Volume 5, Issue 11, November 2015 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

A Survey to Identify the Behavior of Consumer Ratings in Internet Commerce

N. Madhu Suganya, N. Tamilelakkiya
Assistant Professor, K.Ramakrishnan College of Engineering, Trichy, Tamilnadu, India
PG Scholar, Anna University, Bit Campus, Tamilnadu, India

Abstract- Nowadays people prefer online shopping over conventional shopping. People enjoy online shopping experiences by publishing, browsing, or sharing product reviews written by themselves or others. Indeed, research shows that people rely on the opinions of others when selecting something for the first time. The average customer rating on a product, (reputation), is one of the key factors in online purchasing decisions. There is, no guarantee of the trust- worthiness of a reputation since it can be manipulated rather easily. In this paper, I propose the survey detail to make trusted online shopping environment system.

Key term- Reputation, trust, shilling attack, social networks

I. INTRODUCTION

Social network analysis has recently gained a lot of interest because of the advent and the increasing popularity of social media, such as blogs, social networking applications, micro blogging, or customer review sites. In online shopping channels, consumers share their purchasing experiences regarding both goods and services with other potential buyers via evaluation. The most common way for consumers to express their level of satisfaction with their purchases is through online ratings. The overall buyers' satisfaction is quantified as the aggregated score of all ratings and is available to all potential buyers.

In this environment, trust is becoming an essential quality among user interactions and the recommendation for useful content and trustful users is crucial for all the members of the network. One common type of analysis is the identification of communities of users with similar interests. Another research direction is the identification of content that could be of potential interest, whether this is a product review, a blog, or a tweet. Collaborative filtering is the most broadly adopted technique used to predict future item ratings based on the user's past behavior as well as ratings of other similar users. It has been shown that incorporating social network relationships (e.g., friendship) and respective opinions/ratings improves.

However, there is always a natural profit incentive to promote one's own products by leaving biased online comments and/or extreme ratings for the competition's products, which has spawned the so-called "shilling attacks". It appears that shilling attacks are successfully threatening recommender systems, and continuously generating huge volumes of spam or misleading review comments. Shilling attacks are therefore emerging as a considerable challenge to the sustainable development of online shopping. Recently, trust has been introduced in the context of recommender systems for social networks.

II. LITERATURE SURVEY

A. A Trust-Aware System for Personalized User Recommendations in Social Networks

Magdalini Eirinaki, Malamati D. Louta,Member, IEEE,and Iraklis Varlamis, Member, IEEE

This paper introduce a framework for handling trust in social networks, which is based on a reputation mechanism that captures the implicit and explicit connections between the network members, analyzes the semantics and dynamics of these connections, and provides personalized user recommendations to the network members.

More specifically, the proposed system provides users with personalized positive and/or negative recommendations that can be used to establish new trust/distrust connections in the social network. It assumes the notion of trust captures both the user's social context (e.g., friends and enemies) expressed through explicit user-to-user connections, as well as users' common interests and desires inferred from explicit and implicit user-to-item connections. The proposed recommender system is based on a reputation mechanism that rates participants using observations, past experiences, and other user's view/opinion.

In order to compute the reputation of each member, it adopts several properties of trust such as, transitivity, personalization, and context, and draw ideas from sociology axioms. Additionally, in order to address the social network dynamics, it has incorporated in our system the element of time. To this direction, it suggests that reputation fades by time; thus, the positive (negative) reputation value of a user tends to zero unless new explicit or implicit trust (distrust)

and liking (disliking) statements are added frequently. Finally, the context of trust is the same among community members. It exploits positive and negative, time-dependent trust-related information, expressed either explicitly or implicitly. The system can be applied to any type of social network, even in the absence of explicit trust connections, since in such cases only the implicit expressions of trust will be considered for the ranking and recommendation of users.

