Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 8, August : 2023

ANALYZING EMOTIONAL SIMILARITY IN ONLINE STORE REVIEWS TO BUILD USER TRUST IN MINING

#1JELLA ANUHYA, ***Department of MCA,***

#2B.ANVESH KUMAR, ***Assistant Professor,***

#3Dr.V.BAPUJI, ***Associate Professor& HOD,***

Department of Master of Computer Applications,

VAAGESWARI COLLEGE OF ENGINEERING, KARIMNAGAR, TELANGANA.

Abstract- - Electronic commerce is the activity of exchanging goods and services over a computer network. Aside from purchasing and selling, many people utilize the internet for research purposes, such as seeing what's new in the market or comparing prices before making a purchase. E-commerce platforms are generally seen as valuable resources that provide users with an experience, feelings, and desire in purchasing things based on consumer evaluations. This form of data includes consumer opinions on products that can show interest, attitudes, and expressions. Several research theories suggest that people who have similar sentiments toward similar topics are more likely to trust one another. We suggest in this paper that asking and accepting recommendations and feelings in e-commerce networks is a form of reciprocal trust. A scientific observer studied consumers while they were shopping. To explore user trust and similarity, a sentiment similarity analysis approach oriented toward E-commerce system reviews is provided. In essence, trust may be separated into two types: direct trust and trust propagation, which results in a trust connection between two people. We present an entity-sentiment word pair mining approach for extracting similarity features. Sentiment similarity is used to calculate the direct trust degree. The transitivity feature is utilized to compute the trust spread. The proposed trust model is used to calculate the shortest path and to present an improved shortest path algorithm to determine the propagation trust relationship between users. A large dataset of E-commerce reviews is gathered to evaluate the efficacy of the algorithms and the viability of the models. Sentiment similarity analysis, according to the experimental data, can be a beneficial tool for generating user confidence in E-commerce systems.

I. INTRODUCTION

The same behavior might also be referred to as "business." However, the phrase is more commonly used to describe how the Internet is influencing fields like marketing, logistics, and communication within enterprises. E-commerce is defined here as the action of transacting business online. Business-to-Business (or B2B) The proliferation of information and communication technologies (ICTs), notably the Internet, has led to the widespread adoption of electronic commerce (e-Commerce) in the commercial sector. Consumers who have access to the global market via the Internet have a distinct advantage because of their ability to shop around, learn about items, and see if order fragmentation influences prices. Customers have easy access to comparing the offerings of various e-commerce businesses due to the transparency of the industry. For instance, when making an online purchase, the purchaser's rivals are literally a mouse click away. Customers have much greater leeway to cancel an online purchase than they would in a physical store if they are dissatisfied with the website's items, prices, or services. As far as the Vendors are concerned, a

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 8, August : 2023
brick-and-mortar shop is unnecessary.

BENEFITSOFE-COMMERCE

- The consumer gains the most from being able to get more done in less time and access information from any location. Customers can place buy orders at their convenience. The most compelling arguments in favor of internet shopping are as follows:
- Reduced transaction costs benefited all market participants.
- More flexibility in scheduling and no need to physically visit the company group to complete transactions. The internet has made it possible for consumers to purchase and sell virtually anything, at any time.
- The availability of information is constant and instant. Customers will appreciate the streamlined experience of searching across many platforms with a single cursor movement.
- Customers are able to conduct business with no need to leave the house, office, or any other location.
- Restructure your company -- Customers have the option to switch to a new service at any time if they are dissatisfied with the current one.
- Previously unavailable in regional or national marketplaces, consumers now have a wider selection from which to choose.
- Before making a purchase, a consumer can share their thoughts on a product by writing a review, browsing other customers' purchases, or reading reviews published by other customers.
- It lessens the overall price of purchases and encourages repeat business from current clients.
- Reduces shipping expenses and enhances communication between businesses boosts sales and closes deals faster.
- The company has better internal and external communication. In addition to bolstering the firm's reputation.
- E-commerce platforms would benefit greatly from customer input. Many e-commerce sites allow buyers to leave feedback on products they've purchased. More and more people are willing to tell their friends and complete strangers what they think about the things they purchase through social networking applications or e-commerce platforms because of the proliferation of these channels.

