IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 21, 2019, accepted September 14, 2019, date of publication September 25, 2019, date of current version October 8, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2943747

Reputation-Based Approach Toward Web Content Credibility Analysis

SABA MAHMOOD[1], ANWAR GHANI[1], ALI DAUD[2], AND SHAHABODDIN SHAMSHIRBAND[3,4]

[1]Department of Computer Science and Software Engineering, International Islamic University Islamabad, Islamabad 44000, Pakistan
[2]Department of Computer Science and Artificial Intelligence, College of Computer Science and Engineering, University of Jeddah, Jeddah 23890, Saudi Arabia
[3]Department for Management of Science and Technology Development, Ton Duc Thang University, Ho Chi Minh City, Vietnam
[4]Faculty of Information Technology, Ton Duc Thang University, Ho Chi Minh City, Vietnam

Corresponding author: Shahaboddin Shamshirband (shahaboddin.shamshirband@tdtu.edu.vn)

ABSTRACT Web content credibility implies finding credible and correct information on the web. Recent studies have shown there is an increasing trend of users turning towards the web for searching information related to a variety of topics including health, stocks, education, politics to name few. Information credibility is a critical factor in these domains for the decision makers. There is no limitation on the authorship of those articles and content. One criterion for evaluating credibility is to check the authority or source of information. However, there are situations when wrong information flows from credible sources. There are various approaches towards credibility assessment, broadly categorized into human-based and computational approaches. Computational approaches utilizing machine learning based techniques are computationally expensive. Reputation based approaches overcome this, however the latest work fails to take into account issue of negative referrals and utilizes simple summation as the calculation structure making it more resilient to attacks. This paper put forth verified hypothesis of direct relationship of credibility to the expertise of entity. Authors proposed a Bayesian based approach using feedback in the form of interaction among the entities to compute their expertise level, thereby showing improved results in terms of Precision, Correlation and Mean Average Error. The experiments are performed on two different datasets, one of the dataset is developed from a survey as the part of the research study. The results from the two experiments show that the reputation ranks are independent of the pattern of ratings and density of data, unlike previous techniques whose results were limited by these factors. The proposed technique gives 27% and 18% more precise results for the two experiments respectively compared to the baseline. The correlation results are also significant in both experiments for the proposed technique with significant values of 0.39 and 0.87 showing a linear relationship between predicted and original data. The paper also discusses the reputation attacks and proposes counter measures to tackle these attacks through simulation results.

INDEX TERMS Web content credibility, information ranking, reputation systems, experts ranking.

I. INTRODUCTION

WEB is a decentralized repository of information, where anyone can contribute regardless of their knowledge and expertise. People are using information on the web from blogs, websites, e-magazine, e-books, e-journals, social networks on variety of subjects. There is no limitation regarding contribution of information over these channels as far as expertise and authority is concerned. Thus the credibility of information is dubious. Examples of people utilizing information from web for their daily decision making includes scenarios such as information related to medicine, illness and other health related issues [1], [2], news related to stocks, business and investments, daily political information, students browsing the web for their knowledge and subject information are some uses. Using the information from these sources users are taking decisions in their daily lives.For example whether to purchase a particular product from company A or company B, users look for recommendations. In such cases and several others credibility is a crucial virtue. Applications like Molesking [3], Answer Garden [4],TwitterWhoToFollow [5], ContactFinder [6], TweetCred [7] are all aiming towards finding

The associate editor coordinating the review of this manuscript and approving it for publication was Xueqin Jiang.

credible information in different domains. The development of these applications shows the importance of requirement of an application for predicting the credibility of information on the web.

One way of checking the credibility of information is by verifying the source of information. Techniques like Digital Signature, Public Key Infrastructure are utilized for the purpose. Authors are however taking into account the situations where false information might flow from credible sources. Researchers [8], [9] have proposed origin based approaches that address issue of evaluating content based upon its origin or source. The recent works towards finding credible information on web are all knowledge base driven, and the analysis is totally dependent on efficient techniques of retrieving facts. This is quite hard, since information is growing in size at rapid speed and keeping track of the facts in that information and building a knowledge base is not an easy approach. Factors [10] that contribute towards evaluation of content for credibility includes context, provenance, popularity, authority, bias, direct experience, incentive, recommendation, related resources. Web content credibility is judged by two major techniques that is credibility evaluation by human judgment and credibility evaluation by computations [11].

The research problem under discussion in this paper is the computationally feasible and efficient framework that could predict the credibility of the web content so a normal user is better equipped for decision making.

The authors have hypothesized that content can be regarded credible if it is reviewed by an expert. The past behavior of the entities and opinion of others referred as Reputation is utilized for expert ranking. Reputation based credibility assessment implies computing the credibility of the content by evaluating the reputation of the reviewer as a subject expert. The existing techniques need ground truth values for classification; the proposed technique is not limited by this and utilizes the history of interactions for predictions. The recent reputation based technique [12] has utilized user activity, user influence and sentiment value to find the reputation rank. User Activity is taken as the measure of the number of tweets by the user. Such tweets might include fake or fun tweets bearing no meaning. Additionally Sybil attackers will gain much benefit out of this parameter. A fake user can create number of positive fake statements to gain high sentiment value. These features are very critical since users can easily obtain thousands of followers by twitter marketers [13], [14]. The technique is highly prone to reputation attacks, carried out by fake identities.

Since this technique is domain specific i.e for twitter platform thus it is not capable to address other micro blogging or interactive platforms where likes, dislikes, upvotes, downvotes negative comments are present. The proposed framework in this paper aims to provide a generic platform applicable to variety of domains. The basic idea behind the proposed technique is to utilize both positive and negative interactions of a user.For example Alrubaian *et al.* [12] utilized number of followers but did not take into account the

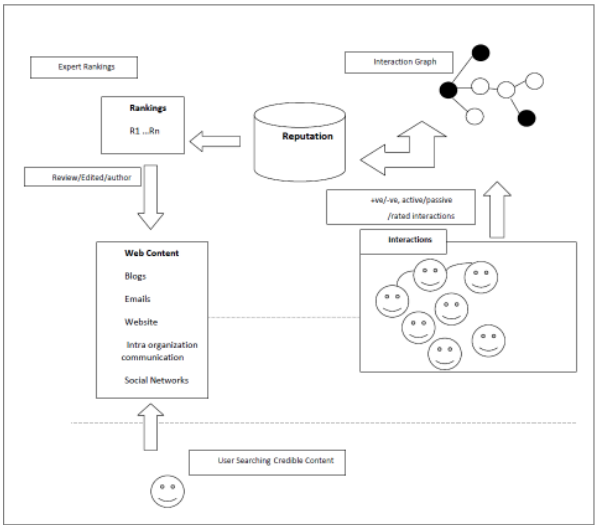

FIGURE 1. The problem domain.

number of unfollowings. It is also important to note that other technqiues that utilized reputation information along with machine learning(ML techniques), if utilized the proposed reputation mechanism in this paper can yield better results. However here in this paper authors are advocating that reputation based techniques are suitable due to lack of ground truth data and high cost associated with the ML techniques.

The experiments are performed on two kinds of dataset. One dataset represents intra organization interaction, the aim of experiment is to rank the employees in order of their expertise as perceived by the co-workers highlighting the functioning of the proposed algorithm. While the second experiment is performed on a dataset acquired through a survey data from the students. The aim of this experiment is again to generate ranking of students in order of their expertise. The two experiments show close correlation of results to the opinion of the users.

The major contributions of the research are:

- Content Credibility Framework based on reputation information with ability to take into account the positive, negative,active, passive behavior.
- Ranking users in order of their expertise gained through reputation information.
- Countermeasures to address reputation attacks of credibility framework.
- Development of a dataset through a survey.
- Experiments on two different datasets, to show how closely the ranking relates to the ranking produced by human judgment treated as ground truth.

The figure 1 shows pictorial representation of the problem under discussion.

