Computing Multi-dimensional Trust By Mining E-Commerce Feedback Comments

[1]Akshay Zaware, [2]Tushar Kamble, [3]Tushar Kardak, [4]Vyankatesh Hardas
[5]Mrs.Deshmukh P. R.(Assistant Professor)

[1,2,3,4,5] *Government College Of Engineering & Research,Avasari(KD),Pune*

Abstract- *Nowdays we use many online shopping sites for shopping.In E-commerce applications various trust models are used based on reputation. Peoples purchase products from listed sellers based on their feedback ratings. All major online shopping sites encourage buyers to provide feedback, often in the form of ratings along with some textual comments to facilitate potential transactions. Sometime buyers gives star ratings from 1 to 5 stars for a particular product but if he/she is not satisfied about any other issue like about delivery, payment options etc. then he/she gives negative comment in text field. So it is not trustworthy for other customers to buy that product or not. Also at a time reading all the comments is not possible.*
To overcome all above issues we proposed model named as "Comm-Trust" based on opinions that user gives in free text comment field. This model includes :
1) Computation of dimension trust scores and dimension weights automatically by extracting dimension ratings from feedback comments. 2) An algorithm for dimension rating and weights combining natural language processing, opinion mining, topic modelling. Finally we can display percentages of +ve, -ve, and neutral comments in histogram form.

Key words: *Text mining, Topic Modelling, NLP, E-Commerce*

I. INTRODUCTION

Users are attracted to online shopping not only due to convenience in accessing information of items on sold, but also because of availability of other buyers feedback on their purchasing experience.

In E-commerce systems like e-Bay, Amazon(for third party sellers), reputation reporting systems are used where reputation trust scores for different sellers are computed by aggregating overall feedback ratings. The ratings indicate the ability of seller to provide satisfactory transactions in future which is beneficial to new buyers. Textual comments can provide detailed information that is not available in ratings. Even though buyers leave positive feedback ratings, they still express some disappointment or negativeness in free text comments. E.g. A comment like "Product were as I expected" expresses +ve opinion towards product dimension whereas the comment "Delivery was little slow but great service" expresses –ve opinion towards the delivery dimension.

Browsing through tens of pages of comments can be time consuming and to digest information is daunting task. So we proposed, a fine grained multi-dimension trust evaluation model for e-commerce application called as "Comm-Trust".We tackle all above issues by implementing two approaches. 1) Topic modelling approach is applied to develop lexical-LDA algorithm for grouping dimension rating extraction and trust computation.

Lexical LDA makes two types of lexical knowledge based on dependency relations for clustering dimension expressions into dimensions so as to produce meaningful cluster.

2) By using seed dimension words we can develop dimension rating mining which is knowledge based approach which includes meta-data, domain knowledge and general grammatical patterns for correctly identifying dimension rating expressions from feedback comments. Trust weights are computed by applying matrix factorization technique automatically.

II. RELATED WORK

J. O'Donovan developed to solve issue of unnatural high trust rating on Amazon, look to text comments and apply classification algorithm to capture keen information of negativity in that comments. Many times situation occurs where buyers are afraid to write negative comments which could damage their self-reputation. In such cases, +ve ratings are made but customer still voices some complaint or disappointment in free feedback comments.

Pooja A. Rangari Discussed that sentiment analysis often referred to as opinion mining is field of computer science which examines peoples sentiments, opinions evaluations, attitudes, issuesand their attributes. The large source of review feedback comments are forums, blogs, social networking sites, online shopping sites etc. Supervised and unsupervised machin learning approach is used. Opinion mining is supervised approach in which it is necessary to train classifier on training set before it is applied on test set. It combines techniques of natural language processing, information retrieval, text analytics and computational linguistics.

Yan Wang, Discussed that trust is major problem in e-service environments and e-commerce. Trust management mechanism is introduced in applications like eBay to provide precious information to customers prior to place orders and making payments. However most applications focus on determining general trustworthiness of individuals but not

provide transaction specific trust that involves factors related with future transactions, a new concept situational transaction trust, which binds existing trust data with new transactions. This can provide more correct trust information to customers and prevent typical attacks.

