

Literature Analysis on Reputation Models for Feedback in E-commerce

Oleena Thomas¹, Sumy Joseph²

Department of Computer Science, Mahatma Gandhi University, Amal Jyothi College of Engineering Kanjirapally,
Kottayam, Kerala, India^{1,2}

Abstract: Trust and reputation are decisive factors in internet service provision. The concept is to let the parties rate each other. Reputation models used in e-commerce have the main intention to guide the customers or buyers in a proper way so that they can choose between the different available trustworthy sellers. The models are meant to project the different sellers as per their services and quality. Models are a real need because most of the e-commerce systems have different qualities. But models differ between each other by the methods they opt to calculate the trustworthiness measure for different sellers. Based on such calculations the ranking of the sellers will also vary. Feedback mining method of trust calculation outperforms all other methods with the idea that users are free to express themselves in textual feedback comments. The method addresses the issue of positive bias which has become a major concern nowadays in the field of e-commerce.

Keywords: reputation system, feedback mining, trust score, positive bias, opinion extraction

I. INTRODUCTION

Trust evaluation for sellers, being the most important decisive factor in e-commerce is calculated in different ways in different models implemented in the e-commerce applications. Trust evaluation is done to rank the sellers properly so that the buyers could choose between them as per their services and qualities. Sellers can be evaluated in different ways using different models.

A well reported problem among the prominent e-commerce sellers is 'all good reputation' problem. This implies to the fact that most of the e-commerce sellers are graded high positive ranks and there exist almost no negative grades. This can be due to such sort of provision by the systems i.e., the users who are supposed to give feedbacks or grades are given only the provision of casting positive votes. In some systems there exist an approach that the users' or customers' reputation will be affected with negative votings they cast for the sellers. Due to these practices the negative votes thereby the real and genuine opinions about sellers cannot be reviewed or expressed. This in turn results in high positive bias.

There exist different models which include individual model, system model and reputation model. In individual model the buyers are guided to choose right sellers. In system model, the trustworthiness and behavior of the sellers are checked to ensure system security and to prevent fraudsters. In reputation model, the sellers are ranked based on some scores which are calculated using some statistical equations or by some other means like rating scores.

The reputation models can again be of different types. PeerTrust [3]-[4], positive feedback rating, average star rating, beta reputation, Kalman inference, Eigen trust model, and CommTrust models are few among them.

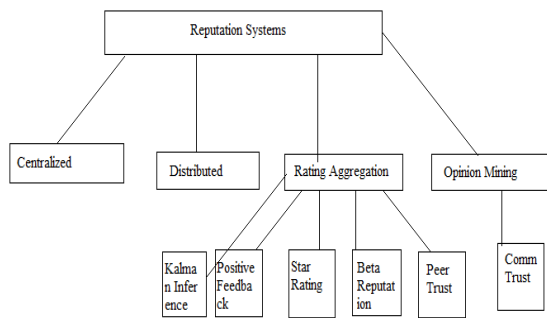
PeerTrust model performs evaluation based on the contextual factors of the transactions. The contextual factors for example, can be the price range \$100-\$200 if a customer intends to buy a camera for \$150. The positive feedback rating is based on the aggregation of only the positive votes given for a product by the buyer whereas average star rating is the aggregation of stars given by the buyers as in case of Amazon and eBay. Beta reputation system depends on beta distribution statistical measures. In Kalman inference the trust scores are calculated using covariance and confidence measures.

The Eigen trust model calculates global trust score from the local trust scores given to each peers. The CommTrust model uses feedback mining technique to calculate the trust scores. Opinion mining techniques are used in this to reduce positive bias present in most of the reputation models. The problem of 'all good reputation' is addressed by this model.

II. REPUTATION PROTOTYPE - A LITERATURE OVERVIEW

There have been rise of different models to attend the ranking of sellers from the user feedback. User feedback can be of any type like star rating, grading or through textual feedback. Depending on the feedback taken into consideration, models can be classified into many.

