Integrating CMP-Blockchain with Natural Language Processing and Machine Learning for Trust Verification and Event Detection

S.Nagamani, S.Saritha, N.Savitha, V.Chiranjeevi

[1] ASSOC.PROFESSOR, [2,3,4] ASSIT.PROFESSOR

Department of CSE, SWARNA BHARATHI INSTITUTE OF SCIENCE & TECHNOLOGY (SBIT), Pakabanda Street, Khammam - 507 002. Telangana, India.

ABSTRACT

An integral part of our everyday lives is spent on social networks, and one of the most important parts of these networks is the so-called social reviewing system (SRS), which allows us to access data, usually in the form of reviews. The significance of social networks necessitates that they be trustworthy and secure, preventing assaults and misuses and allowing users to freely utilise the information they provide. False reviews are a major weapon in the fight against the reputation system. Since even verified members of the network are capable of launching such attacks, a strong defence is to take advantage of trust management by giving each user a trust level and then having them use it to evaluate the collected data. Because it is subjective and difficult to completely automate the process of detecting improper behaviours, trust management within the framework of SRSs is especially complex. Despite several proposals in the existing literature, this matter has not yet been fully addressed. By using the innovative notion of time-dependent and content-dependent crown consensus and modelling trust management as a multicriteria multiexpert decision making, this work proposes a remedy against mendacious reviews that integrates fuzzy logic with the theory of evidence. Even when faced with sockpuppet assaults, our method proved to be more effective than the primary methodologies described in related literature.

INTRODUCTION

AS WELL known, the online social networks [1] are Internet-enabled applications used by people to establish social relations with the other individuals sharing similar personal interests and/or activities. Apart from exchanging personal data, such as photographs or videos, mainly all these applications allow their users to share comments and opinions on specific topics, so as to suggest objects or places of interest (e.g., Trip Advisor,

Foursquare, etc.) or to provide social environments able to facilitate particular tasks (e.g., the search of a job as in LinkedIn, the answer to research questions as in Research Gate, purchases on Amazon, etc.). Due to this comment/opinion sharing, these social applications, which we will refer to as social reviewing systems (SRSs) have been extensively used when people need to make daily decisions, increasing their popularity. As a concrete example, most of us access to a preferable SRS before choosing a restaurant or buying something so as to get reviews and feedback. People are progressively and symbiotically dependent on them as proved by the advanced opinion modeling and analysis, exploiting the impact of neighbors on user preferences or approaching the existing information overload in SRS, such as [2], [3]. For this reason, the trustworthiness of SRS is particularly important, and a key concern for effective opinion dynamics and trust propagation within a community of users [4]. In fact, SRSs suffer from forged messages and camouflaged/fake users that are able to avoid individuals take the right decision. This may raise several issues about privacy and security [5], mainly due to the fact that several personal and sensitive information are shared, and leaked, throughout SRS [6], [7], and that a person may choose to hide its true self and intentions behind a totally false virtual identity [8] or a Bot (short for software robots) may mimic human behavior in SRS [9]. In addition, threats in SRS, such as data leaks, phishing bait, information tampering, and so on, are never limited to a given social actor, but spread across the network like an infection by obtaining victims among the friends of the infested actors. So, an SRS provider needs to provide proper protection means to guarantee its trustworthiness.

Some works in the current literature, such as [10], mostly deal only with forging messages as this can be easily resolved by using cryptography. However, the second kind of malicious behavior caused by camouflaged/fake users is still an open issue. During the last decade, several solutions have been proposed in order to deal with the problem of camouflaged/fake users [11]–[13]. The issue of providing privacy has led to the adoption of access control means, while counteracting forging nodes/identities and social links/connections demanded authentication of users and exchanged messages [14], [15]. Mostly, such mechanisms aim at approaching external attackers or intruders, while thwarting

