IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 24, 2019, accepted September 17, 2019, date of publication September 26, 2019, date of current version October 11, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2943877

Implicit User Trust Modeling Based on User Attributes and Behavior in Online Social Networks

JEBRAN KHAN AND SUNGCHANG LEE, (Member, IEEE)
Department of Information and Communication, Korea Aerospace University, Goyang 10540, South Korea

Corresponding author: Sungchang Lee (sclee@kau.ac.kr)

This work was supported by the Basic Science Research Programs through the National Research Foundation of Korea (NRF) funded by the Ministry of Education under Grant NRF-2017R1A2B1010817.

ABSTRACT In this paper, we present a new user trustworthiness estimation model for social networks (SN), whereas most of existing researches have been focused on the user-user/item relationship trustworthiness estimation. Users share information of their interest on various social media without their trustworthiness verification. Therefore, SN are susceptible to malicious users for misinformation spreading. In SN, the original information source is generally unknown and the user who is sharing the contents is the only known information about the source. Therefore, the user's trustworthiness is an effective criterion for SN content's trustworthiness estimation. However, the users are unable to identify trustworthy/untrustworthy users, and the existing user-user/item relationship models do not provide user trustworthiness information. Our proposed model provides a systematic way to assess the user trustworthiness based on user attributes and interaction behavior. The proposed model is helpful to avoid the trust sparsity (implicit trust model), trust subjectivity (user's objective/collective trustworthiness estimation model) and cold-start user's trustworthiness (user's attributes-based trust modeling) problems. We employ friends-recommendation (FR) as an exemplary application to evaluate the performance of our proposed model in trust-aware recommendations. Simulation results illustrate that our trust-aware FR model outperformed the existing trust and FR models.

INDEX TERMS Credibility, reliability, social-ties strength, trust-aware recommendation, user trustworthiness modeling.

I. INTRODUCTION

Online social networks (OSN) enable users to form relationships, based on homophily, propinquity, popularity, and mutual interests. People exploit OSN to express their views and ideas and share their experiences on a diverse range of topics and items. Over the past decade, OSN emerged exponentially, as per June 30, 2018 according to The World-Stream [1], about 79% of US population and almost 22% of the total world population use Facebook, and about 32% of US population uses Instagram, i.e., the second excessively used OSN platform in the US. Users' connectivity in OSN and decentralized structure of OSN proliferated users' interactions, contents sharing and collaborations. However, it aggravates the risk of uncertainty, lack of reliability and vulnerability to user's malicious manipulation. Trust is getting the attention of researchers because it enhances the quality of cooperation and interactions between social networks (SN) users' and minimize uncertainties and risks from unreliable users and therefore mitigating the trustworthiness and information overload problem.

Trustworthiness is a user's quality of being authentic, reliable, or truthful. Trust is the degree of confidence which the trustor; user which evaluate the trustworthiness of the target user, has on the trustee; target user for trustworthiness evaluation. In OSN, trust reflects social ties strength, credibility, reliability, and self-orientation of a user in SN. Trust in the SN depends on users' interactions, profile information, and relationships.

In many online portals, the users provide ratings and reviews to other users and items which yield an explicit trust of the users and items. However, the proportion to supply feedback and ratings on these portals is significantly low compared to the number of total users', and most users

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

eschew rating other users and items [2]. Moreover, OSN like Facebook, Twitter, and Instagram do not facilitate users to provide explicit trust to other users and contents, which necessitate a mechanism for implicit trust computation to rank users and social contents based on trust on these platforms. Some researchers [3], [4] utilized users interactions for implicit trust computation by combining with users interest [5] and proximity [6], [7]. Such type of trust is classified as interpersonal trust, i.e., trust between two connected users. In social network analysis (SNA), the interpersonal trust can be termed as social ties strength (STS). The STS can be extended to friends of a friend (FOAF) which is also known as propagation trust. The STS and propagation trust are good contributors to the user's trustworthiness, but they are not the exact and only determinants of trust.

Limitations of the previous trust models:

- Current models mostly deal with relational trust (user-user/items). However, for trustworthy contents recommendation, a user's trustworthiness is crucial and unavoidable in OSN like Facebook, Twitter and LinkedIn.
- Most of current trust models lack a systematic mechanism to map SN raw parameters to trust parameters such as credibility and reliability.
- Social interactions can be of many types, i.e., comments, tags, likes, posts and shares etc. These interactions should have a certain weight based on the user individual behavior and behavior of all users in a group.
- User-user/items relationship's trust and user explicit trust are subjective.
- In current explicit trust based recommender systems there is a problem of sparsity and cold start users.

In this paper, to deal with the limitations of the previous state of the art trust models, we proposed a novel model for implicit trust computation of SN user's taking into consideration the SN user's behavior, relationships, and SN parameters importance.

The main contributions of this paper can be summarized as: We proposed a novel trust model for user's trustworthiness computation in OSN based on user's attributes, behavior, and relationships, by mapping these SN raw parameters to SN trust parameters in a systematic manner. We applied the proposed trust model to trust-aware FR.

- Proposed a new user's trustworthiness estimation model in OSN, by taking into consideration the OSN parameters saliency and relevancy to trust.
- We mapped the SN raw parameters i.e. user-user interactions and user profile information (demographics and location information) systematically to the SN trust parameters, i.e. credibility, reliability and strength of social ties.
- The proposed trust model is based on homophily, intimacy (STS) and behavior.
- The proposed model integrates the subjective and objective aspects of trust to assess user's trustworthiness.
- Proposed an trust-aware friends-recommendation (FR) algorithm based on mutual friendship weighted by user's trustworthiness.
- The proposed trust-aware FR system avoid the problems of sparsity and cold start users in recommendation.

The remaining paper is organized as follow: Section II presents an overview of the background study and related work. Section III discuss about proposed model, assumptions, and applications in detail. In section IV we described the datasets, and discussed the simulation results, and analysis of the proposed model. Section V is the conclusion of this work.

II. BACKGROUND AND RELATED WORK

The scope of this paper is implicit trust computation based on human psychology of trust building using SN parameters. Trust is the willing reliance of trustor on the trustee. In psychology, trust is the generalized reliance of trustor on the statements of trustee [8], [9]. Trust is one of the fundamental, and most important elements in human relationships building, and relationship strength assessment. It embeds all the actions, that one performs, and interactions, that one has with others. Trust can be described in two notions, i.e., subjective and objective. Subjective trust is derived from direct interactions between two or more users. Objective trust, also referred to as reputation, is the collective measure of trustworthiness extracted from ratings and reviews.

A. TRUST AND SOCIAL NETWORKS

In SN trust can be explicit (subjective), and implicit (objective). An explicit trust is an approach where the SN users explicitly provide trust scores to other users/items. These scores are provided in form of ratings and reviews. Many publicly available datasets exist with explicit trust scores, e.g., Epinion and Film Trust. These datasets contain binary trust scores (0/1). Therefore, all trust relations have the same value; however, users can have different levels of trust. In real life SN, it is very challenging to urge users for scoring each other. Moreover, human judgments can be influenced by many factors such as interests, demography, spatial proximity, self-orientation, and intimacy with other users. On the other hand, the implicit trust score can be predicted by exploiting the user's activities, attributes, and relationships. In real-life general-purpose SN, the implicit trust models fit well. The explicit trust models suit the online business portals.