III. COMMTRUST

Buyers expresses their opinions more openly and honestly as a feedback comments.Analysis of these feedback comments indicates thateven if a customer gives a +ve ratings for a item, they might still gives comments of mixed opinions about different aspects of purchase. Hence comment based trust evaluation is multidimensional.

Comm-Trust uses Lexical LDA and DR-Mining approaches which uses lexical knowledge of dependency relation analysis to identify dimension ratings in feedback comments. Sentiment orientation is achieved by using SentiWordNet.

Defination A) -By calculating weighed aggregation of dimension trust scores we calculate overall trust score.

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

(a)

Where where wd and td are the weight and trust score for dimension d as in (a).

Defination B) -Given the positive (+1) and (-1) negative ratings towards dimension d, n = $|\{vd|vd = +1 \lor vd = -1\}|$,

$$t_d = \frac{|\{v_d|v_d = +1\}| + 1/2 * m}{n + m}$$
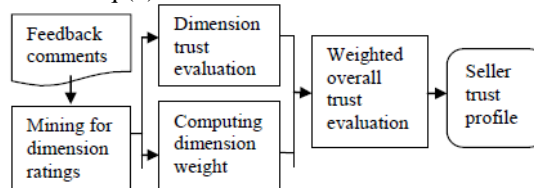
(b)

Eq.(b) is also called as m-estimates.

Fig.1 CommTrust Framework

As in Fig.1 First aspect expressions along with their associated ratings are calculated from feedback comments. Then by clustering aspect expressions into dimensions we calculate dimension trust scores with their associated weights.

Extracting aspect opinion expressions and identifying their associated ratings is based on dependency relation analysis.For clustering dimension expressions into dimensions and computing dimensions weights we use algorithm based on LDA.

A) *Extraction of aspect expressions and rating by Typed Dependency Analysis*

On e-commerce sites like eBay and Amazon, user feedback comments are generally short, andvery often are phrases or short sentences. Our analysis reveal that these short sentences andphrases can be accurately described using dependency relations in natural language parsing.Based on the parsing results of dependency relations, wecan identify dimension-rating patterns that describe dimensions and associated ratings.
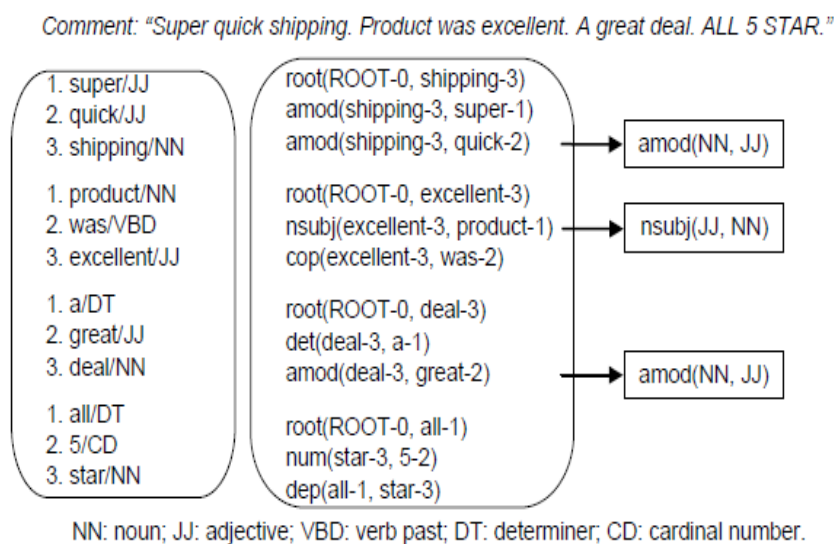

Fig.2 Typed dependency relation analysis example

To understand the grammatical relationships in sentences typed dependency relation is recent NLP tool.Sentences are represented in the form of a set of dependency relations between pairs of words as (head, dependent), where content words are chosen as heads, and other related words depend on the heads.