legitimate participants in the SRS acting in a malicious way is extremely challenging. A naive way to protect against malicious individuals is to have users being careful when choosing with whom to have a relationship. Two users in social networks may have various kinds of relationships: 1) in Face book-like systems users can indicate others as "friends," or 2) in Instagram-like systems a user can "follow" others. However, users are typically not so careful when accepting received joining requests, and selecting other users to be connected with is typically extremely difficult (as malicious users are also experts in camouflaging themselves). Despite the relationships among the social actors within an SRS should be based on the direct knowledge in the real life of the people behind such actors (such as former classmates, colleagues, or member of the same family or group of friends), the majority of the relationships are typically made without such a face-toface knowledge but among users that have never been met in person. *Trust management* is among the most popular solution to fight against such inside attackers [16]. It consists to assign a "trust" value to users based on the direct analysis of their behaviors or indirect trust relationship among social actors. To this aim, it is a soft secure measure implying the revocation of a social link toward those actors with a low trust value, or to strengthen the protection measures for those actors exhibiting a low trust degree, by limiting the data/functionalities that
they can have access to. Despite being a powerful protection means [17], trust management is not explicitly provided by the main SRS platforms, due to the issues related to its automatic computation.

There is the problem to select the data of interest upon which computing the trust degree among the vast amount of shared information, which shows the main features (volume, variety, and velocity) of *big data*. To simplify the problem, a well-investigated aspect is the study of trust network [18]. Specifically, an SRS is seen as a graph, where each vertex is a social actor, and each link models a social relationship between two actors where a trust value is assigned by one to another by means of the previous computation approach [19], [20]. It is not rare that actors may interact with nonadjacent other actors, so it is important to find a trust path among nonadjacent actors to compute trust transitivity (so that they can interact). However, there is still the problem on how measuring the mutual trust of two

actors connected by a social relationship, which is further used for trust transitivity for unrelated users having some related users in common. This can be roughly measured as the ratio of the good iterations over the total number of elapsed interactions, even if more complex models have been proposed. Possible violations against the ethical norms cannot be objectively determined (meaning that such judgments are absolutely true or false and it is possible to assign them a 0 or 1 value) but are strongly based on or influenced by the person making the judgment (i.e., are subjective) and expressed in a partial truth way (i.e., judgement can range from completely true to completely false and it is possible to assign a real number going from 0 to 1), due to uncertain and vague natures of the behavioral data collected on human interactions. This makes the overall trust management an example of the so-called fuzzy decision-making problem [21] and make its fully automation extremely complex. Protecting the overall aggregation from the impact of malicious or fake reputations is an issue to consider [22], and falls within the literature of security and privacy of *Recommender Systems* [23].

Our work aims to contribute to the on-going efforts on the trust computation in SRSs, where the behaviors of social actors are described by means of their submitted "reviews" on specific objects of interest. Specifically, a genuine review may reflect a correct behavior of the actor, while a deceitful review is a sign of a malicious behavior. To deal with the mentioned big data problem, we have considered those social application to recommend objects of interest, since they restrict the kind of behavioral data to reviews as text content in natural language. Despite the current SRS providers are reluctant to disclose their data (since mainly sensitive for their users and/or business), reviews are publicly accessible and user review datasets are largely available. To deal with the subjective review judgement, we leverage on the fuzzy theory as widely used in [24]. To deal with the problem of computing trust computation, we propose a proper robust aggregation means of the outcomes of review analyzers, each evaluating incoming reviews based on a specific criterion. To deal with the problem of mendacious reviews, we estimate which reviews contains opinions deviating from the evaluation of an objective of interest from the majority, as only a small portion of the reviewers is malicious.

The contribution of our work consists in the definition of a proper process to estimate the *trustworthiness* of social actors based on their published reviews and to achieve robustness against possible false opinions as follows.

1) Identifying possible mendacious reviews by exploiting a multi criteria decision making and introducing the novel concept of time-dependent and content-dependent crown consensus, where various criteria are used to evaluate the quality of a given review.