B. Using Machine Learning to Augment Collaborative Filtering of Community Discussions

Michael Brennan, Stacey Wrazien, Rachel Greenstadt

This paper chose to mine data from Slashdot (slashdot.org), a technology news site and online community. Readers of the site submit articles which are reviewed by a team of editors, who select the best ones to post as the news items for that day. The community then discusses the articles and issues posted through a comment system. Each news post has its own comment series. Slashdot has implemented a

collaborative filtering system for users to rank the comments on how relevant they are to the article and to other users on a scale from -1 to 5, with 5 signifying the comments most worth reading.

Comments that receive a very low score are typically hidden, while comments with a higher score are highlighted, allowing the user to easily reach quality commentary. In addition to the numerical rating posts we can also be given a rating description such as "Insightful" if it is good and "Offtopic" if it is bad, among others. The features used to classify Slashdot comments are divided into two groups: linguistic features and contextual and author reputation features. The linguistic set represents features related to the words, their meanings, and the structure of the text.

Most of the linguistic features were extracted from the comments using the Linguistic Inquiry and Word Count (LIWC) software, a text analysis database designed by psychologists to study various emotional, cognitive, and structural components of verbal and written speech. The contextual and author reputation features are based upon information such as when it was posted or how much discussion it generated, or information about the author such as what his or her recent comment ratings have been. A full list of features can be found on the web 1. All classification was performed using a SVM Classifier that used a Gaussian radial basis function kernel. The features were all discretized into four bins before being used for classification (except LIWC sentiment which already had three discrete values).

C. Can You Trust Online Ratings? A Mutual Reinforcement Model for Trustworthy Online Rating Systems

Hyun-Kyo Oh, Sang-Wook Kim, Member, IEEE , Sunju Park, and Ming Zhou

This paper defines false reputation as the problem of a reputation being manipulated by unfair ratings. For this purpose, we propose TRUE-REPUTATION, an algorithm that iteratively adjusts a reputation based on the confidence of customer ratings.

The proposed framework, on the other hand, uses all ratings. It evaluates the level of trustworthiness (confidence) of each rating and adjusts the reputation based on the confidence of ratings. The algorithm that iteratively adjusts a reputation based on the confidence of customer ratings. By adjusting a reputation based on the confidence scores of all ratings, the proposed algorithm calculates the reputation without the risk of omitting ratings by normal users while reducing the influence of unfair ratings by abusers. This algorithm solves the false reputation problem by computing the true reputation, TRUE-REPUTATION.

The computation of a trustworthy reputation starts by measuring the confidence of a rating. To determine the confidence of a rating, therefore, we have adopted three key factors of activity, objectivity, and consistency and defined these factors in the context of online ratings.

First, the user who rates more items displays a higher level of activity. Therefore, we measure user activity in an online rating system based on the number of products he rates. Second, a rating is considered more objective (rating is defined as the deviation of the rating from the general reputation of the item) if it is closer to the public's evaluation. The more similar are the rating and the reputation, the higher is the objectivity of a rating; the more dissimilar they are, the lower the objectivity of a rating.

Additionally, a user whose ratings exhibit higher objectivities should also have a higher level of user objectivity. The objectivity of a rating is calculated based on the deviation of the "rating" from the "reputation" of the product. The difficulty in computing a reputation lies in the fact that the reputation itself is the sum of the ratings adjusted by the confidence, and the confidence of an individual rating is computed using the objectivity of the rating, which uses the reputation in its computation. In other words, the reputation and the confidence of a rating interact with each other in mutual reinforcement. The proposed framework does not require clustering or classification, both of which necessitate considerable learning time.

D. Detecting Product Review Spammers using Rating Behaviors

Ee-Peng Lim, Viet-An Nguyen, Nitin Jindal

This paper aims to detect users generating spam reviews or review spammers. It identifies several characteristic behaviors of review spammers and models these behaviors so as to detect the spammers. In particular, it seeks to model the following behaviors. First, spammers may target specific products or product groups in order to maximize their impact. Second, they tend to deviate from the other reviewers in their ratings of products.

It proposes scoring methods to measure the degree of spam for each reviewer and apply them on an Amazon review dataset. And then select a subset of highly suspicious reviewers for further scrutiny by our user evaluators with the help of a web based spammer evaluation software specially developed for user evaluation experiments.

