ISSN 2229-5518

# Computing Product Rating Using Real-Time Feedback Comments from E-Commerce Portal.

Gaurav Kamble, Rohit Athare, Abhishek Kumar, Neha Upadhyay.

**Abstract**— Different models are used widely used in e-commerce to rate the products on the portal, but the comments are aggregated to compute seller reputation. The "*All Good Reputation*" problem is very prominent in the current e-commerce rating systems. However, these scores are universal and it is difficult for potential buyers to buy from trustworthy sellers. In this study, based on comments that buyers' express in the feedback section, this paper proposes *CommTrust*, for evaluation by mining the feedback comments. The contribution include: (1) This paper proposes a multidimensional trust model for computing user feedback comments; (2) This paper also proposes an Algorithm for Mining Feedback Comments for Dimension Ratings, Combining techniques of NLP, LDA and PLSA. To the best of our knowledge, this study is the pioneer on trust evaluation by mining feedback comments.

**Index Terms**— Aspect Mining, Clustering, Commtrust, LDA, Lexcial-LDA, PLSA, SentiWordNet.

## 1 Introduction

IN the recent years, there are tremendous improvements in various technologies in every field. It is touching every apect of human life. One of the major changing trends now-a-days is e-commerce.

Accurate evaluation of products is crucial for the success of the e-commerce system. Reputation calculation systems are bring used in various e-commerce enterprises. Various methods are employed to calculate ratings & reputations. At eBay, the reputation score for a vendor is calculated based on a 12-month *positive percentage score*, out of the total positive and negative ratings.

"*All Good Reputation*" is a well-reported issue with many e-commerce systems, were feedbacks are 99% on the positive on average. Such strong bias are not helpful to the buyers to select a right product or vendor. At eBay, the system used is called Detailed Seller Ratings (DSRs) [1] on four aspects of transactions, *item as described, communication, postage time,* and *postage handling charges* on a scale of 1 to 5 stars.

Still strong bias feedback is present. One of the possible reason for the lack of negative feedback at e-commerce portal is that users leaving negative feedback comments can attract retaliatory negative ratings and feedback and damage their own reputation thus affecting potential customers.

Buyers may rate the product positively but they highlight some negative aspects of transitions in the feedback text. "*Postage was a little slow but otherwise, great product. Recommend highly.*" expresses negative opinion towards the postage but a positive opinion to general transaction and product. By analysing this hidden wealth of information in feedback comments we can uncover buyers' detailed embedded opinions towards the product, hence compute comprehensive reputation profiles for buyers and sellers.

This paper proposes a *Comment-based Multi-dimensional trust model (CommTrust)*, a multi-dimensional trust evaluation model by mining e-commerce feedback comments. With CommTrust, comprehensive profiles are computed for sellers and vendors, including reputation scores, also overall trust scores by combining dimension reputation scores of vendors. To the best of my knowledge, CommTrust is the first piece of work that computes multi-dimensional trust profiles automatically by mining feedback comments for e-commerce.[1]

In CommTrust, the study proposes an approach that combines dependency relation analysis, a tool in NLP and Lexicon-Based Opinion Mining Techniques to analysis aspect opinion from feedback comments and identify their orientations. This algorithm is called as Lexical-LDA. Unlike conventional methods for text-based documents, clustering is implemented on the dependency relation representations of aspect opinion and expressions. As a result Lexical-LDA makes use of the structures of aspect and opinion, as well as negation defined by relations to achieve more effective clustering of the relations. To address the positive bias dimension weights are computed directly by aggregating aspect opinion expressions rather than regression from overall ratings. The CommTrust reputation profiles comprise dimension reputation scores, as well as overall trust scores for ranking sellers. CommTrust can significantly reduce the strong positive bias comments in e-commerce reputation systems, and solve the "*All Good Reputation*" problem and rank sellers effectively.

## 2 Related Work

### 2.1 Review Stage

This falls into three main categories:
1) Computational approaches to reputation based trust evaluation 2) E-Commerce free text comments analysis and 3) Aspect extraction and summarisation on movie, product reviews and other forms of free text.[1]

### 2.2 Feedback comment analysis

There have been studies in the past on analysing feedback comments in e-commerce applications, though *comprehensive trust evaluation* and sentiment classification of feedback comments was never their focus. Comments and feedback are noisy and therefore analysing them is a challenging. In some missing aspect of feedback are deemed negative and models built from aspect ratings are used to classify comments into positive or negative feedback. In a technique for summarising feedback comments, aiming to filter out polite and respectful comments that do not provide real feedback. *Yue Lu, Universi-*

ISSN 2229-5518

*ty of Illinois* focuses on generating "rated aspect summary" from eBay feedback comments. Their statistical generative model is based regression on the overall transaction.[2]

### 2.3 Aspect Opinion Extraction

The work is related to sentiment analysis and opinion aspect mining on feedback documents. An aspect overview of the field is presented. There has been existing work on aspect and opinion mining on movie and product reviews.