Buyers and sellers are considered as individuals in e-commerce system. Individual level trust models are used in examining the reliability of peers as well as to assist the customers in decision making. System level trust models are used in checking peer behaviours and to ensure the security of the system by preventing fraudsters. Reputation models aim to use public reputation profiles of peers to promote good behaviours and ensure security and reliability of open systems.



A. Centralized Reputation Systems

Information about the performance of a participant is gathered from other members in the community who have had transacted with this given participant as ratings in centralized reputation system. The central authority i.e. the reputation centre collects all the ratings and derives a reputation score for every participant. These scores are made publicly available. Participants can use these scores to choose between the other members to transact with.

A and B of Fig1. show transaction partners with a history of different transactions in the past. They consider transacting present. Ratings are given to the central authority after each transaction. The reputation centre collects all the ratings and continuously updates the scores of the participants and is made visible to all other members.

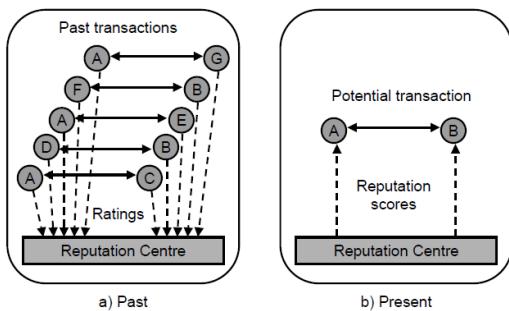


Fig1. Framework of centralized reputation system

B. Distributed Reputation Systems

There is no central location for submitting ratings or for obtaining scores of others. Distributed stores [8] are present where ratings are collected or each participant records their experience with every other member they transacted with. This information is passed only on request from relying parties. Relying parties should find this distributed stores or should get ratings from other members who had transacted with the target parties before committing transactions.

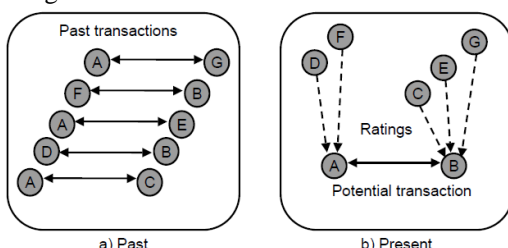


Fig2. Distributed reputation system

Typical distributed reputation system is shown in Fig2. The relying parties compute the reputation scores based on the received ratings. If the relying party has got any direct previous experience with the target party, then that experience can be taken as private information. This private information will have more weights than the received ratings.

C. Classification of Rating Aggregation Systems

In open environments trust relationship [12] is build using ratings. Ratings are also called recommendations, referrals and feedbacks. Rating aggregation algorithms are used to build up trust relationship. Many systems have been proposed using rating aggregation algorithms. Complex algorithms are not always cost effective and resistant to fake ratings. Review aggregator is one such system using rating aggregation algorithm [10]. It stores different reviews and makes use of these to support websites where the users can read the reviews; company databases can be created to evaluate their customers and many other activities. Each review would be assigned a numeric value based on the positive polarity expressed in that particular review and based on that an average assessment is made.

1) *Positive Feedback Systems*: Positive feedback system is a simple system which computes individual reputation scores for each seller. In this positive feedback percentage is calculated. This is based on the total number of positive or negative feedback ratings for transactions in a specified period of time.

Advantages and disadvantages:

The system can calculate reputation with much ease. Positive bias is being highly noticed in this type of approach.

2) *Star Rating System*: Star rating system is that which collects the feedback from the users in the form of star rating. The users are supposed to rate a transaction on the basis of the stars given. If the user gives more stars more will be the rating of the seller and vice versa. The individual reputation scores are calculated in this system for each seller.

Advantages and disadvantages:

The users can cast their voting easily through star rating. This cannot project an accurate seller ranking.

3) *Beta Reputation System*: Beta reputation system [7] depends on the statistical distribution for the ratings. The system is based on the beta probability density function to combine feedback and to derive reputation ratings.

Advantages and disadvantages:

The system is flexible and simple and is based on the theory of statistics. This is dependent on rating aggregation algorithm.