II. LITERATUREREVIEW

2.1 TRUST COMPUTATION

Trust has been examined by academics from various disciplines, including Economics, Management, Computer Science, and Sociology. Trust is explained by the concept of rational choice in economics. After giving it some serious thinking and weighing the costs and benefits, a rational decision can be made regarding the establishment of a trust. "Computational trust" [1] is the term used to describe people's propensity to make choices that are both rational and advantageous. Trust can be determined in two primary ways: directly and indirectly. One way to develop direct trust is to observe past interactions between individuals. Direct trust measures the level of trust between two users who are close relatives. Direct trust is typically determined by analyzing a user's actions, ratings, and other proofs from their most recent encounters.

Dimah investigates if microblogging social data can be mined for actionable user insights. They devised an ISTS approach that makes use of the emotional nature of friends' posts on certain issues and the fact that trust is based on friendship [2]. The trust between peers in P2P networks is managed well thanks to an idea developed by Li and Dai. They allow for post-transaction ratings between parties. A person's trustworthiness

is calculated by adding together all of their ratings. Those in the know will often rely on the trust of their extended network. This type of trust is considered malleable and can arise in the context of familiar events or between normally trustworthy individuals and previously untrustworthy entities.
The method for transmitting data over the semantic web that can be used to calculate the trust number is described by Bo, Yang, and Qiang. The sum of the subjective and objective ratings is used to determine the level of trust in a given situation. Peer relationships provide the subjective component of trust, whereas semantic confidence in information provides the objective component. Things that two coworkers have been through together strengthen their bond and make them trustworthy pals.[3]. The trust management model developed by Faruk and Arnab takes into account the direct, indirect, and worldwide trust in a service in order to determine its total trust value. The distance traveled plays a role in determining the level of indirect trust [4]. Two trust models, one for use within a single domain and one for use between domains, have been developed by Li et al. [5]. Using trust propagation and aggregation in the web of trust, Hong et al. [6] demonstrate Max-aggregation, a new approach that takes a multidimensional and multiattribute view to determine a peer's reputation and obtains the secondary reputation.

2.2 SENTIMENT SIMILARITY ANALYSIS

Similarity analysis is a crucial component of the techniques for determining trustworthiness through the use of emotions. Mood and affective similarity analyses have also received a lot of attention from the fields of statistics, data mining, and natural language processing. There are now three approaches to document level study on mood analysis similarity: the phrase level, the entity and feature level, and the topic level. An opinion lexicon, which is a set of exact words or emotion lexicons identified with their parts of speech, forms the basis for all three tiers and is used to determine the veracity of evaluations. The idea is to classify a whole piece of writing as positive or negative depending on its tone.
Each statement is evaluated to determine whether it presents a positive, negative, or neutral perspective. When something is described as neutral, it lacks an opinion. The study is unable to determine, either at the document or word level, if participants prefer one over the other. This approach goes at the opinion itself at the entity and feature levels rather than the language generation itself (documents, paragraphs, sentences, or phrases). It's predicated on the premise that one's emotions—good or bad—give rise to their perspectives. To learn how individuals felt about various topics on microblogs, Hsu [7] consulted a sentiment word database. He then utilized this data to investigate how various sets of emotion-related terms influenced recommendations for purchases. Existing technologies for evaluating user sentiment, as is customary, primarily aim to classify users' emotions. Its primary function is to categorize the user's preferences according to their likes and dislikes, even at the entity and feature levels. However, these approaches only look at the general trend, which is insufficient when attempting to determine trust based on the degree to which people's feelings are alike. Analyzing the varying opinions on various topics

2.3 ON CORRELATIONS BETWEEN TRUST AND SIMILARITY

Many studies have been conducted in recent years on the correlation between likeness and trust. In order to conduct trustworthy research, sentiment-based similarity analysis has recently gained popularity. Multiple studies have shown that there is a robust relationship between physical similarity and trust. They demonstrated that persons who share similar backgrounds are more likely to trust one another. There are commonalities in their tastes, conversations, and behaviors. Trust and common goals were studied by J. Golbeck and Cai-Nicolas Ziegler. They decided to conduct research into the correlation between physical similarity and interpersonal trust. They provided mathematical techniques for comparing profiles and determining how similar they were. Likeability and trustworthiness were investigated in two separate investigations. Data mining on the Film Trust platform reveals that users' trust in one another fluctuates according to their degree of similarity within a given range. Due to this shift, similarities between people are now a prerequisite for trust [8]. Li proposed an interest-similarity-based trust model in which nodes are interconnected. An interest domain reputation vector was utilized to maintain each node's behavior inside its respective interest domain, taking into account each node's bias and reputation in that domain. Domain-local trust recommendations were evaluated based on the shared interests of similarly situated nodes. These original investigations established a correlation between trustworthiness and likability and provided a scale to quantify the two concepts.