The paper consists of section II that discusses literature and previous approaches. In section III the proposed approach is discussed. Evaluations and experiments are discussed in section IV. Section V concludes the paper.

II. RELATED WORK

Researchers have proposed approaches that are still at their infancy stage regarding the evaluation of the credibility of content present on the web. There are various dimensions to it as well. Shah *et al.* [11] discussed various parameters that can be helpful in doing so. They compared various techniques against those parameters that included categories of authority, accuracy, aesthetics, professionalism, popularity, currency, impartiality and quality. More recently applications like tweetcred [7] are available that try to find credible tweets and separates rumors and false news based on a credibility score. Such systems are restricted by context and trustworthiness of the scores.

A. CONTENT CREDIBILITY TECHNIQUES

Recent survey review [15] highlighted the topic from different aspects that includes the definition, relationship of trust with credibility, and various application domain areas. Content credibility assessment is broadly categorized into human-based and computational approaches.Hybrid approaches as discussed in the literature that combines human judgement with computation techniques and are assumed to produce better results.

Human-based approaches include techniques such as visual appearance, web layout, URL, date, personal belief, site familiarity, etc. Through various studies [16]–[19] it was found that users perceive information differently and thus assess the credibility of it differently and sometimes even the credibility of the source of information is neglected.

Computational approaches includes various techniques such as digital signature, collaborative filtering, machine learning approaches, semantic web and content ratings. A digital signature [20] is an electronic method of providing proof of the authenticity of a document. The signature verifies that the document is indeed created by the mentioned author. However, a situation where even inaccurate content flows from authentic and credible authors, cannot be addressed by this technique.

In collaborative filtering [21], [22], the content is evaluated by the peers and experts. Peer review systems for journals and publications are an example. The credibility rating systems are based upon the ratings given by users on the content [23], [24]. These systems, however, are unable to ascertain the expert level of the users giving the ratings. The semantic web [25] can judge the content using reasoners upon different criteria that cannot be addressed in other techniques. However, it cannot be effectively implemented since all the content is not present in an organized manner.

Recent work towards content credibility is machine learning(ML) based and most of them have targeted twitter as the application domain. ML techniques look into credibility as a classification problem. Olteanu *et al.* [26] proposed a system that looks into social network feature for predicting the credibility of information by utilizing Naïve Bayes classifiers and logistic regression. Although their results are 70% accurate however, the solution is computationally expensive for everyday use of the web content and users might be able to utilize such a service by purchasing it due to its cost. Authors of a latest work utilized [27] supervised learning technique to find credibility of information during high impact events. They used random forest for classification, results showed improved accuracy with the approach.ML techniques require ground truth values for classification, that is a major delimiting factor in the content credibility frameworks. Techniques like [28]–[31] have utilized supervised techniques to find credible tweets.Fairbanks *et al.* [32] has evaluated text based techniques with respect to structural approach in order to identify fake news. The results showed that structural analysis of articles can identify fake news compared to textual that is insufficient in finding credibility.

More recently authors [33] proposed a trust based solution to find about the trustworthiness of the information resources. This technique employed weighted average method whereby weights are dynamically assigned by moving weighted average and ordered weighted average. Authors have compared their technique with the recent model for Twitter data [12]. The research has shown results that reputation based information is more accurate in predicting the credibility of information as compared to machine learning(ML) approaches of Naive Bayes and Logistic Regression. The ML techniques are time-consuming and are dependent upon feature selection. This shrinks down to the comparison between the methods of reputation system adopted by the previous technique and one proposed in this paper. The earlier work [12], [34] utilized the user influence, user engagement and user sentiment information for reputation ranking. In the following sections, the authors have given detail background of the reputation systems in general and the baseline approach of reputation based credibility in detail.

It is to be noted that most of the previous work in the domain of content credibility has utilized supervised learning approaches of classifying data as credible or uncredible. The proposed approach has utilized the reputation information of the user, the approach can avoid the computational cost of the machine learning approaches and the unreliability and inconsistency associated with human judgment. However, it is also important to note that the techniques utilizing reputation information along with certain ML techniques also adopted adhoc reputation algorithms, thus the proposed work has introduced the concept of the technique based on strong mathematics and statistics, that if utilized in combination with other techniques can yield better results. In this work however we are advocating that given limited computation resources, and the ever growing nature of information only reputation based technique can be cost effective and reliable technique opposed to major competitior that is the ML techniques that rely heavly on feature selection and ground truth data.

B. REPUTATION SYSTEMS

Existing reputation systems are categorized either into centralized and distributed system. The centralized systems have

a central entity that records the experience of all the entities in the network, while a distributed one does not has a central store instead all the entities keep a record of their interactions with other entities. Systems like [35]–[37] are based upon distributed reputation mechanism. The distributed systems have to address the issue of propagation of reputation information in addition to the method of calculating the reputation scores. Author's area of interest is the calculation of reputation scores since it is the basis of comparison with respect to other models. Most common ways of calculations are; Summation, Average, Weighted Average and Bayesian.

Summation, the simpler one where positive and negative scores are summed up separately for each individual. The method is employed by eBay [38] and is the easiest method to understand and implement.

Reputation systems employed by Amazon and Epinions [35] use averaging whereby the scores are averaged to present the reputation score of all the entities. In Weighted Average method, the average reputation score is multiplied by a rating factor. The rating factor depends on the age of score, trustworthiness. Bayesian systems [39] are based upon computing reputation score by calculating and updating the beta probability density function. The major advantage of the approach is that it provides the theoretically sound basis for computation of the scores. There is no known disadvantage of this approach.

More recently researchers have utilized PageRank based reputation models [40] to rank volunteered geographic information(VGI) system, thus considered as a new calculation structure of the reputation systems.Calculations can be performed in a centralized manner or decentralized manner for example in the case of peer to peer systems. Both structures have their pros and cons.

Another reputation structure utilized by researchers is based on Normal Distribution Reputation(NDR) [41], irrespective of the domain various reputation models have been widely been researched and utilized in online market places, p2p systems for example eigentrust [42], regret [43]. It is worth mentioning that Bayesian reputation systems, addressed the limitations of the most popular eigen reputation systems [44].

Reputation approach towards Content Credibility technique [12], utilized reputation for credibility assessment of twitter data. The reputation model utilized engagement, popularity and sentimental information for calculating the reputation rank of the user. This work has utilized simple summation as the calculation method that is centralized in structure. The authors of this work compared the result with other machine learning techniques of naive bayes and logistic regression. The results in the form of precision and recall showed improvement with reputation based technique also overcoming computational cost associated with training of data in case of machine learning methods. The technique is however unable to take into account the negative relationships and are prone to reputation attacks.Authors have treated this recent technique as the baseline technique along with other reputation baselines including Page Rank and NDR based approaches for comparison against the results produced through the proposed technique.

C. ATTACKS ON REPUTATION SYSTEMS

Most popular attacks on reputation systems include:

1) SELF PROMOTION/SYBIL ATTACK

The user can augment his/ her reputation with false information. This attack is coupled with the Sybil attack. Systems that consider only positive feedbacks are more prone to this attack. The attacker fabricates false positive interaction about itself. Self Promotion is achieved through Sybil attack whereby a user can start creating fake accounts in order to give positive feedbacks to his own account for increasing reputation [45]. The term Sybil emerged from a book after a woman who had dissocative identity problem.A sybil account can create malicious accounts to subvert the reputation or ranking of a legal user, these days internet bullying, etc all fall in this type of attack. Recently these attacks have been under discussion in the context of social media specifically twitter [46], authors of this work have utilized regression model to predict about the user's profile.

2) SLANDERING

Malicious users could give negative feedback for the users who are positive, thereby affecting the reputation of deserving users. Effect of a single slandering node is less, however it can have an impact when nodes collued to damage positive reputation of a node. Typical defense mechanism have penalty mechanisms once a slandering node is identified. Attaching node with authentic transaction/interaction can also act as the preventive measure.