B. TRUST CONTRIBUTING SOCIAL NETWORKS PARAMETERS

In the last decade, researchers proposed various models for trust computation based on different approaches and factors. These factors can be users' attitudes, experiences, behavior, and users' attributes similarities, such as demographic similarity, spatial proximity, and interests similarity. Attitude of a user is the degree of his/her interest/likeness for something/someone. Experience is the perception about users extracted from mutual interactions. Behaviors is the patterns, and frequency of interactions. Experience affects

users' attitudes and behavior in SN. Positive experiences may change user's opinion about something/someone positively. Similarly, positive experiences may compel users to interact frequently. Most of the current trust models, such as PowerTrust [10], and PeerTrust [11], utilized user's feedbacks/experiences. Paradesi *et al.* [12] proposed a model for composing trustworthy Web services by integrating behavioral trust into Web services compositions. This model leverage user's feedbacks to progressively update the belief of users of a specific Web service. However, this model is based on user's feedbacks/experiences instead of their behavior. Many researches in Psychology, social sciences, and behavior sciences, focused to observe, and analyze the user's behavioral aspects, such as, [13] and [14]. Helbing [15] presents a mathematical model for person's behavior estimation in a social field. Webb *et al.* [16] proposed a model for tracking the recognizable patterns in spammer's behavior. Caverlee and Webb [17] studied characteristics of large scale OSN by tracking user's activities over time and analysing user's profiles. Yan and Yan [18] and Yan *et al.* [19] proposed a trust model by utilizing mobile applications usage behavior; This model assumed that user's experience leads to their attitude, which in result affect their behavior. Nepal *et al.* [20] with the same assumption, studied the effects of unexpected events in consumer's behavior on trust assessment. Adali *et al.* [21] proposed measures of trust which can be derived from interaction behavior of SN users, assuming that trust between two users increase interactions between them which further enhance their mutual trust. Yan *et al.* [22] presented a recommender system based on trust and behavior.

C. GRAPH BASED TRUST MODELS

SN graph topology can affect the trust of SN users and many SN researcher proposed trust models based on the SN graph structure. Buskens [23] observed in his study on SN that high out-degree users and their neighbors are more trusted. Models which exploit the SN structure have usually utilized the notions of Web of trust (WoT) and FOAF. Golbeck *et al.* [24] extended the FOAF schema and proposed a method to create a Trust network over the semantic web which allowed the users to specify a trust level for other users they know. Ziegler and Lausen [25] proposed a novel method to calculate local group trust for semantic web trust management. They motivated using partial trust graph exploration to reduce computational complexity. Golbeck [26] proposed TidalTrust to derive trust relationships between users in SN by exploiting the FOAF vocabulary. Zhang *et al.* [27] expanded this trust model by incorporating user's ratings and reliability for trust computation, using the edge-weighted graph. Maheswaran *et al.* [28] proposed a trust estimation model using gravity/attraction. Kim *et al.* [29] modeled WoT without using user ratings. The main factors of this trust model are context-aware user's reputation and affinity. Zuo *et al.* [30] used a set of trust chains and graph for trust computation in SN.

D. INTERACTIONS BASED TRUST MODELS

People interact and share information with people they trust. Therefore, most of the current implicit trust estimation models are based on user's interactions [31], [32] and communication behavior. Liu *et al.* [33] presented trust estimation model based on user's interactions and behavior in SN communities. They classified and represented the user action and interactions, between pairs in communities, into two classes: i) user actions metrics which contain number of interactions, frequency of interactions, and length of interactions (comments, posts), and ii) user relationships metrics such as connection between contents creator and responder, content creator and contents creator, and responder and responder. They also considered the time difference for the interactions. They provide a supervised learning technique to predict trust between SN users in communities based on user-user interactions. They train classifiers by using these interactions to predict the user-user trust. The user-user trust is referred to as STS. Nepal *et al.* [4] proposed a model to compute user's trustworthiness based on interactions and similarities. They named the model as STrust. This model incorporates popularity trust and engagement trust. Adali *et al.* [21] proposed a model for behavioral trust based on two types of trust: conversational trust and propagation trust.

E. SIMILARITY BASED TRUST MODELS

Literature in sociology and psychology comprehensively addressed the factors which affect trust; positively or negatively, but the evidence of direct relationship between trust and similarities are vague. However, strong linkage between similarity and friendship, or user-user attraction is reported in [34] and [35]. These studies confirmed positive effects of attitude similarity on relationship strength. In SN, users are more likely to establish relationship with similar users [36]. This property is known as homophily in social sciences [37]. The similarity can be spatial, demographic or interest similarity.

Several studies are available to show a strong relationship between trust and similarity. Ziegler et. al. [5] in their study analyzed correlation between trust and interest similarity. In [38] and [39], conclude that users trust on recommender systems (RSs) is directly related to the type of recommendation it makes. RSs which recommend items based on user's preferences are more trustworthy. In online/offline social circles users/people prefer recommendations from their close friends which are more trusted. Another study [40], established a similarity-trust connection. They analyzed two systems and observed that the higher user-user similarity portrays higher trust between them. Based on these correlations between trust and similarity, some researchers [4], [29], [41] exploit user similarity in integration with interactions and behavior to estimate user-user trust in SN. O'Donovan and Smyth [42], proposed a trust model for contents recommendation based on user similarity. A social

contents recommendation method was proposed based on user interactions and similarities in [43].

F. TRUST AWARE FRIENDS RECOMMENDATION

The SN service providers aim to grow their network fast. To achieve this, they developed an FR system which suggests users with similar interest or like-minded users to other users as their potential friends. The problem is generally referred to as missing links or potential links problem. FR is mining of potential links or discovering missing links between users [44]. Many factors can influence FR simultaneously such as user's demography, location, network topology, social relationships, and user's interactions. However, most of the researches utilize single or some of the factors such as relationship [45], profile information [46], and interactions [47], [48] and trust [24], [49]–[51].

Currently the trust-aware FR systems utilize the user's friendship information [51] and user's interactions [47], [48] to compute trust between SN users. In [52] and [53] the trust information is utilized to reduce sparsity in user-item matrix and improve neighborhood set by recommendation. Some researchers directly employ the trust information for FR. In many SN a mechanism for explicit trust assignment is not available and only implicit trust can be calculated by using SN parameters. McAllister and Daniel [54] rated the users' difference for implicit trust computation and described trust propagation. They combined the implicit trust and trust propagation with user information for FR to a target user. Many other researches such as [55]–[57] used trust for FR in SN. Cheng et. al. [58] proposed a scalable FR model based on D-S evidence theory based on user influence, direct and indirect trust. They considered the mutual friendship and mutual interactions for trust scoring.

The existing trust models are mainly based on the user's friendship, interactions, popularity, and trust propagations. However, many other factors can influence trust along with user interactions and relationships. In this work we present a multi-facet generic trust computation model by mapping the SN parameters such as interactions, profile information, user's interests and relationships to SN trust parameters that are STS, credibility, and reliability.

