based Multi-dimensional trust (CommTrust), a fine-grained multi-dimensional trust evaluation model by mining e-commerce feedback comments. With CommTrust, comprehensive trust profiles are computed for sellers, including dimension reputation scores and weights, as well as overall trust scores by aggregating dimension reputation scores. To the best of our knowledge, CommTrust is the first piece of work that computes fine-grained multi dimension trust profiles automatically by mining feedback comments. In later discussions, we use the terms *reputation score* and *trust score* interchangeably. In CommTrust, we propose an approach that combines dependency relation analysis, a tool recently developed in natural language processing (NLP) and lexicon based opinion mining techniques to extract aspect opinion expressions from feedback comments and identify their opinion orientations. We further propose an algorithm based on dependency relation analysis and Latent Dirichlet Allocation (LDA) topic modeling technique to cluster aspect expressions into dimensions and compute aggregated dimension ratings and weights. We call our algorithm Lexical-LDA. Unlike conventional topic modeling formulation of unigram representations for textual documents our clustering is performed on the dependency relation representations of aspect opinion expressions. As a result we make use of the structures on aspect and opinion terms, as well as negation defined by dependency relations to achieve more effective clustering. To specifically address the positive bias in overall ratings, our dimension weights are computed directly by aggregating aspect opinion expressions rather than regression from overall ratings.

This system calculates trust for sellers to provide trust between sellers and third party sellers. With this trust profiles one can allow the sellers who has high trust scores to upload their products in the website. With this buyers can select trustworthy sellers to transact with.

B. Commtrust: Comments-Based Multi-Dimensional Trust Evaluation

If a buyer gives a positive rating for a transaction, s/he still leaves comments of mixed opinions regarding different aspects of transactions in feedback comments. These salient aspects are called dimensions of e-commerce transactions. Comments-based trust evaluation is therefore multi-dimensional. The trust score on a dimension for a seller is the probability that buyers expect the seller to carry out transactions on this dimension satisfactorily. The trust score for a dimension can be estimated from the number of observed positive and negative ratings towards the dimension.

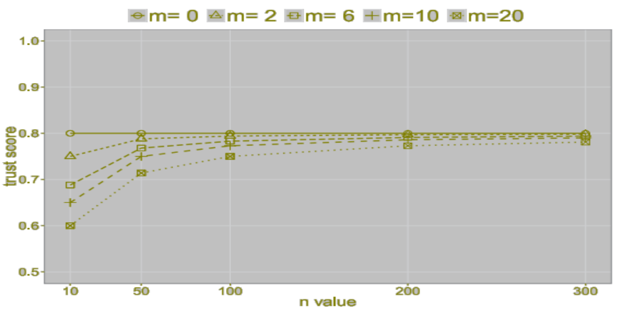


Fig. 2 Dimension Trust Model

Fig. 2 plots trust score td in relation to different settings of total number of ratings n and pseudo counts m. The figure is plotted for y/n = 0.8 and similar trends are observed for other values of y/n. It shows that when the total number of observed ratings n is large ($n \geq 300$), td is not very sensitive to the settings of m and converges to the observed positive rating frequency of 0.8. When there are a limited number of observed ratings, that is $n < 300$, an observed high positive rating frequency y/n is very likely an overestimation and so m is set to regulate the estimated value for td. With m = 2, td $\approx$ 0.8 when $n \geq 50$. On the other hand, with m= 20, td $\approx$ 0.8 only when $n \approx 300$. From our experiments, settings of m = 6... 20 typically give stable results. By default, we set m = 6.

C. Dependency Relation Analysis

Natural language processing (**NLP**) is a field of computer science, artificial intelligence, and linguistics concerned with the interactions between computers and human (natural) languages. As such, NLP is related to the area of human–computer interaction. Many challenges in NLP involve natural language understanding, that is, enabling computers to derive meaning from human or natural language input, and others involve natural language generation.

