

A Survey on CommTrust: Computing Multi-Dimensional Trust by Mining E-Commerce Feedback Comments

Swetha Sam¹, Vani V Prakash²

¹ PG Scholar, Dept. of Computer Science and Engineering, Sree Buddha College of Engineering for Women, Ayathil, Elavumthitta, Kerala

² Assistant Professor, Dept. of Computer Science and Engineering,
Sree Buddha College of Engineering for Women, Ayathil, Elavumthitta, Kerala

Abstract: Reputation-based trust models are widely used in e-commerce applications, and feedback ratings are aggregated to compute sellers' reputation trust scores. The "all good reputation" problem, however, is prevalent in current reputation systems—reputation scores are universally high for sellers and it is difficult for potential buyers to select trustworthy sellers. In this paper, based on the observation that buyers often express opinions openly in free text feedback comments, a survey on CommTrust for trust evaluation by mining feedback comments is done. Main contributions include: 1) a multidimensional trust model for computing reputation scores from user feedback comments; and 2) an algorithm for mining feedback comments for dimension ratings and weights, combining techniques of natural language processing, opinion mining, and topic modeling. Extensive experiments on eBay and Amazon data demonstrate that CommTrust can effectively address the "all good reputation" issue and rank sellers effectively.

Keywords: Electronic commerce, text mining, feedback. Commtrust

1. Introduction

Feedback comments can be viewed as a source where buyers express their opinions more honestly and openly. 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.

According to P.Resnick and K.Kuwabara [2], role of reputation system in online auction play a very crucial role. Many auction users refuse to provide negative feedback to others in order to maintain their reputation. Main problem with an online auction is that buyers do not have the ability to directly inspect the products they are buying. They faces "Lemons" market which means that a high amount of uncertainty about the quality of information is created. This issue of uncertainty is addressed using a concept commitment. It means how a seller promises to sell his goods or services. Reputation is defined as an overall quality or character seen by people in general. Increase in reputation will lead to high price. Only information a buyer has is an individual reputation rating. They also says that retaliatory feedback is the idea of leaving a negative rating for a transaction partner. This causes a partner to leave a negative rating. Retaliatory feedback is the reason behind high level of positive feedback on eBay. It occurs only when one party in a transaction believes that the other party will leave them a negative feedback if they do the same. Retaliator feedback problem is classified in to four categories.

- First category includes the problem of non-paying winner bidders threatening to leave negative feedback for sellers who leave them a negative for non-payment.
- Second includes the problem of who should leave feedback first after a completed transaction. Buyers believe

that seller should leave feedback first and seller believe that buyer should do it.

- Third includes the advice on whether and how to leave negative feedback.
- Last category includes the problem of feedback extortion which occurs when one party to a transaction demands something from the other party and uses the threat of negative feedback as part of the demand.

Feedback escrow is a very simple solution to the retaliatory feedback problem in which parties to a transaction would be allowed to leave feedback only for a specified interval of time. Sometimes there will be a time limit or an expiry period. All feedbacks will be kept in secret until both parties have left their feedback or until time has expired. This solution can be used in first, second and third category but feedback extortion cannot be eliminated using above mentioned solution. Thus an important role is played by the reputation system in providing a level of trust between sellers and buyers. They proposed, a class of trust models which aims at using public reputation profiles of peers are called reputation based trust models. They are used to promote good behaviours and ensure security and reliability of open systems. They are widely used in ecommerce systems, peer to peer networks and multi-agent systems. Simple positive feedback percentage or average of star ratings are used in rating aggregation algorithms for computing individual reputation scores.

B.Smyth et al. [3] says that throughout the world their exists an increased access to online markets. Some of the main problems affecting the reputation of eBay systems are observed by Smyth et al. They are as follows:

- Reputation of buyers matters less because they hold goods until they are paid.
- First comment is always important. Feedback can be

Volume 5 Issue 7, July 2016

www.ijsr.net

Licensed Under Creative Commons Attribution CC BY

affected by the first comment.