III. IMPLICIT USER TRUST MODELING IN OSN

A. PRELIMINARIES

We represented the SN graph by $G(V, E)$, where V is the set of users, and E is the set of weighted edges. The set of users V contain user's identity (ID) and user's attributes. The user's ID are numbers, ranging $1 - N$, where N is the number of users in the graph G. We classified the user's attributes in three categories, i.e., spatial information, interests information and demographic information, which are represented as Sp, In and Dm, respectively. Sp is the set of user's geographic locations, where Sp_i represents the location of i^{th} user in the network, and each Sp_i consists of latitude (ϕ) and longitude (φ) values. In is the set of users' interests, and In_i represent the interests of i^{th} user. Dm is set of users' demographic information, where Dm_i is the demographic information of each i^{th} user. Dm_i contains age, gender, occupation, religion, language, and political views. The set of edges E contain social interactions represented as I_p, and relationships represented as $E(u, v)$. The social interactions are categorized as social actions $I_p(u \rightarrow v)$ and social responses, where social actions are directed from target users u to its friends v and social responses $I_p(u \leftarrow v)$ are directed from v to u.

We derived a set of intermediate trust parameters from SN user's attributes and interactions. The intermediate trust parameters include, user similarities with its neighbors such as spatial similarity ($SpSim(u, v)$), demographic similarity ($DmSim(u, v)$) and interest similarity ($InSim(u, v)$), social influence due to similarity ($SI_s imilarity(u, v)$), social intimacy of user v to u ($SI_i nteractions(u \leftarrow v)$), dependability of user v on u ($Dep(u, v)$), openness ($Op(u)$), accessibility ($Acc(u)$), popularity ($Pop(u)$), expertise ($Exp(u)$), competence ($Compt(u)^{domain}$), consistency ($Cons(u)$), and availability ($A_T(u)$). Where u is the target user and v represent its neighbors. We defined these intermediate trust parameters at their appropriate places, when used in Section III-C.

The set E is tagged by STS. The STS between users u and v is calculated by using interaction and similarities between them. The set of users is weighted by trust, which is the average STS of users with its neighbors. The SN trust parameters which we considered in this work are STS, credibility, and reliability which we represented as $ST(u \leftarrow v)$, $Cr(u)$, and $Rel(u)$, respectively. User trust is represented as $UT(u)$, u is the target user. These notations are listed with a brief description in Table 1.

B. PROPOSED SN TRUST MODEL

In this paper, we propose a new model for SN user's trustworthiness assessment based on user's interaction's behavior, attributes, and relationships. Figure 1 provides a block diagram of our proposed model. The underlying theme of our proposed model is to calculate the SN user's trust by using SN raw parameters. The whole model is divided into four phases; namely, i) data collection/generation ii) weights estimation iii) trust computation and iv) trust application. The first step is the SN raw data collection/generation. The SN raw data consists of user interactions, profile information, interests information, and user relationships. In the second phase, the importance weight for each trust contributing factor is provided/estimated. We divided the weights into three categories, i.e., application specific weights, interactions weights, and similarities weights, where the interactions weights are estimated from users' behavior while the other two weights are application specific, which can be provided by users of this model. The application-specific weights and similarities weights can also be calculated automatically by using machine learning algorithms provided the application specific data is available. We assigned equal values to the application-specific weights and similarities weights. The next two phases are, trust computation and trust application

TABLE 1. List of notations and description.

	Notations	Description
Terminologies	$G(V, E)$	Synthetic graph with set of users V and edges E
	Sp_i	Spatial information of user i
	In_i	Interest information of user i
	Dm_i	Demographic information of user i
	$I_p(u \rightarrow v)$	Social actions: Interaction (p) from user u to v, where p is the type of interaction
	$I_p(u \leftarrow v)$	Social responses: Interaction (p) from user v to u, where p is the type of interaction
	$SpSim(u, v)$	Spatial similarity of user u with its neighbor v
	$InSim(u, v)$	Interest similarity of user u with its neighbor v
	$DmSim(u, v)$	Demographic similarity of user u with its neighbor v
	$SI_similarity(u, v)$	Social influence of user u on v due to similarity
	$SI_interactions(u \leftarrow v)$	Social intimacy of v to u
	$Dep(u, v)$	Dependability of user v on u
	$Op(u)$	Openness of user u or availability of user information
	$Acc(u)$	Accessibility of user u or availability of user contact information
	$Pop(u)$	Popularity of previous contents generated by user u
	$Exp(u)$	Expertise of user u
	$Compt(u)^{domain}$	Domain specific competence of user u
	$Cons(u)$	Consistency of user u
	$A_T(u)$	Availability of user u in time T
	$ST(u \leftarrow v)$	Social ties strength between user u and v.
	$Cr(u)$	Credibility of user u
	$Rel(u)$	Reliability of user u
	$UT(u)$	Trust of user u
	$frRecScore(u \mid v)$	friendship similarity score of user u for v
	$frshipSim(u, v)$	Friendship similarity between u and v
Weights	w_{st}	Weight of social ties strength
	w_{cr}	Weight of credibility
	w_{rel}	Weight of reliability
	W_p	Contribution weight of interaction parameter p, where p is the type of interaction such as likes, comments etc.
	W_{up}	User specific weight which is calculated based on tendency of user's activities
	W_{gp}	Parameter general weight which is calculated on the basis of activities in the dataset
	α	Weight of spatial similarity
	β	Weight of demographic similarity
	γ	Weight of interest similarity

(FR in our experiments), are explained in Section III-C and III-D, respectively.

C. TRUST COMPUTATION

We calculated the SN user trustworthiness $UT(u)$ using Equation 1, which shows the trust as commulative sum of weighted STS $ST(u \leftarrow v)$, credibility $Cr(u)$ and reliability $Rel(u)$.

$$UT(u) = w_{st} \cdot \frac{1}{NN} \sum_{v \in Neighbors} ST(u \leftarrow v) + w_{cr}.Cr(u) + w_{rel}.Rel(u) \quad (1)$$

where w_{st}, w_{cr} and w_{rel} are the weights of STS, credibility, and reliability in trust computation, respectively. We considered equal weights for all parameters. These weights may vary according to their importance in the target application. The sum of weights is 1. *NN* is the number of neighbors of user u.