2) Performing reputation aggregation based on the Dempster–Shafer (D-S) combination rule [25] so as to infer the user trustworthiness.

3) Implementing the proposed approach in a cloud-based platform by crawling reviews from heterogeneous datasets, preprocessing (by performing data cleaning) and storing the acquired reviews in a NOSQL database, and realizing the envisioned trust computation by using an analytics engine for big data processing.

4) Experimenting the proposed approach and implemented solution on two different datasets, one from the Yelp Dataset Challenge and the other from Amazon Customer Review Datasets. We have also adopted the *Yelp NYC* dataset so as to run the effectiveness evaluation of the approach against some of the main works within the literature. Such experiments proved the higher degree of precision and the user ranking challenges than the other approaches.

Several similar approaches have been proposed in the literature for evaluating user trustworthiness by using the textual review, especially in the context of spam detection. They can be generally classified in three groups: 1) linguistic based [26], focusing on the identification of linguistic features of malicious reviews; 2) behavioral based [27], leveraging metadata information of submitted review and user profile for identifying fake reviews; and 3) graph-based approaches [28], analyzing users and objects ties. The proposed framework exploits a behavioral analysis by combining in a novel way reviews' metadata, user compliments and rate's variation over time by leveraging fuzzy logic and the theory of evidence. A novel set of criteria has been formulated in order to determine the trustworthiness of reviews, and they are aggregated so as to determine the overall user trust degree. The evidence theory has been used to compute trust by aggregating binary evidences, such as in [29] and [30]. Our novelty is represented by its application to users'

trustworthiness assessment based on reviews' quality scores.

In the proposed approach, users' trustworthiness is not computed by considering their relationships but only the reviews' features. This is because users' relationships have a limited consideration for SRSs. However, it is possible to integrate our approach with one of those in the literature computing user trustworthiness based on the established relationships, as it may be seen as an additional criterion to be aggregated with the D-S combination rule within our proposed multi criteria decision making process.

EXISTING SYSTEM

Yu *et al.* [38] described an approach for computing user trustworthiness by leveraging on the "familiarity" and "similarity" concepts and considering the influence of user actions on the trustworthiness computation. The aim of this methodology is to detect malicious users-based also on a security queue to record users' historical trust information. Afterward, Yu *et al.* [39] proposed an approach based on deep learning techniques in conjunction with user trustworthiness characterization for configuring privacy settings for social image sharing. In addition, a two-phase trust-based approach based on deep learning techniques has also been proposed by Deng *et al.* [40] for social network recommendation, so as to determine the users' interests and their trusted friends' interests together with the impact of community effect for recommendations.

Rayana and Akoglu [27] presented a system, namely, *SpEagle*, that uses metadata (i.e., text, timestamp, and rating) in conjunction with relational data to spot suspicious users and reviews.

Other related approaches exploit reviews' evaluation for detecting and/or characterizing spam in social media. Shehnepoor *et al.* [28] proposed a framework named *NetSpam* that models reviews in online social media, as a case of heterogeneous networks, by using spam features for detection purposes. Ye *et al.* [41] described an approach based on the temporal analysis by monitoring selected indicative signals of opinion spams over the time, for detecting and characterizing

abnormal events in real time.

A system based on four integrated components, specifically: 1) a reputation-based component; 2) a credibility classifier engine; 3) a user experience component; and 4) a featureranking algorithm, has been designed and implemented by

Alrubian *et al.* [42] for assessing information credibility on Twitter. In [43], the *CommTrust* framework has been introduced for trust evaluations by mining feedback comments. More in detail, it is based on a multidimensional trust model for computing reputation scores from user feedback comments, which are analyzing combining natural language processing techniques, opinion mining, and topic modeling.