The following is a list of some of the most commonly researched tasks in NLP:

1. **Automatic summarization**-Produce a readable summary of a chunk of text. Often used to provide summaries of text of a known type, such as articles in the financial section of a newspaper.
2. **Machine translation**-Automatically translate text from one human language to another. This is one of the most difficult problems, and is a member of a class of problems colloquially termed "AI-complete", i.e. requiring all of the different types of knowledge that humans possess (grammar, semantics, facts about the real world, etc.) in order to solve properly.
3. **Natural language generation**-Convert information from computer databases into readable human language.
4. **Natural language understanding**-Convert chunks of text into more formal representations such as first-order logic structures that are easier for computer programs to manipulate. Natural language understanding involves the identification of the intended semantic from the multiple possible semantics which can be derived from a natural language expression which usually takes the form of organized notations of natural languages concepts. Introduction and creation of language metamodel and ontology are efficient however empirical solutions. An explicit formalization of natural languages semantics without confusions with implicit assumptions such as closed world assumption (CWA) vs. open world assumption, or subjective Yes/No vs. objective True/False is expected for the construction of a basis of semantics formalization.
5. **Sentiment analysis**-Extract subjective information usually from a set of documents, often using online reviews to determine "polarity" about specific objects. It is especially useful for identifying trends of public opinion in the social media, for the purpose of marketing.

Etc.…

The typed dependency relation representation is a recent NLP tool to help understand the grammatical relationships in sentences. With typed dependency relation parsing, a sentence is represented as a set of dependency relations between pairs of words in the form of (head, dependent), where content words are chosen as heads, and other related words depend on the heads. Fig. shows an example of analyzing the comment “Super quick shipping. Product was excellent. A great deal. ALL 5 STAR.” using the Stanford typed dependency relation parser. The comment comprises four sentences, and the sentence “Super quick shipping." is represented as three dependency relations. Shipping does not depend on any other words and is at the root level. The adjective modifier relations amod (shipping-3, super-1) and amod (shipping-3, quick-2) indicate that super modifies ship-ping and quick modifies shipping. The number following each word (e.g., shipping-3) indicates the position of this word in a sentence. Words are also annotated with their POS tags such as noun (NN), verb (VB), adjective (JJ) and adverb (RB).

If a comment expresses opinion towards dimensions then the dimension words and the opinion words should form some dependency relations. It has been reported that phrases formed by adjectives and nouns, and verbs and adverbs express subjectivity. Among the dependency relations expressing grammatical relationships, we select the relations that express the modifying relation between adjectives and nouns, and adverbs and verbs, as determined by the dependency relation parser. These modifying relations are listed in Table 1. It can be seen that with the modifying relations generally the noun or verb expresses the target concept under consideration whereas the adjective or adverb expresses opinion towards the target concept. The modifying relations thus can be denoted as (modifier, head) pairs. With the example comment in Fig. 3, the dependency relations adjective modifier amod (NN, JJ) and normal subject nsubj (JJ, NN) suggest the (modifier, head) pairs including (super, shipping), (quick, shipping), (excellent, product) and (great, deal). We call these (modifier, head) pairs dimension expressions.

TABLE 1. Dependency relations for dimension expressions

Dependency relation pattern	example
adjective modifier: *amod*(NN, JJ)	Super *quick shipping*.
adverbial modifier: *advmod*(VB, RB)	Great dealer *fast shipping*
nominal subject: *nsubj*(JJ, NN)	*Product* was *excellent*
adjectival complement: *comp*(VB, JJ)	Great CD, *arrived quick*.

NN: noun, VB: verb, JJ: adjective, and RB: adverb.
Dimension expressions are highlighted.

Ratings from dimension expressions towards the head terms are identified by identifying the prior polarity of the modifier terms by SentiWordNet, a public opinion lexicon. The prior polarities of terms in SentiWordNet include positive, negative or neutral, which corresponds to the ratings of +1, -1 and 0. Negations of dimension expressions are identified by the Neg () relation of the dependency relation parser. When a negation relation is detected the prior polarity of the modifier term is inverted.