4) *Kalman Inference*: Kalman inference is another advanced model which computes trust score variance and confidence. Kalman filter operates recursively on streams of noisy input data to produce statistically optimal estimate of underlying system state. It uses weighted average. Weights are calculated from covariance which is a measure of estimated uncertainty of the prediction of the system's state.

Advantages and disadvantages:

The system can work on noisy inputs. The dimension weights are assumed to be given in this approach.

5) *PeerTrust System*: PeerTrust is a framework used in peer to peer systems. The contextual factors are taken into account for computing trust scores and weights for peers. The context of a transaction can be described by contextual transaction attributes. For instance, a buyer plans to buy a ‘Canon EOS T3i Digital Camera’ at the price ‘\$700’ from a seller. The customer will be concerned with the trustworthiness of the seller selling ‘Digital Cameras’ in a price range of ‘\$500-\$900’ i.e., a query with respect to a higher layer in hierarchical product category with a specified price range and time range (e.g. in the last one week). Thus the contextual factors includes transaction item, transaction amount and transaction time. Transaction item refers to the product in traded in a transaction, the properties like product qualities, product categories of which determine the nature of the transaction. Transaction amount refers to sum of prices of all products in a transaction. Larger the amount more is the chance for fraud to happen. A transaction with multiple transaction items are taken as several transactions with one item each. Transaction time is the time when a transaction happens. Transaction time has a specific feature in trust computation. Any query on temporal dimension should start from a previous point (e.g. one week ago) and end at present time.

Table I shows the transaction information of a seller in eBay [5]. With a contextual transaction trust query, the computation time for trust value of a seller at a specific product category in a price range and time period and the trust value of a seller in specific price range.

Advantages and disadvantages:

This approach uses a bit large amount of data space as well as computation time. The contextual factors are not flexible because the factors are chosen while the system is designed. The users are not a given a provision to choose the aspects of transactions by themselves and therefore they cannot express on different aspects they intend to convey. As a result of that, the exact ranking of the sellers cannot be ensured.

6) *EigenTrust*: EigenTrust algorithm [6] uses rating matrix representation for local trust scores and computes global ratings for peers from the rating matrix. This is a reputation management algorithm for peer to peer networks. Algorithm provides each peer in the network with a unique global trust value based on peer’s history of uploads and reduce the number of unauthenticated files in peer to peer network.

If peer i trust peer j , then all peers trusted by j can also be trusted. i calculates the local trust value for all peers that have provided it with authenticate or fake downloads based on the satisfactory or unsatisfactory transactions that it had.

TABLE I
 TRANSACTION INFORMATION OF A REAL SELLER AT EBAY

Seller id	Transaction item	Price	Transaction time	Subcategory	Category	Rating
1001	Canon EOS Rebel T3i Digital Camera	693.75	2011-09-16T15:10:45.000Z	Digital Cameras	Cameras & Photo	+1
1001	Canon Metal Neck Strap	3.99	2011-09-16T16:00:05.000Z	Camera& Photo Accessories	Cameras & Photo	+1
1001	B+W 52mm Ultraviolet (UV) Filter	22.0	2011-09-16T16:23:26.000Z	Lenses & Filters	Cameras & Photo	+1
...

Advantages and disadvantages:

Authentication of files is ensured. The system assumes that feedback ratings are given already and the aggregation algorithms are given priority.

D. *Opinion Mining Systems*

Opinion mining is also called sentiment analysis. It refers to natural language processing (NLP) and computational linguistics which identify and extract subjective information [9] from the source. This also involves text analysis. The opinion mining targets to determine the polarity of a document with respect to some context. It also targets in determining different opinions expressed by different authors about some topics. The opinions can be of different types for the same topic. The accuracy in each opinion can be measured from human judgment activities. Accuracy is measured using precision or recall functions.

The CommTrust [2] system consists of a multidimensional trust model for computing reputation scores from user feedback comments. The system is dependent on the idea that users will express themselves freely in textual feedbacks about different aspects of the transactions. Aspect opinion expressions and their corresponding ratings are extracted from the feedback comments of the users. Then the dimension trust scores as well as the weights are computed by clustering aspect expressions into dimensions and aggregating dimension ratings. The trust score for a dimension is calculated from the number of observed positive and negative ratings towards a dimension.