E. Shin: Generalized Trust Propagation with Limited Evidence
Chung-Wei Hang, Zhe Zhang, and Munindar P. Singh

To address the need for trust, even when no suitable forward path exists, we created Shin (the Chinese word for trust), a generalized propagation technique that uses a probabilistic paradigm to estimate trust by comparing assessments from acquaintances that the truster and trustee have in common. In developing Shin, we included two of CertProp's three trust propagation operators, but we also extended CertProp to improve prediction accuracy for backward paths. Our evaluation of Shin's capabilities shows that it is superior to CertProp and other existing approaches when only a few trustworthy forward paths exist from the truster to the trustee.

Shin is based on the idea that it is possible to compute the trust relationship between a truster and trustee using the known direct trust relationships between agent pairs in the network that are proximal to the truster and trustee. Mathematically, a trust network T(V,E,d) captures agents as vertices V and direct trust relationships as directed, weighted edges E, with the weight d(a,b) of an edge from a to b expressing the amount of direct trust placed by truster a in trustee

b. Shin measures direct trust as a value between zero and one, and assigns a trust network an edge if and only if the corresponding direct trust is nonzero. In addition, Shin computes and uses the function t: $V\times V\rightarrow[0, 1]$ such that for $a,b\in V$, t(a, b) is the amount of (direct or indirect) trust that truster a places in trustee b.

In simple terms, trust propagation is the problem of computing the amount of trust for a nonadjacent truster and trustee, or t(a,b). As part of that computation, Shin uses CertProp's concatenation operator ($\otimes$), which discounts trust values along a referral path, and its aggregation operator ($\oplus$), which combines trust from referral paths. The "Trust as Evidence and Belief Representations" sidebar describes the mathematical background of Shin's propagation approach in more detail.

F. iCLUB: An Integrated Clustering-Based Approach to Improve the Robustness of Reputation Systems
Siyuan Liu, Jie Zhang, Chunyan Miao, Yin-Leng Theng, Alex C. Kot

Proposes an integrated CLUstering-Based approach called iCLUB to filter unfair testimonies for reputation systems using multi-nominal testimonies, in multiagent based electronic commerce. It adopts clustering and considers buying agents' local and global knowledge about selling agents.

In Algorithm 1, the Local component first collects the local information regarding

S_t. DBSCAN, a density based clustering routine , is then applied on the collected

Testimonies L^B_{St} to generate a set of clusters. After that, the Local component returns as honest witnesses the set of witnesses whose rating vectors are included in the same cluster as the buying agent's rating vector.

In Algorithm 2, the Global component first finds the honest witnesses for each seller with which the buyer has transactions, using the Local() procedure . Then, a set of common honest witnesses W_F are formed as the intersection of the set of the honest witnesses for each seller except S_t. The Global component obtains the clustering result for S_t. It then calculates the intersection of WF with the witnesses whose rating vectors are in each cluster achieved if W_F is not an empty set. Finally, it returns as honest witnesses the ones whose rating vectors are in the cluster which has the largest intersection result with W_F. Our iCLUB approach further integrates the Local and Global components using a threshold ε. If the number of transactions between B and S_t is greater than ε, Global () procedure will be triggered, otherwise Local () procedure will be called.

G. Preventing Shilling Attacks in Online Recommender Systems
Paul-Alexandru Chirita, Wolfgang Nejdl, Cristian Zamfir

This paper proposes several metrics for analyzing rating patterns of malicious users and evaluate their potential for detecting such shilling attacks . Building upon these results, we propose and evaluate an algorithm for protecting recommender systems against shilling attacks. The algorithm can be employed for monitoring user ratings and removing shilling attacker profiles from the process of computing recommendations, thus maintaining the high quality of the recommendations.

More specifically, the following metrics suitable to address problem of detecting shilling attacks:

1.Number of Prediction-Differences (NPD), 2.Standard Deviation in User's Ratings, 3. Degree of Agreement with Other Users, 4.Degree of Similarity with Top Neighbors.