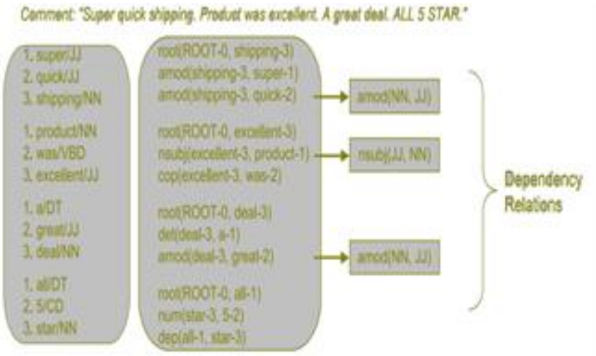


Fig. 3: Typed Dependency Relation Analysis

Clustering Dimension Expressions into Dimensions

We propose the Lexical-LDA algorithm to cluster aspect expressions into semantically coherent categories, which we call dimensions. Different from the conventional topic modeling approach, which takes the document by term matrix as input, Lexical-LDA makes use of shallow lexical knowledge of dependency relations for topic modeling to achieve more effective clustering. We make use of two types of lexical knowledge to "supervise" clustering dimension expressions into dimensions so as to produce meaningful clusters.

• Comments are short and therefore co-occurrence of head terms in comments is not very informative. We instead use the co-occurrence of dimension expressions with respect to a same modifier across comments, which potentially can provide more meaningful contexts for dimension expressions.

• We observe that it is very rare that the same aspect of e-commerce transactions is commented more than once in the same feedback comment. In other words, it is very unlikely that the dimension expressions extracted from the same comment are about the same topic.

With the shallow lexical knowledge of dependency relation representation for dimension expressions, the clustering problem is formulated under topic modeling as follows: The dimension expressions for a same modifier term or negation of a modifier term are generated by a distribution of topics, and each topic is generated in turn by a distribution of head terms. This formulation allows us to make use of the structured dependency relation representations from the dependency relation parser for clustering. Input to Lexical-LDA is dependency relations for dimension expressions in the form of (modifier, head) pairs or their negations, like (fast, shipping) or (not-good, seller).

Latent Dirichlet Allocation

In natural language processing, latent Dirichlet allocation (LDA) is a generative model that allows sets of observations to be explained by unobserved groups that explain why some parts of the data are similar. For example, if observations are words collected into documents, it posits that each document is a mixture of a small number of topics and that each word's creation is attributable to one of the document's topics.

Topics in LDA -In LDA, each document may be viewed as a mixture of various topics. This is similar to probabilistic latent semantic analysis (pLSA), except that in LDA the topic distribution is assumed to have a Dirichlet prior. In practice, this results in more reasonable mixtures of topics in a document. It has been noted, however, that the pLSA model is equivalent to the LDA model under a uniform Dirichlet prior distribution.[2]

For example, an LDA model might have topics that can be classified as CAT_related and DOG_related. A topic has probabilities of generating various words, such as milk, meow, and kitten, which can be classified and interpreted by the viewer as "CAT_related". Naturally, the word cat itself will have high probability given this topic. The DOG_related topic likewise has probabilities of generating each word: puppy, bark, and bone might have high probability. Words without special relevance, such as the, will have roughly even probability between classes (or can be placed into a separate category). A topic is not strongly defined, neither semantically nor epistemologically. It is identified on the basis of supervised labeling and (manual) pruning on the basis of their likelihood of co-occurrence. A lexical word may occur in several topics with a different probability, however, with a different typical set of neighboring words in each topic. Each document is assumed to be characterized by a particular set of topics. This is akin to the standard bag of words model assumption, and makes the individual words exchangeable.

Topic Modeling in LDA- In machine learning and natural language processing, a topic model is a type of statistical model for discovering the abstract "topics" that occur in a collection of documents. Intuitively, given that a document is about a particular topic, one would expect particular words to appear in the document more or less frequently: "dog" and "bone" will appear more often in documents about dogs, "cat" and "meow" will appear in documents about cats, and "the" and "is" will appear equally in both. A document typically concerns multiple topics in different proportions; thus, in a document that is 10% about cats and 90% about dogs, there would probably be about 9 times more dog words than cat words. A topic model captures this intuition in a mathematical framework, which allows examining a set of documents and discovering, based on the statistics of the words in each, what the topics might be and what each document's balance of topics is.