In pervious works, common nouns and phrases used frequently are considered for product reviews and an opinion terminology is developed to identify acclimatization of comments.

Future proposed to apply expressed knowledge patterns to improve the accuracy of aspect extraction of feedback. Dependency relation parsing is used to extract aspect for free text. However, the existing systems do not group *opinion expressions* into clusters.

## 3. COMMTRUST: COMMENTS- BASED MULTI DIMENSIONAL TRUST EVALUATION

Feedback comments are viewed as a source where users can express their assessment, speculation and conclusion more honestly and openly. Analysis of feedback comments on various e-commence portals reveals that even if a buyer gives a positive rating for product, they might still leave comments of mixed opinions regarding different aspects of purchasing. For example, a buyer gave a positive rating for a purchase, but left the comment: "*Bad communication, will not buy again. Super slow postage,* ***but item as described***." The buyer has negative opinion towards the communication and postage aspects of the purchase, but overall positive feedback towards the purchase. This striking aspects dimension of ecommerce transactions. Comments-based trust evaluation is therefore multi-dimensional in nature.

### DEfINITION 3.1.

The overall trust score T for a vendor is the weighted collection of dimension trust scores for the vendor,

The definition of trust in by Audun Jøsang, the trust score on one of the dimension for a vendor is the probability that buyers expect the seller to conducting business on this dimension convincingly. The trust score for one dimension can be estimated from the number of observed positive and negative feedbacks towards that dimension.

## 4. MINING FEEDBACK COMMENTS FOR DIMENSION WEIGHT.

This approach is based on the dependency analysis to extracting aspect expressions from the feedback and identifying their associated ratings in the relation. This paper proposes an algorithm based on LDA for clustering of expressions into dimensions and computing dimension weights of the clusters.

### 4.1 Extracting aspect expressions and rating by typed dependency analysis

The typed dependency relation is a recent tool in NLP help us understand the grammatical relationships in a sentence. With typed dependency relation parsing, a sentence is represented as a set of association between pairs of words in the form of (head, dependent), where content words are chosen as head words, and other related words are chosen as depend on the heads. An example of analysing the comment "*Super quick postage. Product was excellent. Awesome deal. 5 STAR*." using Stanford typed dependency relation parser. The feedback comprises four sentences, and the sentence "*Super quick postage*." is represented as three dependency relations. *Postage* does not depend on any word and is at root level. The adjective modifier relations amod (postage -3, super 1) and amod (postage -3, quick 2) indicate that super modifies postage and quick modifies postage.

The number following each word (e.g., postage -3) indicates the position of this word in a sentence. Words are also annotated with their POS tags such as Verb (VB), Noun (NN), Adjective (JJ) and Adverb (RB). If a comment expresses opinion towards any aspect then the dimension words and the deliberation words should form some dependency relation. It has been studied that phrases formed by adjectives, nouns, verbs and adverbs express perspicacity. Among the dependency relations expressing semantic relationships, we select the relations that express the modifying relation between adjectives, nouns, adverbs and verbs as determined by the dependency relation expression parser. It is observed that with the modifying relations mostly the noun or verb indicate the destination concept under consideration whereas the adjective or adverb indicate opinion or aspects towards the destination concept. The modifying relations thus can be denoted as (modifier, head) pairs. Dependency relations adjective modifier amod (NN, JJ) and normal subjects nsubj (JJ, NN) suggest the (modifier, head) pairs along with the (super, quick, postage), (great, product) and (great, deal). We call these (modifier, head) pairs dimension expressions.

Ratings from dimension interpretation towards the head terms are recognized by identifying the former polarity of the modifier terms by SentiWordNet, lexical resource for opinion mining. The former polarities of terms in SentiWordNet include neutral, negative and positive analysis, which corresponds to the ratings of 0,-1, and +1. Negations of dimension expressions are identified by the Neg() relation of the dependency relation parser.

### 4.2 Clustering Dimension Expressions and Aspects

ISSN 2229-5518

### Into Dimensions

This study steers Lexical-LDA, to cluster side expressions into semantically logical division referred to as dimensions. Completely different from pattern input, Lexical-LDA makes use of flat lexical data of dependency relations for topic modelling to attain simpler clump of feedback comments. The study makes use of 2 classes of lexical data to supervise clump dimension expressions into dimensions to yield relevant clusters.

•Comments area unit short and so co-occurrence of head terms in comments is not very informative. We tend to instead use the concurrency of dimension expressions with reference to a same modifier across feedback that doubtless will offer additional substantive background for dimension expressions.