The STS is the degree of intimacy between two connected users in SN. The social intimacy is the degree and frequency of interactions of SN user to the contents shared by another SN user. In SN, users are more responsive to users of similar demography, proximity, and interests. The higher the user similarity, the stronger is the STS. Therefore, we calculated the STS $ST(u \leftarrow v)$ from users v and u as social intimacy ($SI_{interactions}$) weighted by social influence due to similarity ($SI_{similarity}$) as shown in Equation 2.

$$ST(u \leftarrow v) = SI_{similarity}(u, v) \times SI_{interactions}(u \leftarrow v) \quad (2)$$

The STS is directed and asymmetric due to the asymmetry of social intimacy. The social intimacy ($SI_{interactions}$)$(u \leftarrow v)$, of user v to u is the weighted sum of interactions from user v to u, as shown in Equation 3. It depends on the type and frequency of interactions exchanged between SN users u and v.

$$\begin{aligned} SI_{interactions}(u \leftarrow v) &= \sum_{p \rightarrow parameter} W_p \times I_p(u \leftarrow v) \\ &= W_{Likes}(v) \times I_{Likes}(u \leftarrow v) \\ &\quad + W_{Comments}(v) \times I_{Comments}((u \leftarrow v)) \\ &\quad + W_{Share}(v) \times I_{Share}(u \leftarrow v) \end{aligned} \quad (3)$$

In Equation 3, $I_p(u \leftarrow v)$ is the total number of interactions p from user v to u, where p is the type of interaction i.e., likes, comments, shares. In the dataset we represented these interactions as *R_likes*, *R_comments* and *R_shares*. W_p is the weight of interaction p, which is equal to the average sum of user-specific weight W_{up} and parameter-general weight W_{gp}, as in Equation 4.

$$W_p(u) = \frac{W_{up}(u) + W_{gp}}{2} \quad (4)$$

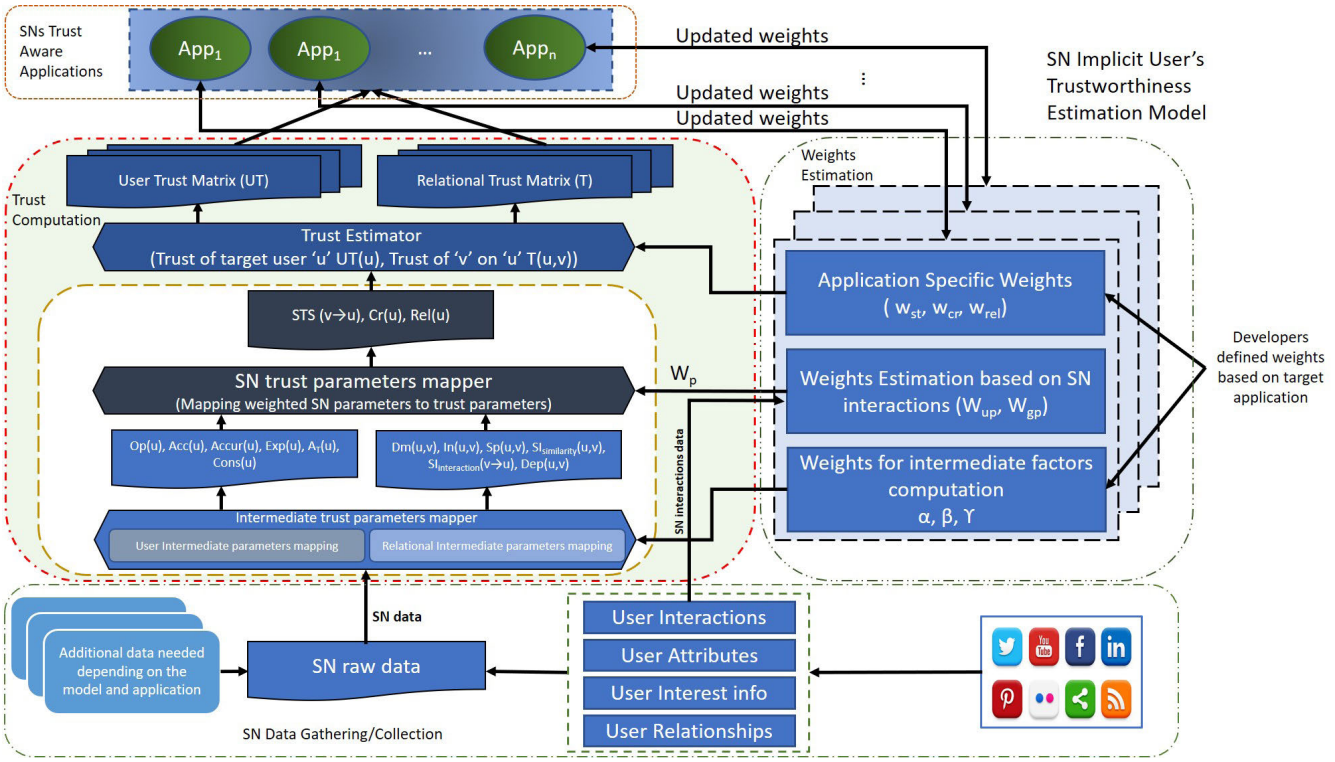

FIGURE 1. Descriptive block diagram of the proposed social network trust model.

The user-specific weight $W_{up}(u)$ for user u is the tendency of the user to generate specific types of interaction, e.g., some users like almost everything on their timeline but rarely comment, for such users the weight of likes is much less than the weight of comments and vice versa. The user-specific weight $W_{up}(u)$ is calculated by using Equation 5 and it is equal to 1 minus the ratio of interaction p to total number of interactions generated by user u. We have normalized these weights so that the total sum of these weights is equal to 1.

$$W_{up}(u) = 1 - \frac{I_p(u)}{I(u)} \tag{5}$$

We calculated the parameter-general weights W_{gp} by using the Random forest (RF) [59], a machine learning technique, in R language. RF can find the feature's importance weights robustly with the slight tuning of its parameters. It generates an ensemble of classification and regression trees (CARTs) using bagged samples of the training data and finds each feature's importance weight in relation to the target class. It has two parameters, the number of trees (ntree) and the number of features required to split the tree node (mtry), which need tuning according to the problem size. In this work, our problem set is limited, and we have a smaller number of features. We selected arbitrary values in repeated experiments to find optimal values of these two parameters (i.e., mtry and ntree). The parameters that we used as predictors are, *Likes*, $R-Likes$, *Comments*, $R-Comments$, *Shares*, $R-Shares$, and *Trust* was used as a target class in the RF method. The actual values of parameters and other related settings that we used with RF to compute the feature's weights are as follows, in Table 2:

TABLE 2. Random forest settings for parameter general weights estimation.

Setting Variable	Value
Number of trees ($ntree$)	50
Variables used to split the tree node ($mtry$)	2
RF Model	Regression
Training Data	User-Interactions-Trust-Data
Predictors	$Likes, R-Likes,$ $Comments, R-Comments$ $Shares, R-Shares$
Target class	$Trust$

The normalized parameter-general weights for parameter p used in trust computation are given in Table 3. The weights of actions and responses is equal because response to user u is an action from user v.

TABLE 3. Parameters general weights.