The algorithm computes for each user the values for all statistical metrics, and then decides, based on her assessed probability of being an attacker, whether her profile will be discarded from the computation of recommendations or not

H. HySAD: A Semi-Supervised Hybrid Shilling Attack Detector for Trustworthy Product Recommendation
Zhiang Wu, Junjie Wu,JieCao, Dacheng Tao

Shilling attackers apply biased rating profiles to recommender systems for manipulating online product recommendations. Although many studies have been devoted to shilling attack detection, few of them can handle the hybrid shilling attacks that usually happen in practice, and the studies for real-life applications are rarely seen. Moreover, little attention has yet been paid to modeling both labeled and unlabeled user profiles, although there are often a few labeled but numerous unlabeled users available in practice.

This paper presents a Hybrid Shilling Attack Detector, or HySAD for short, to tackle these problems. In particular, HySAD introduces MC-Relief to select effective detection metrics, and Semi-supervised Naive Bayes (SNBλ) to precisely separate Random-Filler model attackers and Average-Filler model attackers from normal users. Thorough experiments on Movie-Lens and Netflix datasets demonstrate the effectiveness of HySAD in detecting hybrid shilling attacks, and its robustness for various obfuscated strategies.

A real-life case study on product reviews of Amazon.cn is also provided, which further demonstrates that HySAD can effectively improve the accuracy of a collaborative-filtering based recommender system, and provide interesting opportunities for in-depth analysis of attacker behaviors. These, in turn, justify the value of HySAD for real-world applications

IV. CONCLUSION

In this paper I have studied the problems in E-Marketplaces and various solutions how to overcome some of the problems. There are more factors (other than those addressed in this paper) known to be elemental in assessing the trust of users in the field of social and behavioral sciences. I plan to study how to incorporate them into our model to compute the reputation of items more accurately. In the e-market place such as Amazon.com and eBay.com, buyers give ratings on items they have purchased. I note, however, that the rating given by a buyer indicates the degree of his satisfaction not only with the item (e.g., the quality) but also with its seller (e.g., the promptness of delivery). In a further study, I plan how to develop an approach to accurately separate an item score and a seller score from a user rating. Separating the true reputation of items and that of sellers would enable customers to judge items and sellers independently

REFERENCES

[1] A Trust-Aware System for Personalized User Recommendations in Social Networks, Magdalini Eirinaki, Malamati D. Louta,Member, IEEE,and Iraklis Varlamis, Member, IEEE, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS, VOL. 44, NO. 4, APRIL 2014

[2] Using Machine Learning to Augment Collaborative Filtering of Community Discussions, Michael Brennan, Stacey Wrazien, Rachel Greenstadt, Proc. of 9th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2010)

[3] Can You Trust Online Ratings? A Mutual Reinforcement Model for Trustworthy Online Rating Systems, Hyun-Kyo Oh, Sang-Wook Kim, Member, IEEE , Sunju Park, and Ming Zhou, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS ,YEAR 2015

[4] Detecting Product Review Spammers using Rating Behaviors,Ee-Peng Lim, Viet-An Nguyen, Nitin Jindal, CIKM'10,October 26–30, 2010

[5] Shin: Generalized Trust Propagation with Limited Evidence, Chung-Wei Hang, Zhe Zhang, and Munindar P. Singh, EEE Computer Society 2013.

[6] iCLUB: An Integrated Clustering-Based Approach to Improve the Robustness of Reputation Systems, Siyuan Liu, Jie Zhang, Chunyan Miao, Yin-Leng Theng, Alex C. Kot, Proc. of 10th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2011).

[7] Preventing Shilling Attacks in Online Recommender Systems, Paul-Alexandru Chirita, Wolfgang Nejdl, Cristian Zamfir, WIDM'05,November 5, 2005.

[8] HySAD: A Semi-Supervised Hybrid Shilling Attack Detector for Trustworthy Product Recommendation, Zhiang Wu, Junjie Wu,JieCao, Dacheng Tao, KDD'12, August 12–16, 2012