3) WHITE WASHING

This attack is also called a self-service attack. In this attack, the malicious node starts its original behavior after gaining good reputation initially. The systems that rely on long historical data are more prone to this type of attack. The malicious behavior can be of a sybil or slandering.

Systems like Sybil Guard [47], SumUp [48], Sybil Limt [49] have proposed techniques to reduce number of attacks. However they are not specifically designed for expert ranking scenario. SybilGuard is based on the ''social network'' among user identities, where an edge between two identities indicates a human-established trust relationship. Malicious users can create many identities but few trust relationships.

One research towards fighting against Sybil attacks in expert ranking systems MHITS [50] utilizes SumUp algorithm. In this system, the nodes are removed through the Sum Up strategy before the ranking process. SumUp is a sybil resilient online content rating system that uses the trust network among users to defend against sybil attacks. It uses the concept of max-flow.

Another stream of research in the field of defense mechanism is related to getting truthful feedback from the users is the peer prediction method [51], [52]. They provide proper reward system to agents who provide truthful reports of other agents in a nash equilibrium manner. These systems have theoretically tried to address the Sybil attacks by monitoring the agent behavior related to incurring of cost by giving opinions.

III. PROPOSED APPROACH

In our daily life, we always seek an expert for an advice and recommendation. Thus the authors have hypothesized that content evaluated by an expert can be regarded as credible. The results by Kakol *et al.* [53] also verify this hypothesis.However, to further test this hypothesis a survey was conducted. The survey was conducted by a third party without the involvement of the authors. Authors had no knowledge about the participants. Similarly participants were also unaware of the identity of the authors and the research thus, a double-blinded survey. The sample was selected randomly, the participants were informed about the research and they were allowed to leave the questionnaire unfilled if they do not have the consent.

The survey was conducted with 40 students of the University of Engineering and Technology Peshawar. They were asked to rank different tools that are utilized for checking the credibility of the web information. These tools are

- The source of information
- Popularity of information.
- Expert analysis of information
- Recommendation by others

The ranking ranged from 1to 4.With1 as most important and 4 as least important. The question was asked in the context of blogs and social media. Out of 40 participants 34 participants filled the forms correctly. These 34 participants understood the questions and answered according to the given choices. The results showed that after the source of information, expert analysis is an effective tool for information credibility assessment. Thus mathematically we can write it as follows:

let $U = u_1, u_2, \cdots u_n$ presents set of users.

$R = r_1, r_2, \cdots r_n$ presents reputation information of the users in the set U.

Then $Max(R) = E$ represents the Expert User. Thus Opinion of E is directly proportional to credibility (C) of Web information, presented mathematically as follows;

$$E \propto C \tag{1}$$

Thus content associated with an expert is considered credible. The content credibility framework is a four layered architecture. The bottom layer of interactions is responsible to categorize types of interaction that users had. The reputation layer utilizes this interaction information to compute the reputation rank of the users. The second last layer ranks the users in order of expertise given reputation rank information. Finally the top layer associates the credibility according to the expertise of the user. The architecture is closed one since a layer is dependent on the immediate layer below. Given in the figure 2 is the four layered architecture.

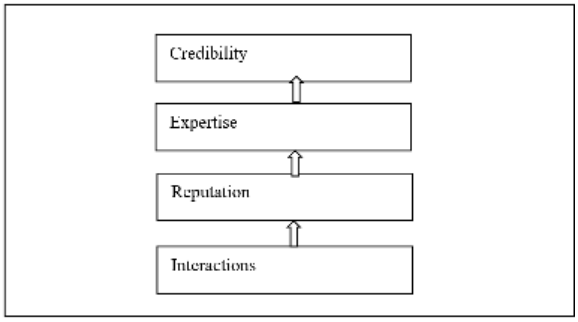

FIGURE 2. The layered architecture.

A. INTERACTION LAYER

In case of social network an interaction can be a post, private message, like, tweet, retweet, following, upvotes, downvotes, rankings etc. It is a generic term can be adopted according to the target platform. Interactions as discussed in the work by Nepal *et al.* [54] can be positive or negative/active passive. For example an interaction between two users A, B through a post. User A posts a message, user B interacts by positively accepting it, shows a positive relationship however this could be negative as well. Another user C chooses not to respond thus remains passive. In another scenario such as in a large organization emails, employee rankings, can all be treated as an interaction. The interaction data can be in the form of text, ratings, votes.

The interaction layer is responsible for categorizing the interaction. If the interactions are present in the form of text, its sentimental value is used to categorize as either positive or negative. While if the interaction is a value, the cutoff decides its category as given in the algorithm 1 from lines 8–9. The layer is responsible to calculate total number of positive and negative interactions.

B. REPUTATION LAYER

Reputation is defined as ''Overall quality as seen and judged by users'' according to the Merriam-Webster's [55] online dictionary. The past behavior of the entities and opinion of others is utilized to find the reputation. The opinion is based on the past history of interactions.

1) MATHEMATICAL MODEL

The reputation layer is based on Bayesian [56] based reputation algorithms which are easy to understand and can be easily applied to wide variety of application domains. Bayesian based systems are binomial or multinomial. Binomial Bayesian reputation systems apply to the binary state space i.e. Bad, Good, that reflects corresponding performance

Algorithm 1 Interaction Categorization

```
1: if Interactions = Text then
2:     result ← Senti(Text, i)
3:     if result = positive then
4:         p ← p + +;
5:     else
6:         n + +
7:     end if
8: else if Interaction = value then
9:     Enter cutoff
10:    if value > cutoff then
11:        p + +
12:    else
13:        n + +
14:    end if
15: else if Interaction = explicit then
16:
17:    if positive then
18:        p + +
19:    else
20:        n + +
21:    end if
22: end if
```

of a service entity. The Beta distributions is a continuous distribution functions over a binary state space indexed by the two parameters alpha and beta. Bayesian based reputation systems have shown promising results and is easy to implement. Bayesian reputation systems have sound basis in classical statistics that makes them effective and adaptable to various contexts.

The beta probability density function can be used to represent probability distributions of binary events. This provides a sound mathematical basis for combining feedback and for expressing reputation ratings. The mathematical analysis leading to the expression for posteriori probability estimates of binary events can be found in many text books on probability theory,e.g. [57] Casella Berger 1990 p.298, we will be using only the results here.Posteriori probabilities of binary events can be represented as beta distributions. The beta-family of probability density functions is a continuous family of functions indexed by the two parameters alpha and beta. The expected value of the distribution is given by

$$E(v) = \frac{\alpha}{\alpha + \beta} \tag{2}$$

A node in a social network might be having positive interactions in the form of number of tweets or number of comments received on a facebook post in a particular context or subject. There can be negative interactions as well.

Let the positive interactions be represented by alpha and negative interactions by beta. Let 'x' represents a particular context and 'A' represents number of activities in a particular context. 'M' is the total number of participants/nodes in the network. Thus we can compute the expert node in a particular

Algorithm 2 Reputation Based Expert Rank

```
1: Load Dataset
2: loop
3:     Compute no. of positive interactions p for node i
4:     Compute no. of negative interactions n for node i
5:     Compute Expected value for node i as
6:     Ev ← p + 1/p + n + 2
7:     Let T represent total number of nodes
8:     Compare Reputation of node i with T-i nodes
9:     max ← i
10: end loop
11: Compute max as Expert
```

context 'A' from equation 2, let's suppose y, y_1 represent number of outcomes of alpha and beta respectively, that means after every y outcome we can expect y_1 outcome. In our framework, let p represent the observed number of outcomes for y and n represent observed number of outcomes for y_1, then we can derive following equations.

$$\alpha = y + 1 \tag{3}$$

$$\beta = y1 + 1 \tag{4}$$

$$E(v) = \frac{y + 1}{y + y_1 + 2} \tag{5}$$

Substituting the number of outcomes for y and y_1, we get

$$E(v) = \frac{p + 1}{p + n + 2} \tag{6}$$

The Expert Reputation of a single node is then given by

$$E = \sum_{i=1}^{M-1} E(v) \tag{7}$$

here 'E' indicates the expected value(Reputation) of a node to behave in a particular context.