Interaction Type	General Weight (W_{gp})
Likes	0.1
Comments	0.22
Posts	0.18
R-likes	0.1
R-comments	0.22
R-Shares	0.18

The social influence due to similarity $SI_{similarity}(u, v)$ of user u and v is calculated as sum of weighted similarities, as shown in Equation (6) [43], where $SpSim(u, v)$, $DmSim(u, v)$ and $InSim(u, v)$ are the spatial, demographic and interest similarities between user u and v, respectively.

$$SI_{similarity}(u,v)=\alpha.SpSim(u,v)+\beta.DmSim(u,v)+\gamma.InSim(u,v) \quad (6)$$

In Equation 6, α, β and γ are the importance weights of spatial, demographic and interest similarities, respectively, such that $\alpha + \beta + \gamma = 1$. The spatial similarity is the geographic distance in km between user u and v. The $SpSim(u, v)$ is calculated by using cosine similarity [60]. The spatial similarity is normalized by the geo-distance between users, as shown in Equation 7.

$$SpSim(u, v) = \frac{1}{\left[1 + \frac{Dist(u,v)}{1000}\right]} \left(\frac{Sp_u.Sp_v}{\parallel Sp_u \parallel . \parallel Sp_v \parallel}\right) \quad (7)$$

In Equation (7), the $Dist(u, v)$ is the geo-distance between the users u and v, which is calculated by using Harvesine formula [61], [62], as shown in Equation (8).

$$a = \sin^2\left(\frac{\Delta\phi}{2}\right) + \cos(\phi_{Sp_u}) \cdot \cos(\phi_{Sp_v}) \cdot \sin^2\left(\frac{\Delta\varphi}{2}\right);$$
$$where\ \Delta\phi = \mid \phi_{Sp_u} - \phi_{Sp_v} \mid, \Delta\varphi = \varphi_{Sp_u} - \varphi_{Sp_v}$$
$$c = 2\arctan\left(\frac{\sqrt{a}}{\sqrt{1-a}}\right)$$
$$Dist(u, v) = R \times c \quad (8)$$

where Sp_u and Sp_v represents the GPS coordinates of the users u and v, respectively. Sp_u and Sp_v contains pair of latitude (ϕ) and longitude (φ) values. ϕ_{Sp_u} and ϕ_{Sp_v} are the latitudes of the users u and v and φ_{Sp_u} and φ_{Sp_v} are the longitudes of users u and v, respectively.

The demographic similarity ($DmSim(u, v)$) of u and v is computed by using cosine similarity [60], as shown in Equation (9).

$$DmSim(u, v) = \frac{\sum_{i=1}^{D} Dm_{i_u}.Dm_{i_v}}{\sqrt{\sum_{i=1}^{D} Dm_{i_u}^2} \cdot \sqrt{\sum_{i=1}^{D} Dm_{i_v}^2}} \quad (9)$$

where Dm_{i_u} and Dm_{i_v} are the i^{th} demographic attribute of users u and v and i ranges from $1 - D$, D is the number of attributes in Dm.

In SN, users express their interests in different items/events such as some users have interest in sports, some users like books or movies, some follow celebrities or television shows, and some may like to follow news and politics. We used the Jaccard similarity formula [63] to compute the interests similarity ($InSim(u, v)$) between users u and v, as shown in Equation 10. The interests similarity between users u and v is the ratio of common interests to the total number of interests they have. The user's interest similarity is increased with an increase in the number of common interests.

$$InSim(u, v) = \frac{[In_u \bigcap In_v]}{[In_u \bigcup In_v]} \quad (10)$$

where In_u and In_v are the set of interests of users u and v, respectively.

Credibility is the quality of being believable. For any piece of contents, the credibility is observed in three dimensions; message credibility, source credibility, or media credibility. The message credibility is the perceived trustworthiness of the message itself. The source credibility is the trustworthiness of the information source. Also, medium credibility is the type of medium used for information sharing. In the case of conventional information sources such as newspaper and television, the source and media are known, and the owner of the media take responsibility for the correctness of the information. However, in social media, the original information source is generally unknown, and only the user information is available. This information can be incomplete or even fake. Therefore, it is inevitable to measure the user's credibility contents trustworthiness assessment. The factors that influence the user's credibility are openness, expertise, accessibility, popularity, and competence. Openness is the availability of user credentials, such as their demographic information. Expertise is calculated by using user education, profession, age, and experience. Accessibility is calculated by the availability of user location and contact information. Popularity is the ratio of responses to the user's previous contents.

$$Cred(u) = Op(u) + Exp(u) + Acc(u) + Pop(u) + Compt(u)^{domain} \quad (11)$$

We calculated the credibility $Cred(u)$ of user u by using Equation 11, where $Op(u)$, $Exp(u)$, $Acc(u)$, $Pop(u)$ and $Compt(u)^{domain}$ represent openness, expertise, accessibility, popularity, and competence, respectively. The user's competence is the ability of a user to have a thorough knowledge/understanding of a specific/target domain. The user's competence is context/domain dependent. In this work, we do not consider the SN user competence, as our model is context independent.

Reliability is the expected response from a user on contents of its neighbors. The SN user is considered to be reliable if he/she responds as expected/desired within a short time. In OSN, the user's reliability can be computed from user dependability and consistency. In this work, we computed the user reliability as the user average dependability weighted by consistency, as shown in Equation (12);

$$Rel(u) = Cons(u) \times \frac{1}{N} \sum_{v \to Neighbors} Dep(u, v) \quad (12)$$

where $Dep(u, v)$ is the dependability of user v on u and $Cons(u)$ is consistency (the average response rate) of user u. NN is the total number of neighbors of user u. In SN like Facebook the user v dependability on u can be calculated from responsiveness of u to the contents of v, as shown in Equation (13);

$$Dep(u, v) = \sum_{p \to parameter} W_p \times \frac{I_p(u \to v)}{I_p(u, v)} \quad (13)$$

where $I_p(u \rightarrow v)$ is the total number of interactions p from user u to v and $I_p(u, v)$ is the total number of interactions between user u and v. W_p is the weight of interaction p and p is the type of interaction i.e., likes, comments, shares. The weights W_p is calculated by using Equation 4 and Table 3.

The consistency is measured from the response to availability ratio, as shown in Equation (15), where $I_T(u)$ is the total number of interactions from user u in time T and $A_u(T)$ is the average availability of user u in time $[0 - T]$ as shown in Equation (15).

$$Cons(u) = \frac{I_T(u)}{A_u(T)} \tag{14}$$

$$A_u(T) = \frac{1}{T} * \int_0^T A_u(t)dt \tag{15}$$

The proposed algorithm for SN user's trust computation is shown in Algorithm 1. The algorithm takes as the SN graph $G(V, E)$ weighted by users' interactions I_p, set of users V and users' attributes, as input. The users' attributes consist of users' demographic, spatial, and interests information. The proposed algoritm in lines $4 - 7$ compute similarity based users' social influence $SI_{similarity}(u, v)$, which is a weighted sum of users' spatial, interests and demographic similarities by using Equations 6-10. The proposed algorithm in lines $8 - 11$ calculate the social intimacy $SI_{interaction}(u \leftarrow v)$ based on users' interactions by using Equations 3-5. The STS $ST(u \leftarrow v)$ is calculated by using Equations 2 in line 12. The STS, as discussed earlier in this Section III-B is $SI_{interaction}(u \leftarrow v)$ weighted by $SI_{similarity}(u, v)$.

User's credibility $Cr(u)$, and reliability $Rel(u)$ are calculated in line $13 - 14$ by using Equations 11 and 12, respectively. In line 15 the proposed algorithm compute SN user's trustworthiness $UT(u)$ by using Equations 2. A users' trustworthiness matrix UT is returned as output of the proposed algorithm.