Furthermore, another framework, namely, *LiquidCrowd*, has been proposed by Castano *et al.* [44] exploiting consensus and trustworthiness techniques for managing the execution of collective tasks. Kumar *et al.* [45] proposed a system, namely, *FairJudge*, to identify fraudulent users based on the mutually recursive

definition of the following three metrics: 1) the user trustworthiness in rating products; 2) the rating reliability; and 3) the goodness of a product. Moreover, Kumar *et al.* [46] described a system for identifying fraudulent users based on six axioms to define the interdependency among three intrinsic quality metrics concerning a user, reliability and goodness of a product by combining network and behavior properties. Hooi *et al.* [47] developed an algorithm, called *FraudAR*, aiming at being resistant to the camouflage attacks, for identifying fake reviews and users. Furthermore, *Birdnest*, an approach combining Bayesian model of user rating behavior and a likelihood-based suspiciousness metric [normalized expected surprise total (NEST)], has been proposed in [48].

Liu *et al.* [32] investigated the sockpuppet attacks on reviewing A system based on four integrated components, specifically:1) a reputation-based component; 2) a credibility classifier engine; 3) a user experience component; and 4) a feature ranking algorithm, has been designed and implemented by Alrubian *et al.* [42] for assessing information credibility on

Twitter. In [43], the *CommTrust* framework has been introduced for trust evaluations by mining feedback comments. More in detail, it is based on a multidimensional trust model for computing reputation scores from user feedback comments, which are analyzing combining natural language processing techniques, opinion mining, and topic modeling. Furthermore, another framework, namely, *LiquidCrowd*, has been proposed by Castano *et al.* [44] exploiting consensus and trustworthiness techniques for managing the execution of collective tasks.

Kumar *et al.* [45] proposed a system, namely, *FairJudge*, to identify fraudulent users based on the mutually recursive definition of the following three metrics: 1) the user trustworthiness in rating products; 2) the rating reliability; and 3) the goodness of a product. Moreover, Kumar *et al.* [46] described a system for identifying fraudulent users based on six axioms to define the interdependency among three intrinsic quality metrics concerning a user, reliability and goodness of a product by combining network and behavior properties.

Hooi *et al.* [47] developed an algorithm, called *FraudAR*,

aiming at being resistant to the camouflage attacks, for identifying fake reviews and users. Furthermore, *Birdnest*, an approach combining Bayesian model of user rating behavior and a likelihood-based suspiciousness metric [normalized expected surprise total (NEST)], has been proposed in [48]. Liu *et al.* [32] investigated the sockpuppet attacks on reviewing systems by proposing a fraud detection algorithm, called RTV, that introduces trusted users and also considers reviews left by verified users.

Disadvantages

- An existing methodology doesn't implement multicriteria multiexpert decision-making (MCME-DM) method.
- The system not implemented Dempster–Shafer (D-S) theory.

PROPOSED SYSTEM

The contribution of our work consists in the definition of a proper process to estimate the *trustworthiness* of social actors based on their published reviews and to achieve robustness against possible false opinions as follows.

1) Identifying possible mendacious reviews by exploiting a multicriteria decision making and introducing the novel concept of time-dependent and content-dependent crown consensus, where various criteria are used to evaluate the quality of a given review.

2) Performing reputation aggregation based on the Dempster–Shafer (D-S) combination rule [25] so as to infer the user trustworthiness.

3) Implementing the proposed approach in a cloud-based platform by crawling reviews from heterogeneous datasets, preprocessing (by performing data cleaning) and storing the acquired reviews in a NoSQL database, and realizing the envisioned trust computation by using an analytics engine for big data processing.

4) Experimenting the proposed approach and implemented solution on two different datasets, one from the Yelp

Dataset Challenge and the other from Amazon Customer Review Datasets. We have also adopted the *YelpNYC* dataset so as to run the effectiveness evaluation of the approach against some of the main works within the
literature. Such experiments proved the higher degree of precision and the user ranking challenges than the other approaches.

Advantages

1) Effectiveness: To examine the accuracy of the proposed approach by varying expert and criteria weights and to compare our technique with the other ones proposed in the literature.

