IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 10, October 2016

Comm Trust by Mining Tweets, Feedbacks & Comments

Mrs. Vijaya Sayaji Chavan[1], Dr. M.Z. Shaikh[2]

Lecturer, Computer Engineering, BVIT, Kharghar, Navi Mumbai, India [1]

Principal, BVCOE, Kharghar, Navi Mumbai, India [2]

Abstract: Most of the e-commerce applications uses trust reputation trust models, and sellers' reputation trust scores are computed with the help of the users feedbacks and comments. The "all good reputation" problem is arises in todays systems as the reputation scores are high for sellers so it becomes difficult for buyers to trust the sellers. In this paper, depend on opinions of the buyers openly given in free feedback comments, We propose Comm Trust for trust evaluation by mining tweets, feedbacks, comments. Our main proposition is: (1) We propose a multidimensional trust for computing reputation scores from user feedback comments or tweets. (2) We propose an algorithm for mining feedback comments and opinion mining and by using topic modelling. Most experiments on twitter, eBay and Amazon data demonstrate that Comm Trust can effective for the "all good reputation" issue and rate sellers effectively. This paper gives a brief introduction to this lime lighted topic and explains few of its application, in our information world and also touch topic likes trust score, reputation trust. This paper will also include the algorithms based on topic modelling.

Keywords: E-Commerce, topic modelling, Feedback Mining, Trust Score, Comments, tweets.

I. INTRODUCTION

There has been a huge growth in e-commerce applications such as Amazon, Flipkart, Twitter. On this site there are many transactions on internet between sellers and buyers. Customers are attracted to E-commerce sites because of convenience and also availability of other buyers feedback on the purchased items related to different aspects [3]. There are many E-commerce websites takes buyers feedback, in the form of rating in the form of comments so that the other customer can review this comments. In this system the overall reputation trust scores for sellers are computed by aggregating feedback ratings [2].

Users provide these ratings on the basis of services they got from the seller. These rating are useful to new buyers. In E-commerce site like Ebay the reputation score for a seller is computed by aggregating buyer feedback ratings in the past 12 months [4]. They calculate this rating and is mostly average rating. [5]. we also refer this rating as trust. We can use these comments to check the trusts of e-commerce sellers. To calculate the Comment-based Multi-dimensional trust we use the positive and negative opinion of the other buyers. Our purpose is to provide a trust profiles for sellers that allows buyers to conduct their online shopping based on previous buyers.

"All good reputation" problem is the most noted issue with eBay [2], [6] where feedback ratings are over 99% positive on average. This guide buyer to select sellers. At eBay detailed seller ratings for sellers (DSRs) on four aspects of transactions, namely item as description, communication, delivery time, and delivery and handling charges. DSRs are aggregated rating scores on a 1- to 5-star scale. Still the strong positive rating is present and they are mostly 4.8 or 4.9 stars. Buyers gives positive feedback ratings, and also they express some disappointment and negativeness in free text feedback comments, often towards specific aspects of transactions. For example, comments like "The product was very good." Expresses positive opinion towards the product aspect, whereas the comment "Postage time was little extended but otherwise, nice service". This gives negative opinion towards the postage time but a positive opinion to the transaction. We propose Comment-based Multi-dimensional trust (Comm Trust) by mining e-commerce feedback comments.

In Comm Trust, extensive trust profiles are computed for sellers, It also computes dimension reputation scores and weights, and overall trust scores by aggregating dimension reputation scores. In Comm Trust, we propose an approach that combines dependency relation analysis [7], [8], natural language processing (NLP) and lexicon-based opinion mining techniques to extract aspect opinion expressions from feedback comments and identify their opinion orientations. To compute aggregated dimension ratings and weights we also propose Latent Dirichlet Allocation (LDA) algorithm and topic modelling technique. Comm Trust is used to reduce the strong positive bias in eBay and Amazon reputation systems, and solve the "all good reputation" problem and rank sellers effectively. we are going to make use of this kind of posts or tweets to help the companies by informing them about the fault in their free products so that their reputation will never go down. So the objective of this project is to get the

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 10, October 2016

real-time Twitter data, filter the data and analyse the data so that it can be reported to the companies if their free products does not work as expected.

II. RELATED WORK

The strong positive rating bias in the eBay reputation system is well mentioned in literature [2]–[6], but no effective solutions are provided. S. Ramchurn, D. Huynh, and N. Jennings has provided a extensive overview of trust models [9]. To compute the reliability of peers and assist buyers in their decision making Individual level trust models are used.

X. Wang, L. Liu In [11] has presented trustworthiness which is build using ratings. Customer recommendations, feedbacks provided by users. There are many Rating aggregation algorithms are used to build up trustworthiness for sellers
Feedback and comment analysis has been done by analysing feedback comments in e-commerce applications by focussing sentiment classification of feedback comments.

back. Lu. et al [3] has mentioned "rated aspect summary" from eBay feedback comments.