Ratings from dimension expressions towards the head terms are determined by identifying the prior polarity of the modifier by SentiWordNet, a public opinion lexicon. The prior polarities of the terms in SentiWordNet includes +ve, -ve , or neutral which corresponds to the ratings of +1, -1 & 0.

B) *Clustering Dimension Expressions into Dimensions*

Wederived the Lexical-LDA algorithm to group aspect opinion expressions into semantically consistentcategories, which is nothing but dimensions.The dimension expressions for a same modifier term or negation of a modifier term are generatedby a distribution of topics, and each topic is generated in turn by a distribution of headterms.

This formulation allows us to make use of the structured dependency relation representationsfrom the dependency relation parser for grouping. Input to Lexical-LDA aredependency relations for dimension expressions in the form of (modifier, head) pairs or theirnegations, like (fast, shipping) or (not-good, seller).

C) *Rating evaluation*

We apply a general opinion word lexicon SentiWordNet, which isa widely used public domain NLP resource to identify opinion polarities. When (modifier,head) pairs are grouped indimensions, the related modifier term express the opinionpriority of dimensions.SentiWordNet has a total of 155,181 words and each word (together with a POS tag) isannotated with positive, negative and objective scores that are summed to one. In applying SentiWordNet, words with a sum of positive and negative scores greater than or equal to 0.5 is consideredas express subjectivity, and if the positive score equals negative score for a subjective word,the word carries prior positive polarity.

V) GUI OF PROJECT
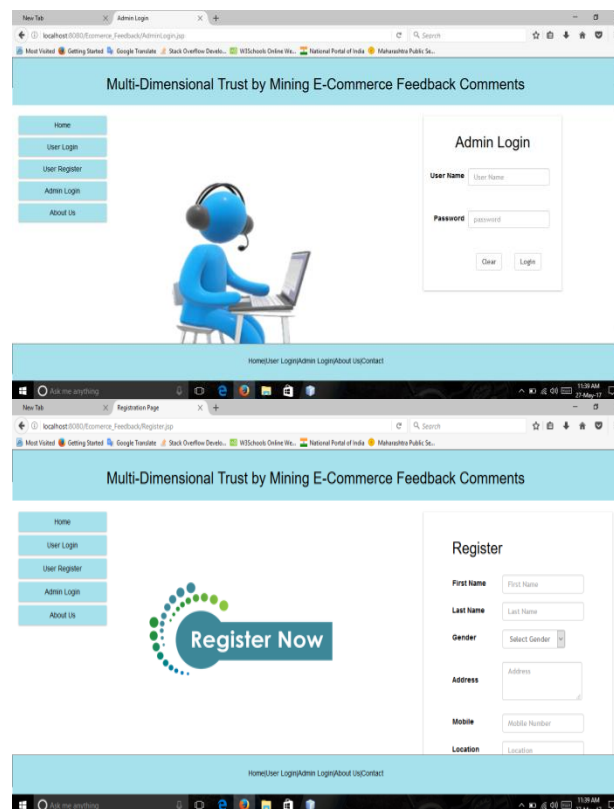
IV)REFERENCES

1. J. O'Donovan, B. Smyth, V. Evrim, and D. McLeod,"Extracting and visualizing trust relationships from online auction feedback comments," IJCAI'07, 2007, pp. 2826–2831

2. Pooja A. Rangari, Prof. Kalyani C. Waghmare "A Survey on Aspect based Opinion Mining" International Journal of Science, Engineering and Technology Research (IJSETR) Volume 4, no 3, March 2015.

3. Xiuzhen Zhang, Lishan Cui, and Yan Wang, "Comm-Trust: Computing Multi-Dimensional Trust by Mining E-Commerce Feedback Comment," vol. 25, no. 7, pp. 449454, July 2014.

4. P. Resnick and R. Zeckhauser, "Trust among strangers in internet transaction: Empirical analysis of E-Bays reputation system," Internet E-Commerce.