2) Robustness With Respect to Sockpuppet Attacks: For analyzing how the proposed approach deals with this particular attack by varying the percentage of most suspicious accounts considered fraudsters and to compare the obtained results with respect to the state-of-the-art ones.

IMPLEMENTATION

A system based on four integrated components, specifically:

1) a reputation-based component;

2) a credibility classified engine;

3) a user experience component; and

4) a feature ranking algorithm, has been designed and implemented.

Targeting and connecting with potential customers by exploiting an SRS as a simple method of advertising is the first example of such a business-related use, but actually it is their weakest business use. Support Vector Machine or SVM is one of the most popular Supervised Learning algorithms, which is used for Classification as well as Regression problems.The goal of the SVM algorithm is to create the best line or decision boundary that can segregate n- dimensional space into classes so that we can easily put the new data point in the correct category in the future.

Service Provider

In this module, the Service Provider has to login by using valid user name and password. After login successful he can do some operations such as
Login, Browse Product Review Datasets and Train & Test Data Sets, View Trained and Tested Accuracy in Bar Chart, View Trained and Tested Accuracy Results, View Prediction Of Product Review Trust, View Prediction Of Product Review Trust Ratio, Download Predicted Data Sets, View Prediction Of Product Review Trust Ratio Results, View All Remote Users.

View and Authorize Users

In this module, the admin can view the list of users who all registered. In this, the admin can view the user's details such as, user name, email, address and admin authorizes the users.

Remote User

In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like REGISTER AND LOGIN, PREDICT PRODUCT REVIEW TRUST TYPE, VIEW YOUR PROFILE

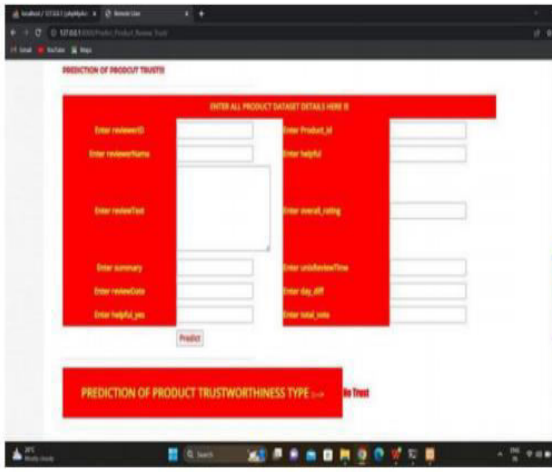

Fig.1. Home page.

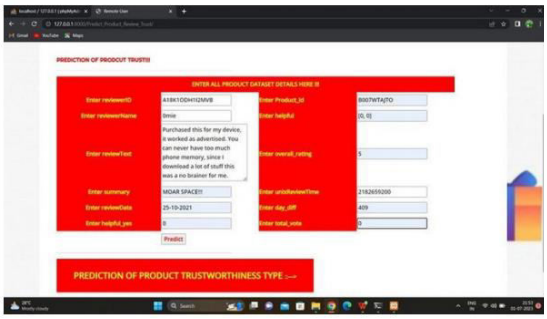

Fig.2. Product details page.

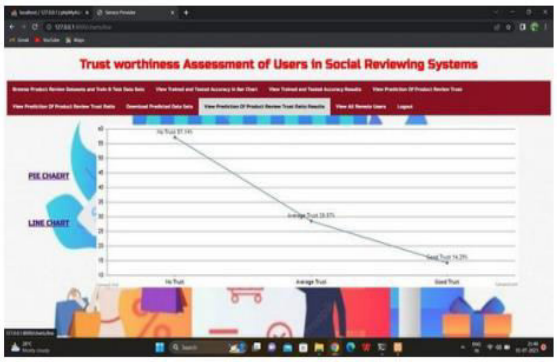

Fig.3. Accuracy details.