Fig.3 shows the CommTrust framework. The aspect opinion extraction is done using typed dependency analysis notion. The parser [11] used is Stanford parser. The dependency relation between words in a sentence is exploited to form a pair (modifier, head) which is given as the input to the clustering algorithm. The head terms are content words in a sentence while the modifiers are related words. The ratings are done using SentiWordNet, a public opinion lexicon. The prior polarities of terms are considered here. The clustering [1] of dimension expressions into dimension is dependent on lexical knowledge. This is done by extending the Latent Dirichlet Allocation algorithm with Gibb’s sampling inference. Clusters so formed relates with the manual clustering by a satisfactory amount. The CommTrust reputation profiles have dimension reputation scores and weights as well as

overall trust scores for sellers. Thus the sellers can be ranked efficiently as if ranking is done manually. The correlation between natural (manual) clustering and CommTrust clustering [13] was checked using Kendall's τ measure. The correlation was found positive i.e. both manual and CommTrust rankings were found equivalent.

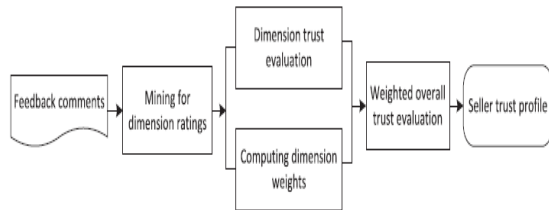


Fig.3. Framework of CommTrust

The CommTrust system uses SentiWordNet general lexicon to decide the prior orientation of the modifier terms. Being a general lexicon some word polarity annotation would not be found accurate in e-commerce applications. This at times can result in wrong polarity calculation of feedback comments and can thus affect the ranking of the sellers.

There have been works on aspect opinion mining on product as well as movie reviews. Frequent nouns and noun phrases are considered aspects for product reviews and opinion orientations are identified by opinion lexicon. Lexical knowledge patterns are also used to improve aspect extraction accuracy. Dependency relation parsing is used to mine aspect opinions for movie reviews. But these works do not group aspect opinion expressions into clusters.

Assuming aspect opinion expressions are given already, works have been there which group aspects into clusters. Semi-supervised algorithm based system extracts aspects and groups them into meaningful clusters as supervised by user input. Unsupervised topic modeling technique based systems model opinions and aspects together based on Latent Dirichlet Allocation or probabilistic Latent Semantic Analysis. They differ in how aspects and opinions interact.

All these systems make use of unigram representation of documents and they are not based on lexical knowledge. Aspect ratings in some works are computed from overall ratings in e-commerce feedback comments. Such aspect ratings and weights are calculated from overall ratings using regression methods. The problem of positive bias is left unaddressed.

Advantages and disadvantages:

The system mines the user textual feedback and the opinion expressions are processed to rank the sellers which model the reality more closely. Human clustering and this automated clustering hold much similarity. The general lexicon parser used in the system cannot process all the expressions.

III. OBSERVATIONS CONCLUDED

The reputation models like positive feedback system, star rating system, beta reputation system, Kalman inference system, PeerTrust and EigenTrust do not support opinion mining. But CommTrust system uses opinion mining method. Feedback ratings are assumed given in all systems except in CommTrust. In that the feedback ratings are calculated using SentiWordNet polarity distribution. Positive bias is found lowest in CommTrust whereas it is high in positive feedback system and star rating system and medium in PeerTrust, EigenTrust, Beta reputation and Kalman inference systems. All the systems other than CommTrust use aggregation algorithm. The dimension weights are either assumed given or not used in all systems other than CommTrust. In CommTrust dimension weights are calculated using dimension ratings.

The Table II shows the comparison among different systems.