Given below is the detailed algorithm 2 utilized by the reputation layer, where 'i' is the node, 'T' represents total number of nodes, Ev represents expected value of node i.

In the Lines 3, 4 the interactions are categorized into positive or negative domains. Line 6 uses beta probability expected value to generate the expert rank of the node. Lines 7-9 compares the expert rank of a particular node to the rest of nodes in the network.

C. DEFENSE MECHANISM

The literature review discussed various attacks to the reputation systems. Thus the reputation layer of the proposed framework has also incorporated defense algorithms to encounter the attacks.

The proposed scheme based upon Bayesian reputation system is able to prevent the slandering, sybil and whitewash attack.

Algorithm 3 Slandering Attack Defense

```
1: loop
2:   if interaction == negative then
3:     N++
4:   else
5:     P++
6:   end if
7: end loop
8: loop
9:   LetX ← i
10:  if X.N > 10 then
11:    Let s represent slandering node thus s ← X
12:    Filter out s
13:  end if
14: end loop
```

Algorithm 4 Sybil Attack Defense

```
1: Load Dataset
2: loop
3:   Compute no. of positive interactions p for node i
4:   Let z represent node having positive interactions with
     node i
5:   Compute no. of positive and negative interactions of
     node z
6:   Compute Expected value for node z as
7:   Ev(z) ← p + 1/p + n + 2
8:   Compute no. of negative interactions n for node i
9:   Compute Expected value for node i as
10:  Ev(i) ← p + 1/p + n + 2
11:  Ev(i) ← Ev(i) ∗ Ev(z)
12: end loop
```

Algorithm 5 White Wash Attack Defense

```
1: Find time t node i interacting node n-i
2: if t = 0 then
3:   Latest interaction only node i,n-i
4: else
5:   All interactions node i,n-i
6: end if
```

1) SLANDERING ATTACK DEFENSE ALGORITHM

In case of slandering attack if the user is giving false negative feedback, after a certain amount of interaction this reflects that the feedback is malicious. Since a true user will not have any further interaction, after the negative interactions. Such users can also be filtered out thus preventing the slandering attack. Thus the attack resistant reputation algorithm 3 is given below.

2) SYBIL ATTACK DEFENSE ALGORITHM

The scheme ranks every user in the network according to the reputation value. The reputation value is calculated by the feedback given by all the rest of the members. Thus if a Sybil attack is launched with fake ids, not all of the users of the network would give good feedback about them or in other words they might encounter isolation. This also holds true for feedback in the form of opinions or through text. The proposed scheme weighs the reputation of the node according to the reputation of the interacting node, lines 10–11 of algorithm 4.

3) WHITE WASH ATTACK DEFENSE ALGORITHM

The whitewash attack is countered by having a time factor with the reputation value, thus older feedback is given less importance than the recent feedback values. Related concept has also been discussed in the dishonesty detector based on historical information [58]

D. EXPERT LAYER

The expert layer of the framework is responsible for generating expert ranking according to the reputation score given by the Reputation Layer.

Researchers [59] proposed a technique whereby a user is ranked as an expert based on reputation information of co-existing users. The calculation structure adopted by this technique is simple summation.

Reputation mechanisms already employed for expert ranking are ad-hoc based. The proposed technique has sound mathematical basis. The Bayesian-based reputation is a most effective mechanism in comparison to other techniques. Here the quality of interactions is utilized to find the reputation score of the user.

E. CREDIBILITY LAYER

Using Equation 1, this layer is responsible to associate expert rank of the user with the content, that was either produced/ reviewed/ edited by him. The user interacts with this layer to find the credibility of information. Thus this layer can be considered equivalent to the user interface layer.

IV. EVALUATION

Authors conducted two experiments one experiment is carried out to evaluate the effect of categorizing the type of interactions upon rankings and its correlation with actual rankings.

In the first experiment, a comparison is performed to the latest technique of content credibility [12]. The second experiment is conducted on a dataset extracted through a survey. The dataset is developed to have ground truth values that can be used for comparison purposes.

A. PERFORMANCE EVALUATION METRICS

The estimated reputation values generated through the proposed model are compared to the real values so as to evaluate the ability of the model in predicting rankings close to real rankings. Two performance indicators [60] within this context are reported, the average absolute error [61] between real and predicted values and the correlation between real and the

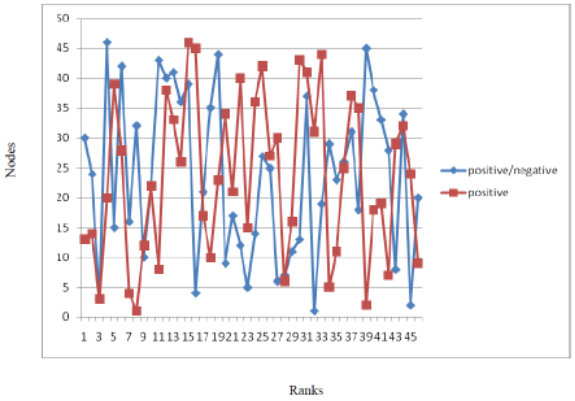

FIGURE 3. Interaction overlap graph.

predicted values.In addition to these indicators, Precision is also utilized. Precision is given by(tp stands for true positives, fp stands for false positives)

$$Precision = \frac{t_p}{t_p + f_p} \tag{8}$$

Precision means the probability with which the algorithm accurately generates the ranking i.e. it truly ranks an expert.Recall and F Measure is also calculated for the baseline and the proposed technique. Recall is given by

$$Recall = \frac{t_p}{t_p + f_n} \tag{9}$$

In order to evaluate the defense mechanism of the proposed model, change in reputation rank is used as a performance indicator.

B. EXPERIMENT 1

We conducted an experiment on Dataset [62] that is of intra-organizational network, where the interactions are weighted on the scale from 0-5 that defines the frequency of advice requested.

0 : I Do Not Know This Person
1 : Never
2 : Seldom
3 : Sometimes
4 : Often and
5 :Very Often

First, a ranked list of nodes was generated using only the positive interactions. In the second instance using the cutoff, both positive and negative interactions were utilized in calculations. The two ranked list differed from each other as can be seen in the given figures and the tables.

The table 1 shows the ranked list through two methods in ascending order, the top nodes are at the bottom of the table.

The Mean Average Error of top3 nodes given in the table 2 was calculated to find how close they reflect original opinions. The results showed that ranked list generated through both positive and negative interactions were closer to mean as compared to list generated through only positive interactions. The figure 3 shows that both ranked lists are not overlapping with very few instances. Thus combining categories of interaction has an effect on rankings. Since some nodes that are ranked higher due to only positive interaction lose their rank when both positive and negative interactions are utilized.

TABLE 1. Ranking with(+ve/−ve) and without(+ve) proposed reputation scheme.

S.No	+ve/-ve	Positive	S.No	+ve/-ve	Positive
1	30	13	24	14	36
2	24	14	25	27	42
3	3	3	26	25	27
4	46	20	27	6	30
5	15	39	28	7	6
6	42	28	29	11	16
7	16	4	30	13	43
8	32	1	31	37	41
9	10	12	32	1	31
10	22	22	33	19	44
11	43	8	34	29	5
12	40	38	35	23	11
13	41	33	36	26	25
14	36	26	37	31	37
15	39	46	38	18	35
16	4	45	39	45	2
17	21	17	40	38	18
18	35	10	41	33	19
19	44	23	42	28	7
20	9	34	43	8	29
21	17	21	44	34	32
22	12	40	45	2	24
23	5	15	46	20	9

TABLE 2. Mean Average Error(MAE) of Top Nodes.