•We study that it is very rare that an equivalent dimension of e-commerce transactions is commented quite once. It is not possible that the size expressions computed from an equivalent comment are regarding an equivalent topic. With less lexical data of dependency relation illustration for dimension expressions, the cluster downside is outlined beneath modelling as follows:

The dimension expressions for a same modifier term or negation of a modifier term are generated by topic distribution and every topic is generated successively by a distribution of head terms. This production permits to create use of the structured dependency relation representations from the dependency relation computer programme for cluster. Input to Lexical-LDA are dependency relations for dimension expressions within the kind of (modifier, head) pairs or their negations, like (fast, postage) or (bad, vendor).



## 5. Mining Feedback Comments for Dimension Ratings and Weights

For clustering aspect expressions into semantically coherent categories, we propose the lexical LDA algorithm, which we call it as Dimension. To achieve effective clustering, we use shallow lexical knowledge of dependency relation for topic modelling using lexical LDA. Here, we use conventional topic modelling approach which takes the document as an input in terms of matrix form. There are two types of lexical knowledge, which is used to supervise clustering dimension expressions into dimension to produce meaningful clusters. The head terms in comment does not inform anything to the user, as comments are short. Instead using the co-occurrence of dimension expression with respect to the same modifier across comments. It is unlikely that the dimension expressions extracted from the same comment are about the same topic.

## 6. Conclusion

There are many e-commernce web portals with "*All Good Reputation"* problem. For sellers to be trust worthy, the ranking should be effective and should have high reputation scores. All the buyers may give high feedback rating on transaction but in their text feedback comments, they express their negative opinions. Hence, we have proposed an effective algorithm to compute dimension weight and trust scores automatically via extracting opinion expression from feedback comments and clustering them into dimensions

## Acknowledgment

We would wish a thank Prof.Archana Jadhav and Prof.Priyadarshi K(H.O.D Computer Dept.).

## References

[1] Computing Multi-Dimensional Trust by Mining E-Commerce Feedback Comments Xiuzhen Zhang, Lishan Cui, and Yan Wang, Senior Member, IEEE, 2007 , IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING VOL:26 NO:7 YEAR 2014.

[2] Probabilistic Latent Semantic Indexing Proceedings of the Twenty-Second Annual International SIGIR Conference on Research and Development in Information Retrieval Thomas Hofmann International Computer Science Institute, Berkeley, CA& EECS Department, CS Division, UC Berkeley.

[3] Old Wine or Warm Beer: Target-Specific Sentiment Analysis of Adjectives, Angela Fahrni & Manfred Klenner.

[4] http://sifaka.cs.uiuc.edu/~wang296/Data/index.html - Data Set for study and initial test cases.

[5] http://pages.ebay.com/help/feedback/allaboutfeedback.html

[6] B. Liu, Sentiment analysis and opinion mining. Morgan & Claypool Publishers, 2012.

[7] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," the Journal of machine Learning research, vol. 3, pp.993–1022, 2003.

[8] T. Hofmann, "Probabilistic latent semantic indexing," in Proc.the 22nd annual international ACM SIGIR conference on Research and development in information retrieval, 1999, pp. 50–57.

[9] Y. Lu, C. Zhai, and N. Sundaresan, "Rated aspect summarization of short comments," in Proc. the 18th Int. Conf. on WWW,2009.

[10] H. Wang, Y. Lu, and C. Zhai, "Latent aspect rating analysis without aspect keyword supervision," in Proc. the 17th ACM SIGKDD international conference on Knowledge discovery and datamining, 2011, pp. 618–626.

[11] "Latent aspect rating analysis on review text data: a rating regression approach," in Proc. the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, 2010, pp. 783–792.

[12] S. Ramchurn, D. Huynh, and N. Jennings, "Trust in multiagent systems," The Knowledge Engineering Review, 2004.

[13] B. Yu and M. P. Singh, "Distributed reputation management for electronic commerce," Computational Intelligence, 2002.

[14] M. Schillo, P. Funk, and M. Rovatsos, "Using trust for detecting deceptive agents in artificial societies," Applied Artificial Intelligence, vol. 14, no. 8, pp. 825–848, 2000.

[15] J. Sabater and C. Sierra, "Regret: reputation in gregarious societies," in Proc. the fifth international conference on Autonomous agents. ACM, 2001, pp. 194–195.

[16] A. Jøsang, R. Ismail, and C. Boyd, "A survey of trust and reputation systems for online service provision," Decision Support Systems, vol. 43, no. 2, pp. 618–644, 2007.

[17] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina, "The EigenTrust algorithm for reputation management in P2P networks," in Proc. the 12th Int. Conf. on WWW, 2003.

[18] A. Rettinger, M. Nickles, and V. Tresp, "Statistical relational learning of trust," Machine learning, vol. 82, pp. 191–209, 2011.

[19] X. Wang, L. Liu, and J. Su, "RLM: A general model for trust representation and aggregation," IEEE Transactions on Services Computing, vol. 5, no. 1, pp. 131–143, 2012.