III. PROBLEM STATEMENT

However, most of the other methods for categorizing people based on text emotions focus on analyzing the prevalence of a certain mood or emotional trend. They don't give a hoot about how much agreement there is or how much faith one person has in the other. Additionally, people are more likely to open up about their emotions on social media, which can have a significant impact on the quality of relationships between individuals. Therefore, genuine emotion words should be the starting point if we want the writing to reflect reality and inspire confidence in the reader. According to Bloom, reviews always target specific aspects for their emotive language.

The purpose of our present research is to determine what it takes to maintain client trust in a company over time and following a series of occurrences.

E-commerce continues to expand in popularity despite the fact that nothing is done to earn the trust of current and potential customers. Customers shopping online are increasingly concerned about the veracity of reviews and the authenticity of reviewers.

The owners of these companies rely on these ratings so that they can tailor future strategies to reflect consumer feedback.

Misspelled words, incorrect punctuation, superfluous words, the usage of non-dictionary terms or slang, and ambiguous symbols all make it more difficult to process unstructured material.

IV.OBJECTIVES

Key objectives include

- People are more likely to buy from online stores if they display customer feedback based on research into how people feel about items relevant to the product being sold.
- To enhance commercial practices for increased effectiveness in online trade

SENTIMENT SIMILARITY BASED USER TRUST RELATIONSHIP CALCULATION FRAME WORK

The majority of our clients are internet shoppers. Figure 1 suggests they may have purchased and reviewed a product or service. A user is typically allowed to submit multiple reviews for various products. That's why you'll find customer reviews of products in many sorts of literature. These evaluations are typically performed by gathering information from the network. We present a four-stage computational method for determining trust in e-commerce systems, including direct and propagating trust, based on the sentiment similarity of user reviews. To get started, I've taken item-emotion pairs from reviews. The method is crucial for advancing direct trust computing and sentiment similarity analysis.

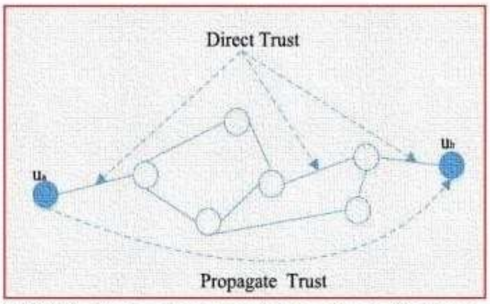

FIG (1). Propagation trust link based on direct trust.

The first step of entity-sentiment word extraction is to examine the text for key words that describe the entity and the sentiment associated with it. We employ the publicly available dictionary of entity words and high-frequency keywords as well as the language parser NL Processor. The relationship between each item and emotion word is calculated using a mutual information formula. Then we can identify meaningful clusters of words.

Then we employ entity-sentiment word pairs to determine the degree of agreement between two reviews written by different people. The second part of the computation for emotional similarity involves determining the degree of similarity between various reviews. The collected entity-emotion word pairs will now be employed in the comparative analysis. Third, we devise a new method for determining the degree of direct trust between reviewers who share an object for evaluation. In the final section, we discuss a method for calculating direct trust from ratings of the same item. How similar the opinions are and how users rank the item are the two most crucial aspects of the calculation technique. The nodes in the trust network are users, and the edges are direct links between users (computed in Part 4).