VIII. ADVANTAGES

1. We Solve the all good reputation problem in proposed system.
2. Assign the ranks for sellers based on trust scores.
3. Very easy for customers to select the suitable seller.

ARCHITECTURE

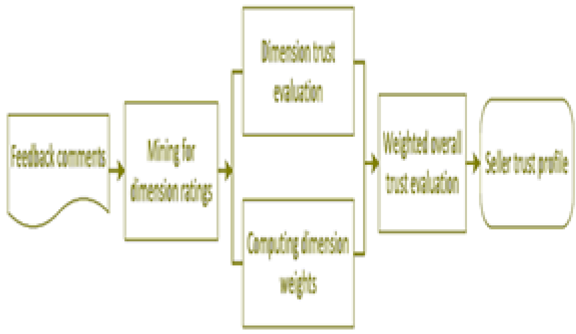


Fig. 4 Commtrust Framework

Fig. 4 depicts the CommTrust framework. Aspect opinion expressions and their associated ratings (positive or negative) are first extracted from feedback comments. Dimension trust scores together with their weights are further computed by clustering aspect expressions into dimensions and aggregating the dimension ratings. With this aggregation a weighted overall trust will be evaluated and trust profiles for sellers will be generated. The following are the required steps to evaluate a seller.

1. Comments Based Multi-Dimensional Trust Evaluation
2. Mining Feedback Comments For Dimension Ratings And Weights
3. Clustering Dimension Expressions into Dimensions
4. Assign the Rank for Sellers Based On Trusted Score

1. Comments Based Multi-Dimensional Trust Evaluation

We view feedback comments as a source where buyers express their opinions more honestly and openly. Our analysis of feedback comments on eBay and Amazon reveals that even if a buyer gives a positive rating for a transaction, s/he still leaves comments of mixed opinions regarding different aspects of transactions in feedback comments. Table 5.1 lists some sample comments, together with their rating from eBay. For example for comment c2, a buyer gave a positive feedback rating for a transaction, but left the following comment: “bad communication, will not buy from again. Super slow ship (ping). Item as described.” Obviously the buyer has negative opinion towards the communication and delivery aspects of the transaction, despite an overall positive feedback rating towards the transaction. We call these salient aspects dimensions of e-commerce transactions. Comments-based trust evaluation is therefore multi-dimensional.

TABLE 2. Sample comments on eBay

Some Sample Comments on eBay

No	Comment	eBay rating
c_1	beautiful item! highly recommend using this seller!	1
c_2	bad communication, will not buy from again. super slow ship(ping). item as described.	1
c_3	quick response	1
c_4	looks good. nice product, slow delivery though.	1
c_5	top seller. many thanks. A+	1
c_6	great price and awesome service! thank you!	1
c_7	product arrived swiftly! great seller.	1
c_8	great item. best seller of ebay	1
c_9	slow postage. didn't have the product asked for, but seller was friendly.	1
c_{10}	wrong color was sent. item was damaged. did not even fit phone.	1

2. Mining Feedback Comments for Dimension Ratings and Weights

We will first describe our approach based on the typed dependency analysis to extracting aspect opinion expressions and identifying their associated ratings. We then propose an algorithm based on LDA for clustering dimension expressions into dimensions and computing dimension weights.

3. Clustering Dimension Expressions into Dimensions

We propose the Lexical-LDA algorithm to cluster aspect expressions into semantically coherent categories, which we call dimensions. Different from the conventional topic modeling approach, which takes the document by term matrix as input, Lexical-LDA makes use of shallow lexical knowledge of dependency relations for topic modeling to achieve more effective clustering.

4. Assign the Rank for Sellers Based On Trusted Score

The CommTrust reputation profiles comprise dimension reputation scores and weights, as well as overall trust scores for ranking sellers. Our extensive experiments on eBay and Amazon data show that CommTrust can significantly reduce the strong positive bias in eBay and Amazon reputation systems, and solve the “all good reputation" problem and rank sellers effectively.