Ev Set1 (+ve/-ve)	Ev Set2 (+ve)	MAE Set1	MAE Set2
3.94	3.6	0.14	0.2
3.93	3	0.12	0.8
3.89	3.72	0.07	0.09

Ev = Expected Value, Positive = +ve, Negative = -ve

1) COMPARISON TO BASELINE

An experiment is performed on the dataset to compare the proposed technique with the web content credibility technique [12] treated as baseline1 technique. The authors have also compared their technique with a reputation calculation structure based on PageRank [40] considered as baseline2 and Normal Distribution based reputation structure NDR [21] as baseline3. Metric for comparison is user opinion. Mean Avergae Error(MAE) of the proposed ranking to the actual user opinion was found. The experiment revealed following results.

The top 3 nodes obtained from the baseline1 showed MAE of 0.5, 0.4 and 0.07. Comparison to the proposed technique showed that the MAE is much lesser. This shows that the proposed technique is more effective in presenting the actual opinion of the user. Thus the ranking of the proposed technique is more accurate in showing the credibility level of the content associated with them. Similarly precision and

TABLE 3. Comparison of ranked lists MAE1 Experiment1.

MAE[baseline1]	MAE [Proposed]
0.5	0.14
0.4	0.12
0.07	0.07

TABLE 4. Comparison of ranked lists MAE2 Experiment1.

MAE[baseline2]	MAE [Proposed]
0.3	0.14

TABLE 5. Comparison of ranked lists MAE3 Experiment1.

MAE[baseline3]	MAE [Proposed]
0.16	0.14

TABLE 6. Comparison w.r.t Precision, Recall, Correlation Experiment 1.

Metrics	Baseline1	Baseline3	Proposed
Precision	0.042	0.021	0.06
Recall	0.11	0.05	0.3
F Measure	0.06	0.029	0.1
Correlation	0.24	-0.076	0.39

correlation tests also revealed that proposed technique performs better than the baseline.

The performance of baseline2 against these metrics could not be reported due to almost zero relationship.Baseline3 also shows improved MAE results for the proposed technique.

C. EXPERIMENT 2

Using the information through the filled survey forms we developed a dataset of the students of a class in graph format. The survey was double blinded, it was conducted with the students of University of Engineering and Technology Peshawar. The sample was selected randomly, the participants were informed about the research and they were allowed to leave the questionnaire unfilled if they do not have consent. The participants were informed verbally as well as through a note on questionnaire. An edge shows the friendship relationship. The weight of the edges represents the interactions of the students. The weights are assigned based upon the ratings provided by the students for their friends regarding the interactions related to subject knowledge. The weights are scaled in the range of 1-5,i.e.

1 : Nil
2 : Fair
3 : Good
4 : Very Good
5 :Excellent

This dataset is undirected where nodes represent the students and the edges represents the rated expert friendship relationship among them. This dataset has been taken as a special scenario where each node has three friends showing equal number of links for each node. Such special scenario would highlight the shortcomings of previous link based reputation based on the PageRank thereby showing the capability of the proposed technique. This dataset is undirected with 40 nodes.

TABLE 7. Rank List2.

Nodes	EV	Nodes	Avg	Nodes	EV	Nodes	Avg
13	6.6	13	0.11	21	1.7	40	0.035
2	5.57	2	0.099	4	1.69	3	0.034
37	5	8	0.087	30	1.69	12	0.03
17	4.99	6	0.08	3	1.66	21	0.03
32	4.9	17	0.08	12	1.66	30	0.029
8	4.2	32	0.08	20	1.66	26	0.027
6	4.15	37	0.077	34	1.66	28	0.026
25	4.06	10	0.058	26	1.65	34	0.025
38	4	7	0.056	28	1.65	27	0.023
14	3.33	38	0.054	16	1.63	31	0.023
10	3.3	5	0.05	23	1.63	16	0.02
7	3.26	14	0.05	27	1.63	20	0.02
33	3.26	25	0.05	31	1.63	23	0.02
40	3.1	33	0.044	19	0.85	19	0.017
35	2.52	35	0.04	9	0.833	9	0.0158
5	2.5	39	0.039	22	0.833	22	0.014
39	2.5	4	0.038	24	0.66	24	0.003
1	0	1	0	18	0	18	0
11	0	11	0	29	0	29	0
15	0	15	0	36	0	36	0

EV = Expected Value

Given in the table 7 is the ranked list of students according to the proposed algorithm. The original dataset presents the ground truth. The neighbors or close relations possess the same level of knowledge. Thus the ratings given by students for their friends are considered the ground truths. The average mean is calculated for every student. Given below is the rank list of nodes according to average weights in comparison with the rank list generated by the proposed algorithm.

1) COMPARISON TO BASELINE

The comparison is made to calculate precision-recall.

Precision@k is an important metric in the field of information retrieval to find the percentage of accurately discovered documents. K stands for a certain number of instance. For example, if k is 5 it implies a number of accurate discovered items in top 5 entries in comparison to the ground truth values. We have calculated P@3, which came as 0.66. Overall precision is 0.32 while of baseline it is 0.05.

Correlation is calculated of the two ranked list, to find if the ranking produced by the proposed algorithm closely correlates to the human judgment. The correlation of top5 ranked list is 0.87 for the proposed technique compared to the baseline1 technique where correlation is 0.55. The correlation result shows that the expert ranked nodes positively relates to human judgment. Figure 4 shows this correlation, whereby the rank list generated though proposed algorithm closely relates to the rank list according to average weights.

Experiment on the dataset to compare the proposed technique with the web content credibility technique [12] treated

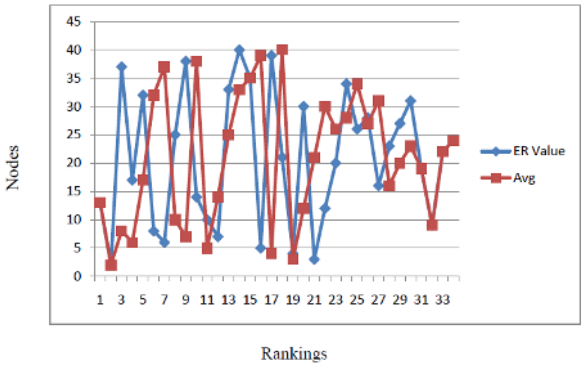

FIGURE 4. Comparison of estimated ranked list to real ranking w.r.t MAE.

TABLE 8. Comparison of Ranked Lists MAE3 Experiment2.

MAE[baseline1]	MAE [Proposed]
1.2	0.2
0.2	1.2
0.86	0.46

TABLE 9. Comparison w.r.t. Precision, Correlation Experiment 2.

Metrics	Baseline1	Baseline3	Proposed
Precision	0.05	0.05	0.32
Recall	0.02	0.09	0.4
F Measure	0.028	0.064	0.36
Correlation	0.55	-0.13	0.87

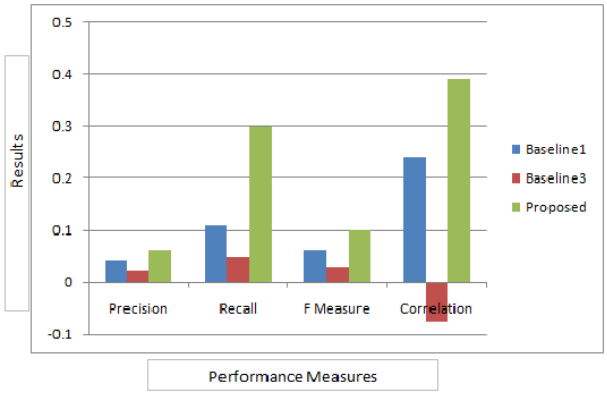

FIGURE 5. Comparison of Ranked lists w.r.t Precision, Correlation Experiment1.