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 8, August : 2023

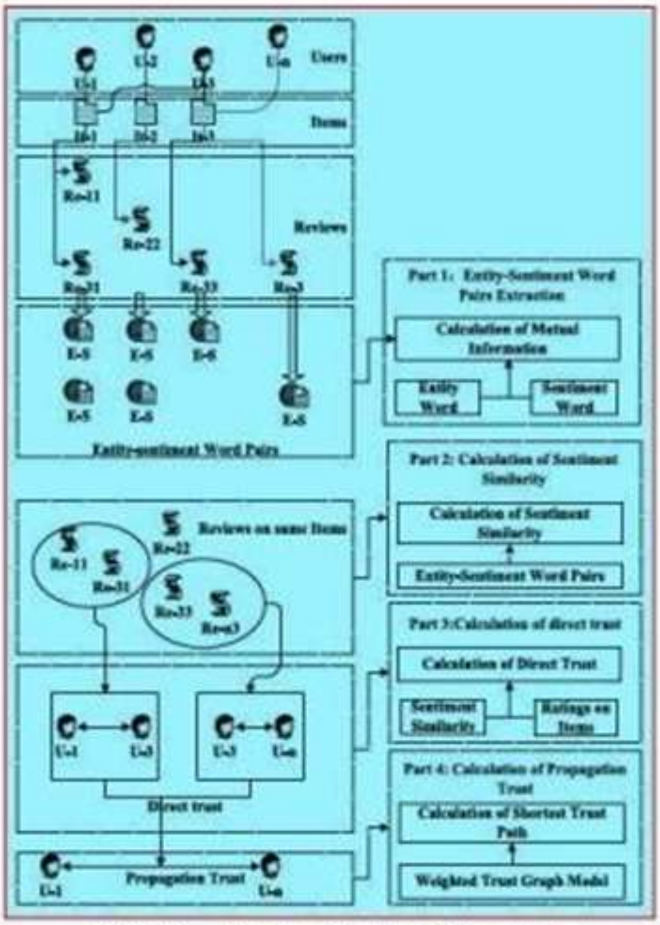

FIG (2). Trust calculation framework

1) Determine the data file's value. There are ratings, comments, and votes.

(2) Data stored about a product, including that product's name, category, price, brand, image quality, and history of being viewed or purchased. We only collect reviews of books, gadgets, sports and leisure activities, video games, and baby goods. The total size of the archive exceeds 350GB.

S.No	Category	Reviews	Items
1	Books	21.5M	2.17M
2	Electronics	8.21M	564K
3	Sports and Outdoors	1.5M	856K
4	Video Games	4.5M	87K
5	Baby	968K	85.3K

Table 1. Selected dataset of reviews from amazon

V.DATASET COLLECTION

Amazon.com provided the data used in the test. There were 143.7 million reviews written for 9.4 million products in 24 categories between May 1996 and July 2014. The dataset contains basic information about reviewers, including their names, IDs, the objects reviewed (such as product names and IDs), the reviews' text and the reviewers' rating statuses. In specifically, the archive contains two documents:

To determine whether there is trust between users, we consider the following three characteristics from the original experiment data set (described in Tabs 1 and 2). We consider two people to be trustworthy if they share the same overall value (product rating) and helpful value (review rating), as well as if they have the same associated items (also bought, also viewed, bought together, bought after seeing). To the best of our knowledge, no previously published studies have utilized a strategy analogous to ours for determining the credibility of an E-commerce platform on the basis of sentiment similarity. To determine the efficacy of our approach, we simply replicated previously conducted studies and provided four metrics by which to evaluate it: precision, memory, F value, and accuracy.

We accept the calculation as true if and only if the sentiment mining algorithm discovers a trust link between any two nodes that is larger than a preset non-zero threshold value. We reject the result as erroneous if the alleged trust is not genuine. The proposed solution does not reveal any trust ties (either directly or via propagation), despite the fact that there is trust between the two users. The term for this kind of connection is "absence of trust." The following formula illustrates the relationship between precision and recall: The F value is used to demonstrate the degree of similarity between two markers. The following is a definition:

$$F - value = \frac{2pecision \times recall}{pecision + recall}$$

$$precision = \frac{CorrectLinksNumber}{CorrectLinksNumber + IncorrectLinksNu}$$

$$recall = \frac{CorrectLinksNumber}{CorrectLinksNumber + MissedLinksNum}$$

VI. EXPERIMENTAL RESULTS AND ANALYSIS

The dataset under scrutiny was partitioned into a training and testing set. These two halves were then divided at random into noncontiguous halves. For training purposes, such as tagging entity-sentiment word pairs, measuring sentiment similarity, and assigning trust levels, we use 80% of the dataset. The remaining 20% of the data set is used to evaluate the efficacy of direct trust and trust that spreads.

The level of trust between users is investigated through experiments that focus on both individual trust trans actions and the spread of trust.