TABLE 1Some sample comments on eBay

No	Comment	eBay rating
1	Reasonable price and great service.	1
2	product arrived swiftly! Great seller.	1
3	good item. best seller of ebay	1
4	slow delivery, but seller was friendly.	1
5	Wrong item was sent but but delivery was on time.	1

Our paper is related to opinion mining, or sentiment analysis on free text documents. An extensive overview of the field is presented in [10]. There has been existing work on aspect opinion mining on product reviews.

In [12] frequent nouns and noun phrases are considered aspects for product reviews, and an opinion lexicon is developed to identify opinion orientations. To improve the aspect extraction accuracy some papers proposed to apply lexical knowledge patterns.

Some work groups uses topic modelling-based techniques and that have been developed to model opinions and aspects. Latent Semantic Analysis (pLSA) or Latent Dirichlet Allocation (LDA) algorithms are u differ in granularities.

There has been some recent work on computing aspect ratings from overall ratings in e-commerce feedback comments or reviews. Based on regression from overall ratings and the positive bias their aspect ratings and weights are computed.

III METHODOLOGY

COMMTRUST: FEEDBACK, COMMENTS BASED MULTI-DIMENSIONAL TRUST EVALUATION
We get feedback comments as a source where buyers express their opinions more freely. On eBay and Amazon most of a buyer gives a positive rating for a transaction, but still he leaves some comments with mixed opinions regarding different aspects of transactions in feedback comments.. Table 1 shows some sample comments, with their rating from eBay. For example for comment 4 , a buyer gave a positive feedback rating for a transaction, but left the following comment: "slow delivery". It means the buyer has negative opinion about delivery but overall positive feedback rating towards the transaction. This is salient aspects dimensions of e-commerce transactions. Comments-based trust evaluation is therefore multi-dimensional.

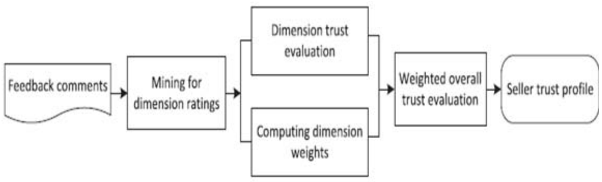

Fig. System Architecture

The overall trust score for a seller is the weighted aggregation of dimension trust scores for the seller,
Aspect opinion expressions and their associated ratings either 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 di-mension ratings.

MINING FEEDBACK COMMENTS FOR DIMENSION RATINGS AND WEIGHTS
To understand the grammatical relationships in sentences NLP tool is used in typed dependency relationship[8]. The sentence is represented as a set of dependency relations between pairs of words in the typed dependency relation, where content words are chosen as heads, and other related words depend on the heads. Sometimes comment like "Super quick shipping. Product was excellent. A great deal.
ALL 5 STAR." uses typed dependency relation parser. The comment consist four sentences, and the sentence "Super quick ship-ping." 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 shipping and quick mod-ifies 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

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 10, October 2016

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.

TABLE 1Dependency 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.

With the example, 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. 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.

CLUSTERING DIMENSION EXPRESSIONS INTO DIMENSIONS

We propose the Lexical-LDA algorithm to cluster aspect expressions into semantically coherent categories, called as dimensions. Different from the conventional topic modelling approach, which takes makes use of shallow lexical knowledge of dependency relations for topic modelling 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 or comment. In other words, it is very unlikely that the dimensions expressions extracted from the same comment are about the same topic.clustering problem is formulated under topic modelling 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 are dependency relations for dimension expressions in the form of (modifier, head) pairs or their negations, like (fast, shipping) or (not-good, seller

> 0 more positives votes,
= 0 same number of positive and negative votes,
< 0 more negative votes.

Extensive experiments on two e-commerce datasets and one hotel review datasets were conducted to evaluate various aspects of CommTrust, including the trust model and the the Lexical-LDA algorithm for clustering dimension expressions. The hotel review dataset is specifically used to demonstrate the generality of Lexical-LDA in domains other than e-commerce. 5.1 Datasets 180,788 feedback comments were crawled for ten eBay sellers on ebay.com, where two sellers were randomly selected for each of five categories on the "Shopby category" list on eBay.com, including Cameras & Photography, Computers & Tablets, Mobile Phones & Accessories, Baby, and Jewellery & Watches. Note that the sellers also sell products in other categories in addition to the listed categories. For evaluation of our trust model, the feedback profile for each seller were also extracted 2:
• The feedback score is the total number of positive ratings for a seller from past transactions. 2. pages.ebay.com/services/forum/feedback.html. TABLE 4 The Amazon dataset

Seller	Category	comments	Avg. rating
Seller 1	Electronics-Computer	4365	4.8
Seller 2	Electronics-Computer	4786	4.9
Seller 3	Electronics-Camera	3202	4.6
Seller 4	Electronics-Camera	8000	4.9
Seller 5	Electronics-Phone	4500	4.7
Seller 6	Electronics-Phone	3000	4.8
Seller 7	Jewelry-Ring	1000	4.9