IX. ALGORITHM

Commtrust: Comments-Based Multi-Dimensional Trust Evaluation

The overall trust score T for a seller is the weighted aggregation of dimension trust scores for the seller,

$$T = \sum_{d=1}^{m} t_d * w_d$$

Where td and wd represent respectively the trust score and weight for dimension d (d = 1... m).

Calculating Trust Score and Weight for Dimensions:

Let S ={ X1, ..., Xn} be ‘n’ observations of binary positive and negative ratings, where y observations are positive ratings. S follows binomial distribution B (n, p). Following the Bayes rule, p can be estimated from observations and some prior probability assumption.

α and β are hyper-parameters expressing prior beliefs, the Bayes estimate of p is formed by linearly combining the mean α/(α+ β) from prior distribution and the mean y/n , as below

$$\hat{p} = (y + \alpha)/(n + \alpha + \beta)$$

The assumption of Beta distribution for the prior belief leads to reasonable trust evaluation. The Beta reputation system adopts constant settings of $\alpha = \beta = 1$ for above equation. We develop the approach further by introducing hyper-parameter settings for α and β to suit for a varying number of observed positive and negative ratings. It is preferable to have only one parameter for trust evaluation. With the prior belief of neutral tendency for trust, it can be assumed that $\alpha = \beta$.

Let $\alpha + \beta = m$,

Then $\alpha = \beta = 1/2 * m$.

The trust score for a dimension is thus defined as follows:

Given n positive (+1) and negative (-1) ratings towards dimension d, $n = |\{vd| \ vd = +1 \vee vd = -1\}|$, the trust score for d is:

$td = (|\{ vd| \ vd = +1\}| + 1/2 * m)/n + m$

Positive percentage score is treated as weight of the dimension d.

wd= no. of positive ratings/ (no. of positive ratings + no. of negative ratings).

Mining Feedback Comments for Dimension Ratings and Weights

If a comment expresses opinion towards dimensions then the dimension words and the opinion words should form some dependency relations. It has been reported that phrases formed by adjectives and nouns, and verbs and adverbs express subjectivity. Among the dependency relations expressing grammatical relationships, we select the relations that express the modifying relation between adjectives and nouns, and adverbs and verbs, as determined by the dependency relation parser.

Ratings from dimension expressions towards the head terms are identified by identifying the prior polarity of the modifier terms by SentiWordNet, a public opinion lexicon. The prior polarities of terms in SentiWordNet include positive, negative or neutral, which corresponds to the ratings of +1, -1 and 0. Negations of dimension expressions are identified by the Neg () relation of the dependency relation parser. When a negation relation is detected the prior polarity of the modifier term is inverted. Based on the trust scores, ranks will be assigned to the sellers.

X. RESULT

The proposed scheme provides better performance when compared to the existing solutions.

Evaluation Metrics

The ultimate goal of trust evaluation for e-commerce applications is to rank sellers and help users select trustworthy sellers to transact with. In this respect, in addition to absolute trust scores, relative rankings are more important for evaluating the performance of different trust models.

We employ metrics Rand index (RI) and Clustering Accuracy (Acc) to evaluate the performance of dimension clustering algorithms. RI measures both within-cluster and between-cluster agreement of two clustering algorithms. Acc measures the level of consistency between clusters produced by a clustering algorithm and the clusters by human annotation.

1. Evaluation of the Trust Model

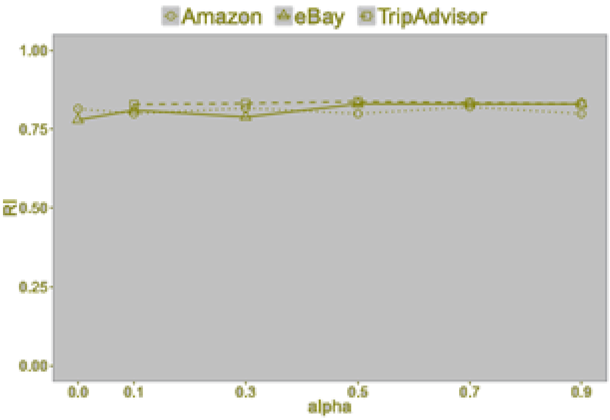
Fig. 5 depicts the dimensional trust profiles for three eBay sellers Seller 1, Seller 2 and Seller 3, where they have the same four dimensions, including shipping, cost/response, item and seller. For each seller, the upward bars represent trust scores for dimensions while the downward bars represent their weights. For example while having a high overall trust score of 0.9771, Seller 3 has a low dimension trust score of 0.9067 for the response dimension (Dimension 2). The figure clearly illustrates the variation of dimension trust for each seller horizontally and those across different sellers vertically. Such comprehensive trust profiles certainly can cater to users preferences for different dimensions and guide users in making informed decisions when choosing sellers.