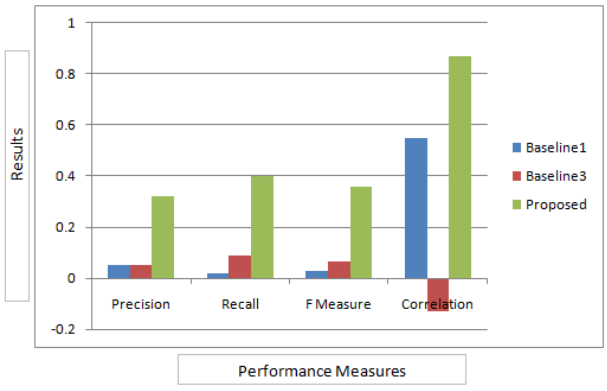

FIGURE 6. Comparison of Ranked lists w.r.t Precision, Correlation Experiment2.

as baseline technique was carried out. Metric for comparison is user opinion. Mean average error(MAE) of the proposed ranking to the actual user opinion was calculated. The experiment revealed results shown in the table 8.

The top 3 nodes obtained from the baseline showed variance of 1.2, 0.2 and 0.86. Comparison showed that the mean error is much lesser for proposed technique except for the second ranked node. This shows the proposed technique is more effective in presenting the actual opinion of the user. Thus the ranking of the proposed technique is more accurate in showing the credibility level of the content associated with them. Figure 5 and Figure 6 shows the comparison of techniques with repect to correlation, precision and MAE.

D. ANALYSIS OF RESULTS

The results from the two datasets show that the mean average error for the proposed technique is lesser as compared to the previous baseline models. In Experiment1, working of the proposed model is evaluated and it was found that inclusion of both positive and negative types of interactions yields lesser mean average error as compared to the scenario when only positive interactions are utilized as is done by the previous technique where only positive interactions are counted towards a user's popularity/engagement.

The second experiment conducted on dataset compiled through a survey showed again that mean average error for the proposed technique is lesser, compared to the baseline techniques. In these experiments Precision and Correlation results also supported effectiveness of the proposed technique. The results of the techniques are however not promising in experiment1, but the performance of proposed technique is still better. The dataset from experiment1 is dense as compared to second dataset that is sparse. The percentage difference of precision of the baseline and proposed technique is 27% as compared to experiment1 where it is 18%, giving us an insight that in case of sparse data, the proposed algorithm has almost 27% more precise results while in case of dense dataset it is 18% more precise only.Similar trend is found for Recall values. Performance of baseline2 in terms of precision and correlation are too poor to report. In the case of baseline3 again the results for the proposed techniques are better following almost same trend. The proposed technique works better with sparse datasets, that actually is the problem scenario, where it is not necessary that a particular node might be having interactions with every other node and vice versa. Baseline2 that is pagerank based looks at the number of connections regardless of the quality of the connections. Baseline3 that is NDR based also performs less, since it produces less accurate results for sparse data then for dense [21]. Observing results from Experiment1 alone, it is evident that

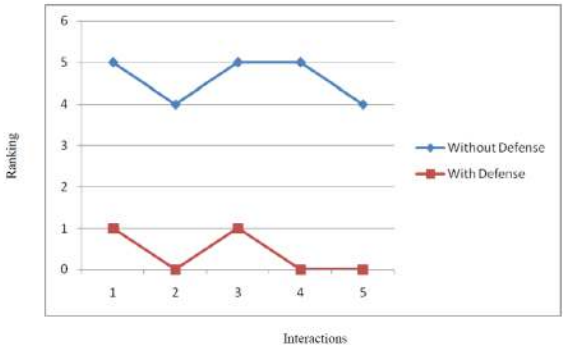

FIGURE 7. Simulation to test defense mechanism.

results from NDR based technique are less then baseline1, while for experiment2 performance of NDR(baseline3) is better than baseline1, due to the fact that although experiment1 dataset is dense, the ratings are distributed normally as compared to dataset from experiment2 where ratings are on higher or lower end and not normally distributed.

Thus we can summarize the findings that the proposed technique produces better results regardless of density of the dataset and pattern of ratings, while other two techniques appear to be dependent on these two factors. The pagerank based baseline has already produced poor result due to inability of taking the quality of interactions and negative referrals into account.

We have also calculated the correlation for the baseline techniques and the proposed against the null hypothesis i.e. ''There is a linear relationship between original data and predicted data for all the four techniques''. The null hypothesis stands true for the three techniques the baselines(baseline1, baseline3) and the proposed technique. However the hypothesis is rejected for the pagerank(baseline2) based technique. The significance value for baseline3 is negative that shows a negative relationship whereby if original data is showing ranking in highest to lowest order the predicted data is mostly in opposite direction. The results of proposed technique are significant with value of 0.39 for experiment 1 and 0.87 for experiment2.

E. SIMULATION TO TEST DEFENSE

Authors performed simulation to test the defense mechanism against Sybil attack. For this some new nodes(fake nodes) were introduced in the dataset from Experiment1. These fake nodes were involved in giving positive feedback for node 8. Reputation rank calculated by proposed algorithm without defense mechanism gives it a high rank. In order to verify the defense mechanism, the reputation rank of fake node is found. Since the fake nodes did not do interactions with other nodes, thus resulting rank was low. The new rank of node 8 is then calculated by weighing it by the reputation of the fake nodes and real nodes. The series at the top of the graph in the figure 7 shows node 8 rank due to fake nodes, while the series at the bottom of the graph shows node 8 after utilizing the defense mechanism in the proposed algorithm. This clearly shows that the node looses its fake reputation due to the proposed defense mechanism.

To check the functionality against the slandering attack, the interactions of node 8 towards a specific node 1 were manipulated. This is done by inducing more negative interactions. As a threshold if number of negative interactions exceed 10, the node 8 is blacklisted and its interactions towards node 1 and other nodes are not recorded.

V. CONCLUSION

The paper has proposed a technique for assessment of the credibility of the content present on the web. The proposed technique is reputation based since it ranks a user as an expert based on the past interactions. The technique has its grounds on the classical concept of the relationship of credibility to the credibility of the author. Since the scenario under consideration has unpopular users thus it was needed to identify their expertise. The authors compared the results with three baselines. These existing techniques used simple summation based calculation structure, pagerank based structure and NDR structures for generation of reputation ranks. The proposed technique is bayesian based and takes into account all kinds of interactions producing results that are independent of rating pattern and density of data. The proposed technique has ability to solve negative refferal problem. The results from the two experiments carried out on two different datasets support the better performance of the proposed technique interms of MAE, Precision, Recall with significance of Correlation. The comparison showed that proposed algorithm closely reflects human perception regarding content credibility. In this paper a brief evaluation of the defense mechanism is also given, in future it is intended to perform more evaluations of defense mechanisms in different scenarios.

REFERENCES

[1] G. Ferreira, A. C. Traeger, G. Machado, M. O'Keeffe, and C. G. Maher, ''Credibility, accuracy, and comprehensiveness of Internet-based information about low back pain: A systematic review,'' *J. Med. Internet Res.*, vol. 21, no. 5, p. e13357, 2019.

[2] A. Z. Nabo ny, B. Balcerzak, and A. Wierzbicki, ''Automatic credibility assessment of popular medical articles available online,'' in *Proc. Int. Conf. Social Informat.* Saint-Petersburg, Russia: Springer, 2018, pp. 215–223.

[3] P. Avesani, P. Massa, and R. Tiella, ''A trust-enhanced recommender system application: Moleskiing,'' in *Proc. ACM Symp. Appl. Comput.*, 2005, pp. 1589–1593.

[4] M. S. Ackerman and T. W. Malone, *Answer Garden: A Tool for Growing Organizational Memory*. New York, NY, USA: ACM, vol. 11, nos. 2–3. 1990.