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 5, Issue 10, October 2016

• The positive feedback percentage is calculated based on the total number of positive and negative feedback ratings for transactions in the last 12 months, that is #positive-ratings #positive-ratings+#negative-ratings. The Detailed seller ratings of a seller are five-star ratings on the following four aspects: Item as described (Item), Communication (Comm), Shipping time (Shipping) and Shipping and handling charges (Cost). The DSR profile shows a sellers average rating and the number of ratingss. Average ratings are computed on a rolling 12-month basis, and will only appear when at least ten ratings have been received. Details of the dataset are as Details of the dataset are as shown in Table 3.On Amazon, for a third-party seller, an average rating in the past 12 months is displayed, together with the total number of ratings. Each rating is associated with a short comment. 40,444 comments for ten third-party sellers with a large number of ratings were crawled from five categories, including Electronics-Computer, Electronics-Camera, Electronics-Phone Jewelry- Ring, and Baby-Tub and Baby-Diaper. Note that these sellers also sell products in other categories. A summary of the Amazon dataset is as shown in Table 4.
As shown in Tables ?? and ??, the strong positive bias is clearly demonstrated on the eBay and Amazon datasets. On the eBay dataset, the positive feedback percentage as well as DSR five-star rating scores have little dispersion and can hardly be used by itself to rank sellers. Similarly on the Amazon dataset, the average ratings for six sellers are 4.8 or 4.9. Evaluation of Lexical-LDA

Informal language expressions are widely used in feedback comments. Some pre-processing was first performed: Spelling correction was applied. Informal expressions like A+++ and thankx were replaced with AAA and thanks. The Stanford dependency relation parser was then applied to produce the dependency relation representation of comments and dimension expressions were extracted. The dimension expressions were then clustered to dimensions by the Lexical-LDA algorithm.

IV CONCLUSION

On the online shopping sites like e Bay and Amazon "All good Reputation" problem arises as there are always good comments and the feedbacks. Sometimes buyers gives good comments but with these again they provide some negative comments also .But we as a buyers review only the good comments and feedbacks. Because of this buyers cant not trust while shopping online So this paper is used to reduce this problem. We have proposed effective algorithms like LDA to compute dimension trust scores and dimension weights automatically via extracting aspect opinion expressions from feedback comments and clustering them into dimensions.

We have proposed a multi-dimensional trust evaluation model for computing trust profiles for sellers. Here we calculate dimension trust scores and dimension weights through extracting dimension ratings from feedback comments and aggregating with feedback rating. Based on this trust scores we can identify the reputable sellers from another seller that have had bad history with previous buyers.

ACKNOWLEDGMENT

We are very thankful to the reviewers for their valuable feedback which help improve the quality of this paper.

REFERENCES

[1] "Computing Multi-Dimensional Trust by Mining E-Commerce Feedback Comments" Xiuzhen Zhang, Lishan Cui, and Yan Wang, Senior Member, IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING VOL:26 NO:7 YEAR 2014
[2] P. Resnick, K. Kuwabara, R. Zeckhauser, and E. Friedman, “Reputation Systems: Facilitating Trust in Internet Interac-tions,” Communications of the ACM, vol. 43, pp. 45–48, 2000..
[3] Y. Lu, C. Zhai, and N. Sundaresan, “Rated aspect summariza-tion of short comments,” in Proc. the 18th Int. Conf. on WWW, 2009.
[4] P. Resnick and E. Friedman, “Reputation Systems: Facilitating Trust in Internet Interactions,” Communications of the ACM, vol. 43, pp. 45–48, 2000
[5] P. Thomas and D. Hawking, “Evaluation by comparing result Set in context," in Proc. 15th ACM CIKM, Arlington, VA, USA, 2006, pp. 94101.
[6] P. Resnick and R. Zeckhauser, “Trust among strangers in internet transactions: Empirical analysis of ebay’s reputation system,” The Economics of the Internet and E-commerce, 2002.
[7] M. De Marneffe, B. MacCartney, and C. Manning, “Generating typed dependency parses from phrase structure parses,” in Proc. LREC, vol. 6, 2006, pp. 449–454.
[8] M. De Marneffe and C. Manning, “The stanford typed de-pendencies representation,” in Proc. the workshop on Cross-Framework and Cross-Domain Parser Evaluation, 2008.
[9] S. Ramchurn, D. Huynh, and N. Jennings, “Trust in multi-agent systems,” The Knowledge Engineering Review, 2004.
[10] B. Liu, Sentiment analysis and opinion mining. Morgan & Claypool Publishers, 2012.
[11] M. De Marneffe, B. MacCartney, and C. Manning, “Generating typed dependency parses from phrase structure parses,” in Proc. LREC, vol. 6, 2006, pp. 449–454.
[12] M. Hu and B. Liu, “Mining and summarizing customer re-views,” in Proc. the fourth Int. Conf. on KDD, 2004, pp. 168–177.