Reputation models	Opinion mining	Attributes taken into Feedback	consideration for	evaluation Use of	Dimension Weights
		Ratings	Positive Bias	Aggregation Algorithm	
Positive feedback	No	Assumed given	High	Yes	No
Star rating	No	Assumed given	High	Yes	No
Beta reputation	No	Assumed given	Medium	Yes	Assumed given
Kalman Inference	No	Assumed given	Medium	Yes	Assumed given
PeerTrust	No	Assumed given	Medium	Yes	Assumed given
EigenTrust	No	Assumed given	Medium	Yes	Assumed given
CommTrust	Yes	Calculated using SentiWordNet polarity	Low	No	Calculated using dimension ratings

IV. CONCLUSION

The reputation systems used in commercial and online applications are prone to vulnerabilities. Thus the reliability is being questioned. When the area of e-commerce is taken into consideration, the sellers need to be ranked accurately so that the customers could find it easy to choose between trustworthy sellers in e-commerce applications. This ranking can be done with the help of the feedback given by the buyers. There are different models to put forward the reputation of the sellers. But the methods adopted by them in reputation score calculation are different. Depending on such methods the rankings given to each seller also vary. The ranking which relate more closely to the manual ranking is the most effective and efficient method i.e. if the correlation between manual and automated rankings is strong enough, then it can be concluded that the corresponding automated ranking is much efficient and effective in ranking sellers and can be used widely as reputation systems in e-commerce applications.

REFERENCES

- [1] Jiawei Han and Micheline Kamber, *Data Mining: Concepts and Techniques*, 2nd ed., Morgan Kaufmann, San Francisco, CA, 2006.
- [2] Xiuzhen Zhang, Lishan Cui and Yan Wang, "CommTrust: Computing Multi-dimensional Trust by Mining E-Commerce Feedback Comments," *IEEE Transactions On Knowledge And Data Engineering*, Vol. 26, No. 7, July 2014.
- [3] Y. Zhang and Y. Fang, "A fine-grained reputation system for reliable service selection in peer-to-peer networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 18, no. 8, pp. 1134-1145, Aug. 2007.
- [4] L.Xiong and L. Liu, "Peertrust: Supporting reputation-based trust for peer-to-peer electronic communities," *IEEE Trans. Knowl. Data Eng.*, vol. 16, no. 7, pp. 843-857, Jul. 2004.
- [5] H. Zhang, Y. Wang, and X. Zhang, "Efficient contextual transaction trust computation in e-commerce environments," in *Proc. 11th IEEE TrustCom*, Liverpool, UK, 2012.
- [6] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina, "The EigenTrust algorithm for reputation management in P2P networks," in *Proc. 12th Int. Conf. WWW*, Budapest, Hungary, 2003.
- [7] A. Josang and R. Ismail, "The beta reputation system," in *Proc. 15th BLED Electron. Commerce Conf.*, 2002, pp. 41-55.
- [8] S. Reece, A. Rogers, S. Roberts, and N. Jennings, "Rumours and reputation: Evaluating multi-dimensional trust within a decentralised reputation system," in *Proc. 6th AAMAS*, Honolulu, HI, USA, 2007, pp. 165172.
- [9] G. Qiu, B. Liu, J. Bu, and C. Chen, "Opinion word expansion and target extraction through double propagation," *Comput. Linguist*, vol. 37, no. 1, pp. 927, 2011.
- [10] P. Thomas and D. Hawking, "Evaluation by comparing result sets in context," in *Proc. 15th ACM CIKM*, Arlington, VA, USA, 2006, pp. 94101. Department
- [11] M. De Marne and C. Manning, "The Stanford typed dependencies representation," in *Proc. CrossParser*, Stroudsburg, PA, USA, 2008.
- [12] X. Wang, L. Liu, and J. Su, "RLM: A general model for trust representation and aggregation," *IEEE Trans. Serv. Comput.*, vol. 5, no. 1, pp. 131143, Jan-Mar, 2012.
- [13] Z. Zhai, B. Liu, H. Xu, and P. Jia, "Constrained LDA for grouping product features in opinion mining," in *Proc. 15th PAKDD*, Shenzhen, China, 2011, pp. 448459