[5] P. Gupta, A. Goel, J. Lin, A. Sharma, D. Wang, and R. Zadeh, ''WTF: The who to follow service at twitter,'' in *Proc. 22nd Int. Conf. World Wide Web*, 2013, pp. 505–514.

[6] B. Krulwich, C. Burkey, and A. Consulting, ''The contactfinder agent: Answering bulletin board questions with referrals,'' in *Proc. AAAI/IAAI*, vol. 1, 1996, pp. 10–15.

[7] A. Gupta, P. Kumaraguru, C. Castillo, and P. Meier, ''TweetCred: Real-time credibility assessment of content on Twitter,'' in *Proc. Int. Conf. Social Informat.* Barcelona, Spain: Springer, 2014, pp. 228–243.

[8] J. M. Kleinberg, ''Authoritative sources in a hyperlinked environment,'' *J. ACM*, vol. 46, no. 5, pp. 604–632, 1999.

[9] M. S. Fox and J. Huang, ''Knowledge provenance: An approach to modeling and maintaining the evolution and validity of knowledge,'' Enterprise Integr. Lab., Univ. Toronto, Toronto, ON, Canada, EIL Tech. Rep., 2003.

[10] Y. Gil and D. Artz, "Towards content trust of Web resources," *Web Semantics, Sci., Services, Agents World Wide Web*, vol. 5, no. 4, pp. 227–239, 2007.
[11] A. A. Shah, S. D. Ravana, S. Hamid, and M. A. Ismail, "Web credibility assessment: Affecting factors and assessment techniques," *Inf. Res.*, vol. 20, no. 1, pp. 1–20, 2015.
[12] M. Alrubaian, M. Al-Qurishi, M. Al-Rakhami, M. M. Hassan, and A. Alamri, "Reputation-based credibility analysis of Twitter social network users," *Concurrency Comput., Pract. Exper.*, vol. 29, no. 7, p. e3873, 2017.
[13] G. Stringhini, G. Wang, M. Egele, C. Kruegel, G. Vigna, H. Zheng, and B. Y. Zhao, "Follow the green: Growth and dynamics in twitter follower markets," in *Proc. Conf. Internet Meas. Conf.*, 2013, pp. 163–176.
[14] C. De Micheli and A. Stroppa, "Twitter and the underground market," in *Proc. 11th Nexa Lunch Seminar*, vol. 22, 2013, pp. 1–43.
[15] M. Alrubaian, M. Al-Qurishi, A. Alamri, M. Al-Rakhami, M. M. Hassan, and G. Fortino, "Credibility in online social networks: A survey," *IEEE Access*, vol. 7, pp. 2828–2855, 2019.
[16] M. J. Metzger, "Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 58, no. 13, pp. 2078–2091, 2007.
[17] M. J. Metzger and A. J. Flanagin, "Credibility and trust of information in online environments: The use of cognitive heuristics," *J. Pragmatics*, vol. 59, pp. 210–220, Dec. 2013.
[18] R. Savolainen, "Judging the quality and credibility of information in Internet discussion forums," *J. Assoc. Inf. Sci. Technol.*, vol. 62, no. 7, pp. 1243–1256, 2011.
[19] F. Calefato, F. Lanubile, M. C. Marasciulo, and N. Novielli, "Mining successful answers in stack overflow," in *Proc. IEEE/ACM 12th Work. Conf. Mining Softw. Repositories (MSR)*, May 2015, pp. 430–433.
[20] N. Asokan, V. Shoup, and M. Waidner, "Optimistic fair exchange of digital signatures," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 4, pp. 593–610, Apr. 2000.
[21] F. Loll and N. Pinkwart, "Using collaborative filtering algorithms as elearning tools," in *Proc. 42nd Hawaii Int. Conf. Syst. Sci. (HICSS)*, 2009, pp. 1–10.
[22] M. A. Parsons, R. Duerr, and J.-B. Minster, "Data citation and peer review," *Eos, Trans. Amer. Geophys. Union*, vol. 91, no. 34, pp. 297–298, 2010.
[23] A. V. Pantola, S. Festin, and F. Salvador, "Rating the raters: A reputation system for wiki-like domains," in *Proc. 3rd Int. Conf. Secur. Inf. Netw.*, 2010, pp. 71–80.
[24] X. Liu, R. Nielek, A. Wierzbicki, and K. Aberer, "Defending imitating attacks in Web credibility evaluation systems," in *Proc. 22nd Int. Conf. World Wide Web*, 2013, pp. 1115–1122.
[25] T. Berners-Lee, J. Hendler, and O. Lassila, "The semantic Web," *Sci. Amer.*, vol. 284, no. 5, pp. 34–43, 2001.
[26] A. Olteanu, S. Peshterliev, X. Liu, and K. Aberer, "Web credibility: Features exploration and credibility prediction," in *Proc. Eur. Conf. Inf. Retr.* Moscow, Russia: Springer, 2013, pp. 557–568.
[27] K. P. A. Gupta, "Credibility ranking of tweets during high impact events," in *Proc. 1st Workshop Privacy Secur. Online Social Media*, 2012, pp. 729–736.
[28] P. B. C. Castillo and M. Mendoza, "Information credibility on twitter," in *Proc. 20th Int. Conf. World Wide Web*, Hyderabad, India, 2011, pp. 675–684.
[29] A. Gupta, H. Lamba, P. Kumaraguru, and A. Joshi, "Faking sandy: Characterizing and identifying fake images on twitter during hurricane sandy," in *Proc. 22nd Int. Conf. World Wide Web*, 2013, pp. 729–736.
[30] N. Y. Hassan, W. H. Gomaa, G. A. Khoriba, and M. H. Haggag, "Supervised learning approach for Twitter credibility detection," in *Proc. 13th Int. Conf. Comput. Eng. Syst. (ICCES)*, 2018, pp. 196–201.
[31] R. El Ballouli, W. El-Hajj, A. Ghandour, S. Elbassuoni, H. Hajj, and K. Shaban, "CAT: Credibility analysis of arabic content on Twitter," in *Proc. 3rd Arabic Natural Lang. Process. Workshop*, 2017, pp. 62–71.
[32] J. Fairbanks, N. Fitch, N. Knauf, and E. Briscoe, "Credibility assessment in the news: Do we need to read," in *Proc. MIS2 Workshop Held Conjuction 11th Int. Conf. Web Search Data Mining*, 2018, pp. 799–800.
[33] Y. Gao, X. Li, J. Li, Y. Gao, and S. Y. Philip, "Info-trust: A multi-criteria and adaptive trustworthiness calculation mechanism for information sources," *IEEE Access*, vol. 7, pp. 13999–14012, 2019.
[34] T. Amjad, A. Daud, A. Akram, and F. Muhammed, "Impact of mutual influence while ranking authors in a co-authorship network," *Kuwait J. Sci.*, vol. 43, no. 3, pp. 101–109, 2016.
[35] P. D. Turney, "Thumbs up or thumbs down?: Semantic orientation applied to unsupervised classification of reviews," in *Proc. 40th Annu. Meeting Assoc. Comput. Linguistics*, 2002, pp. 417–424.
[36] F. Cornelli, E. Damiani, S. D. C. Di Vimercati, S. Paraboschi, and P. Samarati, "Choosing reputable servents in a P2P network," in *Proc. 11th Int. Conf. World Wide Web*, 2002, pp. 376–386.
[37] K. Aberer and Z. Despotovic, "Managing trust in a peer-2-peer information system," in *Proc. 10th Int. Conf. Inf. Knowl. Manage.*, 2001, pp. 310–317.
[38] P. Resnick and R. Zeckhauser, "Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system," in *The Economics Internet E-Commerce*. Bingley, U.K.: Emerald Group, 2002, pp. 127–157.
[39] L. Mui, M. Mohtashemi, C. Ang, P. Szolovits, and A. Halberstadt, "Ratings in distributed systems: A Bayesian approach," in *Proc. Workshop Inf. Technol. Syst. (WITS)*, 2001, pp. 1–7.
[40] C. Lodigiani and M. Melchiori, "A pagerank-based reputation model for VGI data," *Procedia Comput. Sci.*, vol. 98, pp. 566–571, Jan. 2016.
[41] A. Abdel-Hafez, Y. Xu, and A. Jøsang, "A normal-distribution based reputation model," in *Proc. Int. Conf. Trust, Privacy Secur. Digit. Bus.* Munich, Germany: Springer, 2014, pp. 144–155.
[42] S. D. Kamvar, M. T. Schlosser, and H. Garcia-Molina, "The eigentrust algorithm for reputation management in P2P networks," in *Proc. 12th Int. Conf. World Wide Web*, 2003, pp. 640–651.
[43] J. Sabater and C. Sierra, "REGRET: Reputation in gregarious societies," in *Agents*, vol. 1. New York, NY, USA: ACM, 2001, pp. 194–195.
[44] M. Ravi, "Trust and uncertainty in distributed environments: Application to the management of dataand data sources quality in M2M (machine to machine) systems," Ph.D. dissertation, Univ. Grenoble Alpes, Gières, France, 2016.
[45] K. Hoffman, D. Zage, and C. Nita-Rotaru, "A survey of attack and defense techniques for reputation systems," *ACM Comput. Surv.*, vol. 42, no. 1, p. 1, 2009.
[46] M. Al-Qurishi, M. Alrubaian, S. M. M. Rahman, A. Alamri, and M. M. Hassan, "A prediction system of Sybil attack in social network using deep-regression model," *Future Gener. Comput. Syst.*, vol. 87, pp. 743–753, Oct. 2018.
[47] A. B. Potey and A. B. Raut, "Combating Sybil attacks using sybilguard," *Int. J. Adv. Res. Comput. Eng. Technol.*, vol. 2, no. 2, pp. 452–455, 2013.
[48] D. N. Tran, B. Min, J. Li, and L. Subramanian, "Sybil-resilient online content voting," in *Proc. NSDI*, 2009, vol. 9, no. 1, pp. 15–28.
[49] H. Yu, P. B. Gibbons, M. Kaminsky, and F. Xiao, "SybilLimit: A near-optimal social network defense against Sybil attacks," in *Proc. IEEE Symp. Secur. Privacy (SP)*, May 2008, pp. 3–17.
[50] K. A. Rashed, C. Balasoiu, and R. Klamma, "Robust expert ranking in online communities-fighting Sybil attacks," in *Proc. 8th Int. Conf. Collaborative Comput., Netw. Appl. Worksharing (CollaborateCom)*, 2012, pp. 426–434.
[51] N. Miller, P. Resnick, and R. Zeckhauser, "Eliciting informative feedback: The peer-prediction method," *Manage. Sci.*, vol. 51, no. 9, pp. 1359–1373, 2005.
[52] A. Dasgupta and A. Ghosh, "Crowdsourced judgement elicitation with endogenous proficiency," in *Proc. 22nd Int. Conf. World Wide Web*, 2013, pp. 319–330.
[53] M. Kakol, R. Nielek, and A. Wierzbicki, "Understanding and predicting Web content credibility using the content credibility corpus," *Inf. Process. Manage.*, vol. 53, no. 5, pp. 1043–1061, 2017.
[54] S. Nepal, C. Paris, S. K. Bista, and W. Sherchan, "A trust model-based analysis of social networks," *Int. J. Trust Manage. Comput. Commun.*, vol. 1, no. 1, pp. 3–22, 2013.
[55] F. C. Mish, *Merriam-Webster's Collegiate Dictionary*. Springfield, MA, USA: Merriam-Webster, 2004.
[56] A. Jøsang and W. Quattrociocchi, "Advanced features in Bayesian reputation systems," in *Proc. Int. Conf. Trust, Privacy Secur. Digit. Bus.* Linz, Austria: Springer, 2009, pp. 105–114.
[57] G. Casella and R. L. Berger, *Statistical Inference*, vol. 70. Belmont, CA, USA: Duxbury Press, 1990.
[58] L.-H. Vu, J. Zhang, and K. Aberer, "Using identity premium for honesty enforcement and whitewashing prevention," *Comput. Intell.*, vol. 30, no. 4, pp. 771–797, 2014.
[59] M. Faisal, A. Daud, and A. Akram, "Expert ranking using reputation and answer quality of co-existing users," *Int. Arab J. Inf. Technol.*, vol. 14, no. 1, pp. 118–126, 2017.
[60] F. Fouss, Y. Achbany, and M. Saerens, "A probabilistic reputation model based on transaction ratings," *Inf. Sci.*, vol. 180, no. 11, pp. 2095–2123, 2010.