Algorithm 1 Algorithm for SN User's Trustworthiness Computation

Input: Graph $G(V, E)$, Social interactions I_p, Set of Users V in the network with their attributes, Set of weights provided by the program developer or estimated from data available

Output: Trust matrix T

1: **for** each $u \in V$ **do**
2: **for** each $v \in V$ **do**
3: Calculate spatial similarity $SpSim(u, v)$, using Equation 7
4: Calculate interest similarity $InSim(u, v)$, using Equation 9
5: Calculate demographic similarity $DmSim(u, v)$, using Equation 10
6: Calculate social intimacy/influence $SI_{similarity}(u, v)$ based on similarities, using Equation 6
7: Calculate interactions weights W_{gp}, using Random Forest technique
8: Calculate interactions weights W_{up}, based on user behavior using 5
9: Calculate interactions weights W_p by adding W_{gp} and W_{up}, as in 4
10: Calculate social intimacy $SI_{interaction}(u \leftarrow v)$ based on users' interactions, using Equation 3
11: Calculate STS $ST(u \leftarrow v)$, using Equation 2
12: Calculate credibility $Cr(u)$ of user u, using Equation 11
13: Calculate reliability $Rel(u)$ of user u, using Equation 12
14: Calculate trust $UT(u)$ of user u, using Equation 1
15: **end for**
16: **end for**
17: **return** User Trustworthiness matrix UT

D. TRUST BASED USERS' FRIENDS RECOMMENDATION

In heterogeneous SN, there exist many factors which influence a user's FR. These factors range from similarity to popularity, from co-authorship to common interests, from graduating school to working organization, and from FOAF to influential users. Therefore, it is critical to take multiple information for most relevant FR. Any of these factors cannot be used individually to generate efficient FR. In this paper, we consider a combination of these SN parameters to propose an FR based on the user's trustworthiness and the fraction of mutual friends. We calculated SN users FR score by using trust and fraction of common friends, as shown in Equation 16;

$$frRecScore(u \mid v) = UT(u) \times frshipSim(u, v) \tag{16}$$

where $frRecScore(u \mid v)$ is the friendship score of u for user v and $UT(u)$ is trust of u. $frshipSim(u, v)$ is the fraction of common friends. It is calculated by using jaccard similarity [63], as shown in Equation 17, i.e., the number of common friends divided by the total number of friends. $UT(u)$ is calculated by using Equation 1.

$$frshipSim(u, v) = \frac{|\, fr(u) \bigcap fr(v) \,|}{|\, fr(u) \bigcup fr(v) \,|} \tag{17}$$

Algorithm 2, shows the pseudocode for our proposed FR algorithm. The algorithm takes graph $G(V, E)$, user trustworthiness matrix UT, i.e, output of Algorithm 1, and number of friends to be recommended K. The algorithm returns a list of $top - k$ friends as output. The algorithm calculates the friendship similarity score $frshipSim(u, v)$, between the target users u and v, in line 4 using Equation 17. In line 5 friendship recommendation score $frRecScore(u \mid v)$ of user u for v is calculated by from weighted sum of friendship similarity and trust of u, by using Equation 16. Sort the $frRecScore(u \mid v)$ in descending order in line 8. The algorithm returns a list $top - k$ FR $TrustFriRec(v)$ for each user $v \in V$ with FR score $friRecScore(u)$ of each u.

Algorithm 2 Algorithm for SN Friends Recommendation Based on Trust and Mutual Friends

Input: Graph $G(V, E)$, User trustworthiness matrix UT, and number of friends to be recommended K

Output: Sorted list of top K trustworthy friends $TrustFriRec(v)$ with friendship scores $friRecScore(u)$

for each $v \in V$ **do**
  for each $u \in V$ **do**
    Calculate friendship similarity $frshipSim(u, v)$ between user u and v, using Equation 17
    Calculate friendship recommendation score $friRecScore(u \mid v)$ of each user u for v based on friendship similarity and trust of u, using Equation 16
  end for
end for
  descending_order_sort($friRecScore(u \mid v)$)
return $TrustFriRec(v) = \text{top}(K, friRecScore(u \mid v))$

IV. EVALUATION AND SIMULATION RESULTS

In this section, the feasibility and performance comparison of our proposed model with other, state of the art models based on simulation results is discussed. In this section first we discuss the dataset used in our experiments, then we will discuss the feasibility of our proposed model for SN user trust computation, and at the end of this section, we will discuss the performance of our model in FR.

A. DATASET DESCRIPTION

OSN like Facebook, Instagram, and Twitter provide limited information to preserve user privacy. Several datasets are available to train and evaluate SN models for applications in different domains. However, these datasets do not contain all the desired/required information. These datasets insufficient in data, to evaluate multi-facet models. Our model is one of such type, and the data we need for our model evaluation was not available in a single dataset. Therefore, in this work for our model evaluation, we used synthetic data generated by using our previously proposed model in [64], as mentioned in Table 4. The dataset descriptions are available in our previously published paper [64]. For verification of our results, we used twitter data, crawled using Gephi, and two publicly available datasets, i.e., Movielens dataset [65], and Filmtrust dataset [66]. The descriptions of these datasets are in Table 4.

We combined the publicly available datasets by randomly selecting 900 users with their available information from these datasets and interactions crawled from twitter. We assigned these attributes and interactions randomly to the selected users. We formed new relationships between the selected users based on homophily and preferential attachments. The description of the datasets is shown in Table 5.

In the synthetic data generated, there was no trust information available. Therefore, we have tagged the user relationships, in the synthetic datasets, as trusted/untrusted based on user similarities and interactions. We calculated the average of user's relational trust to determine user's trustworthiness.

TABLE 4. Datasets introduction.

Dataset	Properties
Twitter data	Social interactions, followers and followings of 900 users
Movielens	User demographic information along with ratings of 900 [65]
Filmtrust	Interest information and trust of 900 users [66]
Synthetic data	Synthetic dataset generated by using [64]. We generated 1000 nodes and 46,780 edges with other required attributes, as shown in 5.

TABLE 5. Description of the datasets used.

Item	Statistics	
	Synthetic dataset	Combined dataset
User Attributes	Location (synthetic latitude and longitude information generated in North American region), interests, political-orientation, and demographics such as age, gender, religion, language, maritial status, and profession	Interest ratings, demographics such as, age, gender, and location information generated synthetically
Users num.	1000	900
Edges num	46,780	40,875
Min. Edges per user	10	10
Avg. Interactions	15807.44	500
Max. Interactions per user	60398	4832
Min. Interactions per user	418	12
Avg. Degree	46.78	45.41
Avg. Path Length	2.49	1.96
Avg. Clustering Coefficient	0.49	0.46
Graph Density	0.094	0.091
Graph Diameter	6	7
Modularity	0.53	0.47

TABLE 6. Comparison of Top-k Users based on Similarities, Interactions, Popularity and proposed model with Top-k trusted users in the Dataset.