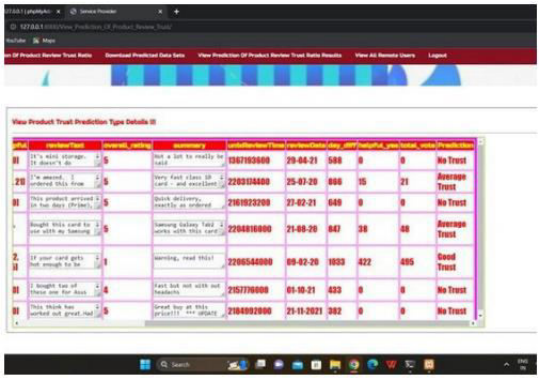

Fig.4. Output results.

CONCLUSION

This study proposed a solution to the problem of trust management within the context of the social networks, where it is important to deal with the subjectivity of the detection of malicious behaviors and the need of objectivity in order to design an automatic process to assign trust degrees to users based on their activity in the social network. To this aim, we have approached the vagueness and subjectivity in the review analysis from the social network by means of the fuzzy theory. We have leveraged on the theory of evidence so as to device a MCME-DM process to aggregate the judgments from multiple perspectives

and optimize the trust estimation.Amazon datasetand showed that aggregating the output of multiple criteria allows achieving higher accuracy in detecting malicious reviews. We have also compared our approach against the main relatedworks in the existing literature and showed that our approach obtained better efficacy by using 80% and 100% of the considered dataset.As this kind of attacks can be conducted by legitimate users of the network, a particularly powerful solution is to exploit trust management, by assigning a trust degree to users, so that people can weigh the gathered data based on such trust degrees. Trust management within thecontext of SRSs is particularly challenging, as determining incorrect behaviors is subjectiveand hard to be fully automatized. Several attempts in the current literature have been proposed.In methodology Navie Bayes algorithm is Best to use in this project.

REFERANCES

[1] M. Faloutsos, T. Karagiannis, and S. Moon, "Online social networks," IEEE Netw., vol. 24, no. 5, pp. 4–5, Sep/Oct. 2010.

[2] J. Castro, J. Lu, G. Zhang, Y. Dong, and L. Martinez, "Opinion dynamics-based group recommender systems," IEEE Trans. Syst., Man, Cybern., Syst., vol. 48, no. 12, pp. 2394–2406, Dec. 2018.

[3] F.Xiong, X. Wang, S. Pan, H. Yang, H. Wang, and C. Zhang, "Social recommendation with evolutionary opinion dynamics," IEEE Trans.Syst., Man, Cybern., Syst., vol. 50, no. 10, pp. 3804–3816, Oct. 2020.

[4] R. Ureña, G. Kou, Y. Dong, F. Chiclana, and E. Herrera-Viedma, "A review on trust propagation and opinion dynamics in social networks and group decision making frameworks," Inf. Sci., vol. 478, pp. 461–475, Apr. 2019.

[5] Y. Xiang, E. Bertino, and M. Kutylowski, "Security and privacy in social networks," Concurrency Comput. Practice Exp., vol. 29, no. 7, 2017, Art. no. e4093.

[6] D. Irani, S. Webb, K. Li, and C. Pu, "Modeling unintended personal information leakage from multiple online social networks," IEEE Internet Comput., vol. 15, no. 3, pp. 13–19, May/Jun. 2011.

[7] A. Nosko, E. Wood, and S. Molema, "All about me: Disclosure in online social networking profifiles: The case of Facebook," Comput. Human Behav., vol. 26, no. 3, pp. 406–418, 2010.

[8] K.Krombholz, D.Merkl, and E. Weippl, "Fake identities in social media: A case study on the sustainability of the

Facebook business model," J. Service Sci. Res., vol. 4, no. 2, pp. 175–212, 2012.

[9] E.Ferrara, O.Varol, C.Davis, F.Menczer, and A.Flammini,"The rise of social bots," Commun. ACM, vol. 59, no. 7, pp. 96–104, 2016.

[10] X.Wang et al., "Game theoretic suppression of forged messages in online social networks.