Fig. 5 Dimension trust profiles by commtrust for sellers

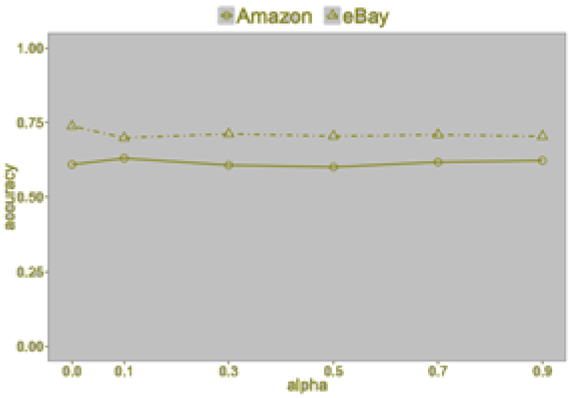
2. Evaluation of Lexical-LDA

We evaluate Lexical-LDA against standard LDA for clustering and against the human clustering result. As there are seven categories by human clustering, K = 7 for Lexical-LDA.

Fig. 6(a) plots the RI of Lexical-LDA at different settings of α. Note that the data point for α = 0 corresponds to the standard LDA. In addition to the eBay and Amazon datasets, to demonstrate the generality of our approach, the performance of Lexical-LDA on the Trip Advisor dataset is also plotted. For eBay and Amazon data, each plotted data point is the average for ten sellers. On eBay data, RI of Lexical-LDA hovers over 0.78 ~ 0.83, and Lexical-LDA significantly outperforms standard LDA for α> 0 except α = 0.3(p-value < 0.05, paired two-tail t-test). Comparable RI is observed on Trip Advisor and Amazon datasets. Our experiment results indicate that Lexical-LDA has steady performance across different domains.



(a)



(b)

Fig. 6 Evaluation of Lexical-LDA dimension clustering. (a) RI of Lexical-LDA. (b) Accuracy of Lexical-LDA.

Fig. 6(b) plots the accuracy of Lexical-LDA with different settings of α. As can be seen in the graph, accuracies hover over 0.70 ~ 0.74 on eBay data and 0.61 ~ 0.63 on Amazon data. There are not statistically significant differences in accuracies between Lexical-LDA with α>0 and standard LDA, on either Amazon or eBay datasets. However clustering accuracy only measures how automatic clustering matches the human clustering, rather than the coherence within clusters by clustering algorithms.

XI. CONCLUSION

The “all good reputation” problem is well known for the reputation management systems of popular e-commerce web sites like eBay and Amazon. The high reputation scores for sellers cannot effectively rank sellers and therefore cannot guide potential buyers to select trustworthy sellers to transact with. On the other hand, it is observed that although buyers may give high feedback ratings on transactions, they often express direct negative opinions on aspects of transactions in free text feedback comments. In this paper we have proposed to compute comprehensive multi-dimensional trust profiles for sellers by uncovering dimension ratings embedded in feedback comments. Extensive experiments on feedback comments for eBay and Amazon sellers demonstrate that our approach computes trust scores highly effective to distinguish and rank sellers.

We have proposed effective algorithms to compute dimension trust scores and dimension weights automatically via extracting aspect opinion expressions from feed-back comments and clustering them into dimensions. Our approach demonstrates the novel application of combining natural language processing with opinion mining and summarization techniques in trust evaluation for e-commerce applications.

XII. FUTURE WORK

Even though this proposed system provides efficient results in calculating trust profiles for sellers, this trust profiles are visible to only third party sellers. But users cannot view this trust profiles, based on the trust of third party sellers users will buy products.