Top-k	Top in dataset	Spatial Similarity based	Interest Similarity based	Demographic Similarity based	Interactions based	Popularity based	Proposed model
1	91	961	451	144	3	138	92
2	89	986	92	26	1	222	119
3	32	886	508	29	5	22	145
4	1	943	63	40	9	51	91
5	119	994	866	86	16	836	137
6	145	944	261	21	137	35	34
7	34	898	153	34	2	70	89
8	92	145	105	68	17	36	344
9	137	34	127	25	60	91	36
10	36	344	36	4	918	99	25

TABLE 7. Similarity based comparison of top-5 users in the dataset.

Trust Basis	Top-5 Users	Spatial Similarity	Interest Similarity	Demographic Similariy	Interaction Based	Popularity Based	Social Ties Strength	Credibility	Reliability	Dataset	**PITM**
Spatial similarity based	961	High	Low	Average	Average	Low	Weak	Low	Low	Un-Trusted	Un-Trusted
	986	High	Low	Low	High	Low	Average	Average	Low	Un-Trusted	Un-Trusted
	886	High	Average	Low	Average	Low	Weak	Average	Low	Un-Trusted	Un-Trusted
	943	High	Low	Average	Low	Low	Weak	High	Low	Un-Trusted	Un-Trusted
	994	High	Low	Low	Low	Low	Average	Average	Low	Un-Trusted	Un-Trusted
Interest Based	451	Average	High	Low	Low	Low	Weak	Average	Low	Un-Trusted	Un-Trusted
	92	High	High	High	High	Low	Strong	High	High	Trusted	Trusted
	508	Low	High	Average	Low	Low	Weak	Average	Low	Un-Trusted	Un-Trusted
	63	High	High	High	High	Low	Strong	High	Average	Trusted	Trusted
	866	Low	High	Average	Low	Low	Weak	High	Low	Un-Trusted	Un-Trusted
Demography Based	144	Low	Average	High	Low	High	Weak	Average	High	Un-Trusted	Trusted
	26	Average	Average	High	High	High	Strong	High	Average	Trusted	Trusted
	29	Average	Low	High	High	High	Average	High	High	Trusted	Trusted
	40	Average	Average	High	High	Average	Strong	High	Average	Trusted	Trusted
	86	Average	Low	High	High	Low	Strong	High	High	Trusted	Trusted
Interaction Based	3	Average	Average	Average	High	High	Strong	High	Average	Trusted	Trusted
	1	Average	High	High	High	High	Strong	High	Avrage	Trusted	Trusted
	5	High	Low	High	High	High	Strong	High	Average	Trusted	Trusted
	9	Average	Low	Average	High	Average	Strong	High	Average	Trusted	Trusted
	16	Low	Low	Average	High	Low	Weak	Low	High	Un-Trusted	Un-Trusted
Popularity based	138	Low	Average	Low	Low	High	Weak	Average	Low	Un-Trusted	Un-Trusted
	222	Low	Low	Average	Average	High	Weak	Average	Low	Un-Trusted	Un-Trusted
	22	Average	Low	Average	High	High	Average	Average	High	Trusted	Trusted
	51	Average	Average	High	High	High	Strong	High	Average	Trusted	Trusted
	836	High	Low	High	High	High	Strong	High	Low	Trusted	Trusted
Proposed Model	92	High	High	High	High	Low	Strong	High	High	Trusted	Trusted
	119	High	Average	High	High	Average	Strong	High	High	Trusted	Trusted
	145	High	Average	High	High	Low	Strong	High	High	Trusted	Trusted
	91	High	Average	High	High	High	Strong	High	High	Trusted	Trusted
	137	High	Average	High	High	Average	Strong	High	High	Trusted	Trusted

Ranges:

High/Strong (0.7< value <= 1.0),
Average (0.4< value <= 0.7),
Low/Weak (0< value <= 0.4).

B. EVALUATION OF THE PROPOSED USER TRUST MODEL

We evaluate our proposed model by applying it to the synthetic dataset. The details of the dataset are discussed in Section IV-A. We also applied our model to publicly available datasets to verify the performance of our proposed model. These datasets have limited information; therefore, we set weights according to the available information. We assigned equal weights to the available parameters, and set weights of the absent factors as zero.

Table 6, shows an exemplary list of top-10 trusted users in the dataset, top-10 users based on similarities, interactions, popularity, and top-10 trusted users predicted by our proposed model. By comparing these lists, we observed that the list of top-10 trusted users identified by our model are more similar to the top-10 trusted users in the dataset.

Table 7 shows the comparison of top-5 trusted users predicted by our proposed model with the top-5 users based on similarities, interactions, and popularity. From Table 7 we can observe that the top-5 users predicted by our proposed model are more close to its neighbors in all dimensions, i.e., spatial similarity, demographic similarity, interest similarity, and high interaction rate, unlike other methods based on similarities, interactions, and popularity. We also show a comparison of our proposed model with other models

TABLE 8. Relationship of interactions, and its direction with STS, and Reliability.

Interactions		$u \rightarrow v$		
		Low	Average	High
$u \leftarrow v$	Low	Weak STS Low Reliability	Weak STS Average Reliability	Weak STS High Reliability
	Average	Average STS Low Reliability	Average STS Average Reliability	Average STS High Reliability
	High	Strong STS Low Reliability	Strong STS Average Reliability	Strong STS High Reliability

Ranges:
High/Strong (0.7< value <= 1.0),
Average (0.4< value <= 0.7),
Low/Weak (0< value <= 0.4).

based on trust parameters, i.e., STS, reliability, and credibility. Unlike other models, the top trusted users predicted by our model are more credible, reliable, and have strong STS. We assigned high/strong, average and low/weak value to ranges of continuous values of similarities, interaction based trustworthiness, popularity, STS, credibility and reliability, to discretize them for better understanding, as shown in Table 7, 8 and 9.

TABLE 9. Relationship of interactions, and similarity with STS.

Similarities/ Interactions	Low	Average	High
Low	Weak STS	Weak STS	Average STS
Average	Weak STS	Average STS	Strong STS
High	Average STS	Strong STS	Strong STS

Ranges:
High/Strong (0.7< value <= 1.0),
Average (0.4< value <= 0.7),
Low/Weak (0< value <= 0.4).

Table 7 is divided into six portions for detail analysis. We observed that some users with high interactions are predicted as untrusted, due to low similarities score, and vice versa. From simulations, we observed that high/low interactions scores do not ensure trustworthiness/untrustworthiness of SN users, because interactions are two way (directed). Some users are active (high reliability), but they do not receive many responses and have lower intimacy with their neighbors, which results in weak STS. User 144 in 3^{rd} portion of Table 7, demography based top users, is untrusted in the dataset, due to fewer interactions, but it is predicted as a trusted user by our proposed model, due to its high reliability score. Similarly, some users with high interactions scores have weak STS/low reliability, i.e., more actions but fewer responses (high reliability and weak STS), and fewer actions but more responses (low reliability and strong STS). User 16 in 4^{th} portion and 836 in 5^{th} portion of Table 7, are their respective examples. Similarly, the top-5 users, predicted by popularity based ranking, are not necessarily similar to its neighbors, or they may not be more interactive with its neighbors. Hence, popularity (high number of friends/followers) also cannot determine trustworthiness.

We analysed the simulation results to extract relationships and effect of SN raw parameters on trust parameters. Table 8, shows the effect of interactions, and its direction on STS, and Reliability. Table 9, shows the effect of interactions, and similarities on STS.