[61] Y. Liu, U. S. Chitawa, G. Guo, X. Wang, Z. Tan, and S. Wang, "A reputation model for aggregating ratings based on beta distribution function," in *Proc. 2nd Int. Conf. Crowd Sci. Eng.*, 2017, pp. 77–81.
[62] R. L. Cross and A. Parker, *The Hidden Power of Social Networks: Understanding How Work Really Gets Done in Organizations*. Brighton, MA, USA: Harvard Business Review Press, 2004.

SABA MAHMOOD is currently pursuing the Ph.D. degree with the Department of Computer Science and Software Engineering, International Islamic University Islamabad. She has served as a Lecturer in computer science with Air University Islamabad and as the Head of the Computer Science Department, Garrison Degree College (GDC), Rawalpindi, Pakistan. She has over 11 years of teaching and research experience at various universities. Her broad research interests include trusted computing, reputation systems, web information systems, and social informatics.

ANWAR GHANI received the B.S. degree from the University of Malakand, Khyber Pakhtunkhwa, Pakistan, in 2007, and the M.S. and Ph.D. degrees from the Department of Computer Science and Software Engineering, International Islamic University Islamabad, in 2011 and 2016, respectively, all in computer science. He was a Software Engineer with Bioman Technologies, from 2007 to 20011. He was selected as an Exchange Student under the EURECA Program, in 2009, for VU University Amsterdam, The Netherlands, and the EXPERT Program, in 2011, for Masaryk University, Czech Republic, funded by the EUROPEAN Commission. He is currently a Faculty Member with the Department of Computer Science and Software Engineering, International Islamic University Islamabad. His broad research interests include wireless sensor networks, next-generation networks, information security, and energy-efficient collaborative communication.

ALI DAUD received the Ph.D. degree from Tsinghua University, in July 2010. He is currently an Associate Professor with the Department of Computer Science and Artificial Intelligence, College of Computer Science and Engineering, University of Jeddah, Jeddah, Saudi Arabia. He has published about 77 articles in reputed international impact factor journals and conferences. He has completed supervising five Ph.D., 24 M.S., and 26 B.S. dissertations/theses. He has taken part in many research projects and was a PI of two projects as well. His research interests include data mining, data science, social network analysis and mining, probabilistic models, scientometrics, and natural language processing.

SHAHABODDIN SHAMSHIRBAND received the M.Sc. degree in artificial intelligence in Iran, and the Ph.D. degree in computer science from the University of Malaya, Malaysia. He was an Adjunct Assistant Professor with the Department of Computer Science, Iran University Science and Technology (IUST). He was also a Senior Lecturer with the Faculty of Computer Science, University of Malaya, and the Islamic Azad University, Iran. He is currently an Adjunct Professor with the Department of Management Science and Technology Development, Ton Duc Thang University, Ho Chi Minh City, Vietnam. He has been listed among the top 1% Researchers by Thomson Reuters (Web of Science) and one of the global peer-review awards winners (top 1% reviewers worldwide), in 2019, based on Publons (Clarivate Analytics). He supervised and co-supervised undergraduate and postgraduate students (master's and Ph.D.) by research and training. He has authored and coauthored articles published in high-impact journals and attended to high ranked A and B conferences. He is a Professional Member of ACM.