- Real name of the person writing the feedback may not be revealed sometimes creating an anonymity.
- Sometimes user can choose not to display feedback comments.
- Feedback cannot be written by unpaid item buyers.
- Users can agree to mutually with feedbacks.
- Negative feedbacks may be hidden from new users.

Processing of free text comments requires four main concepts including morphology, syntax checking, deciding on a semantic meaning, taking some action based on that meaning. A classification algorithm, which can extract both context-specific and personal trust, called an "Auction rule" algorithm is proposed. Its main goal is to correctly classify online auction comments into positive or negative according to a threshold. Seven features by which algorithm compute granular trust score are:

- Item - The quality/condition of the product being bought or sold.
- Person - The person the user makes the transaction with.
- Cost - Cost of item, cost of shipping, hidden costs etc.
- Shipping - Delivery of the item, security, time etc.
- Response - Communication with the other party, emails, feedback comment responses.
- Packaging - The packaging quality/condition of the item.
- Payment - how the payment will be made to the seller, or back to buyer for return.
- Transaction - the overall transaction quality.

This technique helps to compute a personal trust score between individual users. An algorithm for extracting personalized and contextual trust from the wealth of free text comments on online auction sites is proposed. This algorithm operates on the assumption that online auction transactions can be categorized into a relatively small set of features.

C.Manning et.al. [4] proposed a system for extracting typed dependency parses of English sentences automatically. Structure of sentences can be represented in different ways namely typed dependencies and phrase structures. Nesting of multiword constituents is represented by phrase structure parse whereas dependency parse represent dependencies between individual words. The proposed technique for producing typed dependencies is based on rules or patterns. This method requires appropriate rules for each language and tree bank representation. The two phases of generating typed dependencies are: dependency extraction and dependency typing. Dependency extraction is simple in which a sentence is parsed with a phrase structure grammar parser. Stanford parser is used, which helps in identifying head for each sentence. It aims at producing a representation closer to the semantics of the sentence.

According to A. Jøsang and R. Ismail,[5] a beta reputation system is based on using beta probability density function to combine feedback and derive reputation ratings. Its advantages are flexibility and simplicity. Its foundation is on theory of statistics. In ecommerce it is difficult to rely on legal procedures in case of any disagreement of contract; since cost is higher than contractual value. A reputation

system gathers, distributes, and aggregates feedback about participants behavior. This scheme is first introduced in online auction sites like eBay.com. Capturing and distributing of feedback is usually a costly approach. A reputation system has many challenges. If an entity changes its name, feedback can get erased and a dishonest participant can use this to start fresh every time it builds up a bad reputation. People may not bother to provide feedback at all, negative feedback can be difficult to elicit, and it is difficult to ensure that feedback is honest.

Two fundamental aspects are considered:

- 1) An engine that calculates the value of the users' reputation ratings from various inputs including feedback from other users.
- 2) A propagation mechanism that allows entities to obtain reputation values when required.

In the centralized approach, reputation values are stored in a central server, and whenever there is a need, users forward their query to the central server for the reputation value. One example is eBay. In the decentralized approach, everybody keeps and manages reputation of other people themselves. Whenever there is a need, users can ask others for the required reputation values.

A.Jøsang and R. Ismail proposed a new reputation engine called the beta reputation system which is based on the beta probability density function which has a firm basis in theory of statistics. Posteriori probability of binary events can be represented as Beta distributions.

$$\text{Beta}(p|\alpha, \beta) = \frac{\Gamma(\alpha+\beta) \cdot p^{\alpha-1} \cdot (1-p)^{\beta-1}}{\Gamma(\alpha) \cdot \Gamma(\beta)}$$

A reputation rating is interpreted as measure of reputation or an indication of how a particular agent is expected to behave in future transactions. A reputation system must be able to combine feedback from multiple sources.

According to S. Ramchurn et.al [6], trust is a belief an agent has that the other party will do what it says it will or reciprocate, given an opportunity to defect to get higher payoffs. Two main approaches to trust in multi-agent systems are:

- Firstly, to allow agents to trust each other.
- Enable agents to calculate amount of trust they can place in their interaction partners.