This system can be enhanced to view the seller trust profiles to the users to transact with trustworthy sellers. This system cannot restrict the users in giving comments i.e. one can give fake comments, this system has scope to improve in identifying fake comments and restrict such users in giving comments.

REFERENCES

1. Paul Resnick and E. Friedman, “Reputation Systems: Facilitating Trust in Internet Interactions” Commun. ACM, vol. 43, no. 12, pp. 45-48, 2000.
2. Paul Resnick and Richard Zeckhauser, “Trust among Strangers in Internet Transactions: Empirical Analysis of eBay’s Reputation System”, vol. 11, no. 11, 2002.
3. John O’Donovan and Barry Smyth, “Extracting and Visualizing Trust Relationships from Online Auction Feedback Comments” , 2007.
4. Marie-Catherine de Marneffe, Bill McCartney, And Christopher D. Manning, “Generating Typed Dependency Parses from Phrase Structure Parses”, vol. 6, 2006.
5. Bo Pang and Lillian Lee, “Opinion mining and sentiment analysis”, vol. 2, 2008.
6. David M. Blei, Andrew Y. Ng and Michael I. Jordan, “Latent Dirichlet Allocation”, vol. 3, 2003.

7. Jordi Sabater and Carles Sierra, “REGRET: A reputation model for gregarious societies”, 2001.

8. Achim Rettinger, Matthias Nickles,Volker Tresp, “Statistical relational learning of trust”, vol. 82, 2011.

9. Paul Resnick, Richard Zeckhauser , John Swanson , Kate Lockwood, “The value of reputation on eBay: a controlled experiment”, vol. 9, 2006.

10. Audun Josang, Roslan Ismail, “The Beta Reputation System”, 2002.

11. Nathan Griffiths, “Task Delegation using Experience-Based Multi-Dimensional Trust”, 2005.

12. Steven Reece, Alex Rogers, Stephen Roberts, “Rumors and Reputation: Evaluating Multi-Dimensional Trust within a Decentralized Reputation System”, 2007.

13. Michael Gamon, “Sentiment classification on customer feedback data: noisy data, large feature vectors, and the role of linguistic analysis”, 2004.

14. Y. Hijikata, H. Ohno, Y. Kusumura, and S. Nishida, “Social summarization of text feedback for online auctions and interactive presentation of the summary”, vol. 20, 2007.

15. Zhongwu Zhai, Bing Liu, Hua Xu, Peifa Jia, “Constrained LDA for Grouping Product Features in Opinion Mining”, 2011.

16. Q. Mei, X. Ling, M. Wondra, H. Su, and C. Zhai, “Topic sentiment mixture: Modeling facets and opinions in weblogs”, 2007.

17. A. Mukherjee and B. Liu, “Aspect extraction through semi-supervised modeling”, vol. 1, 2012.

18. G. Heinrich, “Parameter estimation for text analysis”, 2005.

19. P. Thomas and D. Hawking, “Evaluation by comparing result sets in context,” in Proc. 15th ACM CIKM , Arlington, VA, USA, 2006, pp. 94–101.

20. A. Fahrni and M. Klenner, “Old wine or warm beer: Target-specific sentiment analysis of adjectives,” in Proc. Symp. Affective Language in Human Machine, AISB, Aberdeen, Scotland, 2008, pp. 60–63.

21. H. Wang, Y. Lu, and C. Zhai, “Latent aspect rating analysis on review text data: A rating regression approach,” in Proc. 16th ACM SIGKDD Int. Conf. KDD, New York, NY, USA, 2010, pp. 783–792.

22. A. Josang, R. Ismail, and C. Boyd, “A survey of trust and reputation systems for online service provision,” DSS, vol.43, no.2, pp. 618–644, 2007.

23. S. Brody and N. Elhadad, “An unsupervised aspect-sentiment model for online reviews,” inProc. HLT, Los Angeles, CA, USA, 2010, pp. 804–812.

24. P. D. Turney, “Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews,” in Proc. 40th ACL , Philadelphia, PA, USA, 2002, pp. 417–424.