FIG (4). Recommended products

VII. CONCLUSION AND FUTURE WORK

The primary purpose of this research is to investigate consumers' attitudes toward online shopping platforms. By using entity-emotion word pairs and establishing two types of trust relationships (direct trust and spread trust), we pivot our research on user trust to determine the degree to which users' reviews have emotional resonance. Taking into account both ratings and feelings, sentiment similarity analysis measures the degree to which review sentiments are similar and identifies direct trust-related linkages between reviews. We can utilize these two things together to investigate the link between sentiment and immediate trust. We develop a model of trust propagation based on a weighted trust graph. Transmission of trust is the practice of facilitating the spread of trust throughout a group. It's the bond of confidence between two Internet users who have never met face to face. This confidence is established by users playing the role of intermediaries.

REFERENCES:

[1] P.-Y. Hsu, H.-T. Lei, S.-H. Huang, T. H.Liao, Y.-C. Lo, and C.-C. Lo, ‘‘Effects of sentiment on recommendations in social network,’’ in Electron Markets. Berlin, Germany:Springer,2018,pp.1–10,doi:10.1007/s12525-018-0314-5.

[2] C. Qin, W. Siyi, and A. Lin, ‘‘The joint beta distribution with refund rate in onlineC2C trust building: A theoretical study on Taobao,’’in Proc.Int. Conf.E-Learn.E-Technol. Educ. (ICEEE), Lodz, Poland, Sep.2012,pp.191–196.

[3] S. Kraou nakis, I. N. Demetropoulos, A.Michalas,M.S.Obaidat,P.G.Sarigiannidis, and M. D. Louta, ‘‘A robust reputation-based computational model for trust establishment in pervasive systems ,’’IEEE Syst. J., vol.9, no. 3, pp. 878–891,Sep.2015.

[4] P.DeMeo,E.Ferrara,D.Rosaci,andG.M.L.Sarné,‘‘Trustandcompactnessinsocialnetworkgroups,’’IEEETrans.Cybern., vol. 45, no. 2, pp. 205–216, Feb.2015.

[5] M. G. Ozsoy and F. Polat, ‘‘Trust based recommendation systems, ’’in Proc. IEEE/ACM Int.Conf. Adv.Social Netw.Anal.Mining(ASONAM),NiagaraFalls,ON,Canada,Aug.2013,pp.1267–1274.

[6] L. Sheugh and S. H. Alizadeh, ‘‘A fuzzy approach for determination trust threshold in recommender systems based on social network,’’inProc.9thInt.Conf.E-Commerce Developing Countries, Focus E-Bus. (ECDC), Isfahan, Iran, Apr. 2015, pp.1–5.

[7] Y. Ruan, L. Alfantoukh, and A. Durresi,‘‘Exploring stock market using twitter trustnetwork,’’inProc.IEEE29thInt.Conf.Adv.Inf.Netw.Appl.,Gwangiu,SouthKorea,Mar.2015,pp.428–433.

[8] C.-N. Ziegler, Social Web Artifacts for Boosting Recommenders, vol. 487. Berlin,Germany:Springer,2013.

[9] C.-N.Zieglerand J.Golbeck, ‘‘Investigating interactions of trust and interest similarity, ’’Decis. SupportSyst.,vol.43,no.2,pp.460–475,2007.

[10] N.Lathia,S.Hailes,andL.Capra,‘‘Trust-based collaborative filtering, ’’in Trust Management II,vol.263, Y.Karabulut, J.Mitchell,P.Herrmann,and C.D.Jensen,Eds.Boston,MA,USA:Springer,2008,pp.119–134.

[11] D. H. Alahmadi and X.-J. Zeng, ‘‘ISTS: Implicit social trust and sentiment based approach to recommender systems,’’ ExpertSyst. Appl., vol. 42, no. 22, pp. 8840–8849,2015.

[12] M.KoufarisandW.Hampton-Sosa, ‘‘The development to finitial trust in an online company by new customers ,’’Inf.Manage.,vol.41,no.3,pp.377–397,2004.

[13] H. Zhong, S. Zhang, Y. Wang, and Y.Shu,‘‘Study on directed trust graph based recommendation for E-commerce system,’’Int.J. Comput.Commun.Control,vol.9,no.4,pp.510–523,2014.