High degree of trust in agent means that there is a high chance that it can be chosen as an interaction partner. Low degree of trust causes not to be selected. Trust models aims to guide an agents decision making in deciding on how, when and who to interact with. With respect to designing agents and open multi-agent systems, trust can be conceptualized in the two ways:

- Individual level trust where an agent has some beliefs about the honesty or reciprocative nature of its interaction partners. It is the view point of an agent in an open environment.
- System level trust where the actors in the system are forced to be trustworthy by the rules of encounter that

regulate the system.

These two are complementary to each other. Protocols aim to ensure the trustworthiness of agents at the system level, they cannot always achieve this objective without some loss in efficiency, and, in such cases, trust models at the individual level are important in guiding an agent's decision making. Similarly, where trust models at the individual level cannot cope with the overwhelming uncertainty in the environment, system-level trust models, through certain mechanisms, aim to constrain the interaction and reduce this uncertainty.

According to Xiuzhen Zhang, Lishan Cui, and Yan Wang [1], in a commtrust framework, aspect opinion expressions and their associated ratings are extracted from feedback comments. Dimension trust score along with their weights are computed by clustering aspect expressions in to dimensions and aggregating the dimension rating. In order to find or extract aspect opinion expression and to identify their associated ratings an approach based on typed dependency analysis is used. Also lexical LDA is used for clustering dimension expressions. Normally feedback comments consist of informal language expressions. These feedback comments with informal language expressions need to be preprocessed. Spelling correction need to be applied to make these informal expressions to formal expressions. For example, informal expressions like gud, tanx were replaced with good, thanks. Dependency relation representation of comments and dimension expressions were extracted using Stanford dependency relation parser. Lexical LDA algorithm is used to cluster dimension expressions to dimensions. Typed dependency analysis is used to derive aspect opinion expression. Typed dependency analysis is a natural language processing (NLP) tool. This tool is used to derive the grammatical relation between sentences. The output of this NLP tool is a set of dependency relations between word pairs in form of (head, dependent). It is a dimension expression where head contains the content words and dependent contains other related words depend on heads. In our work, we use a Stanford typed dependency relation parser.

The dimensions are then clustered using lexical LDA algorithm. Lexical LDA using mallet topic modelling toolkit is used to group dimensions according to conceptual meaning. The overall trust score T for a seller is given by

$$T = \sum_{d=1}^m td * wd$$

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

Trust score on a dimension is the probability that buyers expect the sellers to carry out transactions on this dimension satisfactorily. It can be found from the number of observed positive and negative rating towards the dimension.

References

- [1] Xiuzhen Zhang, Lishan Cui, and Yan Wang, Senior Member, IEEE "CommTrust: Computing MultiDimensional Trust by Mining ECommerce Feedback Comments" IEEE transactions on knowledge and data Engineering, vol. 26, no. 7, July 2014

- [2] P. Resnick, K. Kuwabara, R. Zeckhauser, and E. Friedman, "Reputation systems: Facilitating trust in internet interactions," Commun. ACM, vol. 43, no. 12, pp. 45–48, 2000.
- [3] J. O'Donovan, B. Smyth, V. Evrim, and D. McLeod, "Extracting and visualizing trust relationships from online auction feedback comments," in Proc. IJCAI, San Francisco, CA, USA, 2007, pp. 2826–2831.
- [4] 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.
- [5] A. Jøsang and R. Ismail, "The beta reputation system," in Proc. 15th BLED Electron. Commerce Conf., 2002, pp. 41–55.
- [6] S. Ramchurn, D. Huynh, and N. Jennings, "Trust in multi-agent systems," Knowl. Eng. Rev., Vol. 19, no. 1, pp. 1–25, 2004.

Author Profile

Swetha Sam received the B.Tech degree in Computer Science and Engineering from Cochin University in 2013. She is currently pursuing second year M.Tech in Computer Science Engineering at Sree Buddha College of Engineering for Women.

Vani V Prakash. received the B.Tech degree in Computer Science and Engineering from Cochin University in 2008 followed by M.Tech. from Amrita University in 2011. She is currently working as Assistant Professor in Computer Science and Engineering Department, Sree Buddha College of Engineering for Women.