From the above comparisons we draw two conclusions, i.e., i) similarity, interactions, and popularity alone cannot be a trust basis, and ii) for SN user's trust estimation, it is necessary to map/combine these SN parameters systematically.

We also compared the performance of our proposed trust model based on precision, recall, accuracy, and F-score with similarity-based trust model and interactions based trust model. The similarity-based trust modeling assumes that users with high similarities trust each other. We assumed the same for interactions based trust modeling and popularity based trust modeling, as well.

Precision is the ratio of real trusted users predicted to the total number of users identified as trusted by the proposed model, as shown in Equation 17. Precision shows the exactness of our model in trusted users prediction.

$$Precision = \frac{Predicted \bigcap Trusted}{Predicted} \tag{18}$$

Recall, as in Equation 18, is the ratio of real trusted users predicted by the model to the total number of real trusted users in the dataset. The recall represents the success of trusted users prediction out of all trusted users.

$$Recall = \frac{Predicted \bigcap Trusted}{Trusted} \tag{19}$$

Accuracy is the averaged sum of real trusted, and real untrusted users predicted by the model, as shown in Equation 20. We used balanced accuracy equation, to avoid the problem of imbalance dataset. The balanced accuracy is the mean of true positive rate (TPR) and true negative rate (TNR). Accuracy shows the overall performance of our model, considering both trusted and untrusted users.

$$Accuracy = \frac{\frac{Predicted \bigcap Trusted}{Trusted} + \frac{Predicted \bigcap UnTrusted}{UnTrusted}}{2} \tag{20}$$

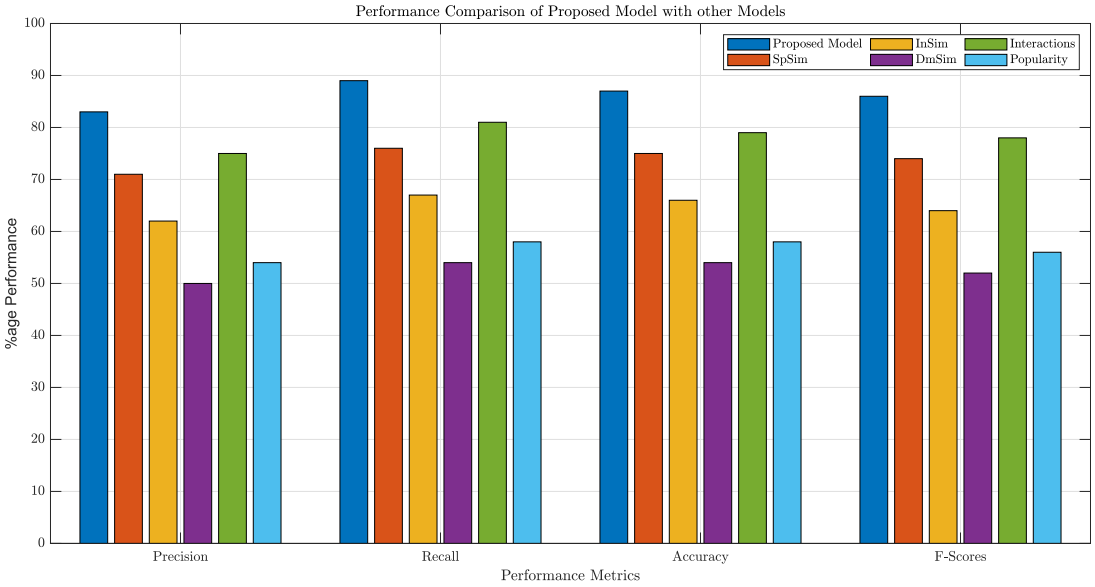

FIGURE 2. Performance comparison of the Proposed Model with other models; the performance metrics considered are precision, recall, accuracy and F-scores.

F1-score is the combination of precision and recall, which is computed by using Equation 21.

$$F - Score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{21}$$

In our experimental analysis, we found that our proposed algorithm outperforms other methods in term of the mentioned performance metrics. Figure 2, shows the performance comparison of our proposed algorithm with other methods. From these results, we concluded that our proposed model is not biased in prediction towards any of the class, i.e., trusted users, and untrusted users.

We used the metrics of mean reciprocal rank (MRR), mean average precision (MAP), normalized discounted cumulative gain (NDCG), and Kendall Tau correlation coefficient, to evaluate the quality of top-k users predicted by the considered models. Table 10, illustrates the experimental results of top-k trusted users predictions by our proposed model with other models. The four metrics measures the quality of the top-k predictions. The higher value of these metrics shows better quality of prediction.

Figure 3 shows the Kendall Tau correlation coefficient of our proposed model in comparison with other models, to show the quality of the top-k predictions. From Figure 3, it is evident that the Kendall Tau correlation coefficient of our proposed model is higher than other algorithms. Hence our algorithm identifies and ranks the trusted users more accurately than the other algorithms.

We have observed from simulation results that the performance of our model is better than the other methods, in accurate prediction of both trusted and untrusted users, in terms of precision, recall, accuracy, and F1-score. We also observed that the quality of predictions, in top-k users, by our model is better than the other models.

TABLE 10. Performance comparison of the proposed model with other models in top-k predictions based on MRR, MAP, NDCG and Kendall Tau.

Trust Basis	Top-k	MRR	MAP	NDCG	Kendall Tau
Spatial similarity based	5	0.000	0.000	0.000	0.20
	10	0.023	0.034	0.378	0.22
	20	0.056	0.144	0.654	0.41
	50	0.058	0.328	0.862	0.45
	100	0.036	0.392	0.895	0.43
Interest based	5	0.100	0.100	0.630	0.19
	10	0.060	0.070	0.564	0.21
	20	0.049	0.103	0.614	0.38
	50	0.028	0.161	0.695	0.39
	100	0.031	0.320	0.850	0.41
Demograpgy based	5	0.000	0.000	0.000	0.08
	10	0.014	0.014	0.333	0.12
	20	0.038	0.091	0.561	0.18
	50	0.024	0.146	0.656	0.15
	100	0.016	0.224	0.741	0.19
Interactions based	5	0.100	0.100	0.630	0.22
	10	0.066	0.083	0.605	0.25
	20	0.110	0.296	0.839	0.35
	50	0.060	0.399	0.867	0.37
	100	0.026	0.467	0.874	0.41
Popularity based	5	0.000	0.000	0.000	0.12
	10	0.023	0.034	0.378	0.18
	20	0.050	0.161	0.625	0.22
	50	0.028	0.191	0.699	0.26
	100	0.020	0.315	0.792	0.29
Proposed Model	5	**0.150**	**0.200**	**0.650**	**0.37**
	10	**0.230**	**0.788**	**0.996**	**0.42**
	20	**0.167**	**0.782**	**0.994**	**0.66**
	50	**0.084**	**0.781**	**0.985**	**0.68**
	100	**0.049**	**0.769**	**0.988**	**0.58**

The proposed model is an offline approach for SN user's trustworthiness estimation. The time complexity of the proposed model depends on the number of users N and number

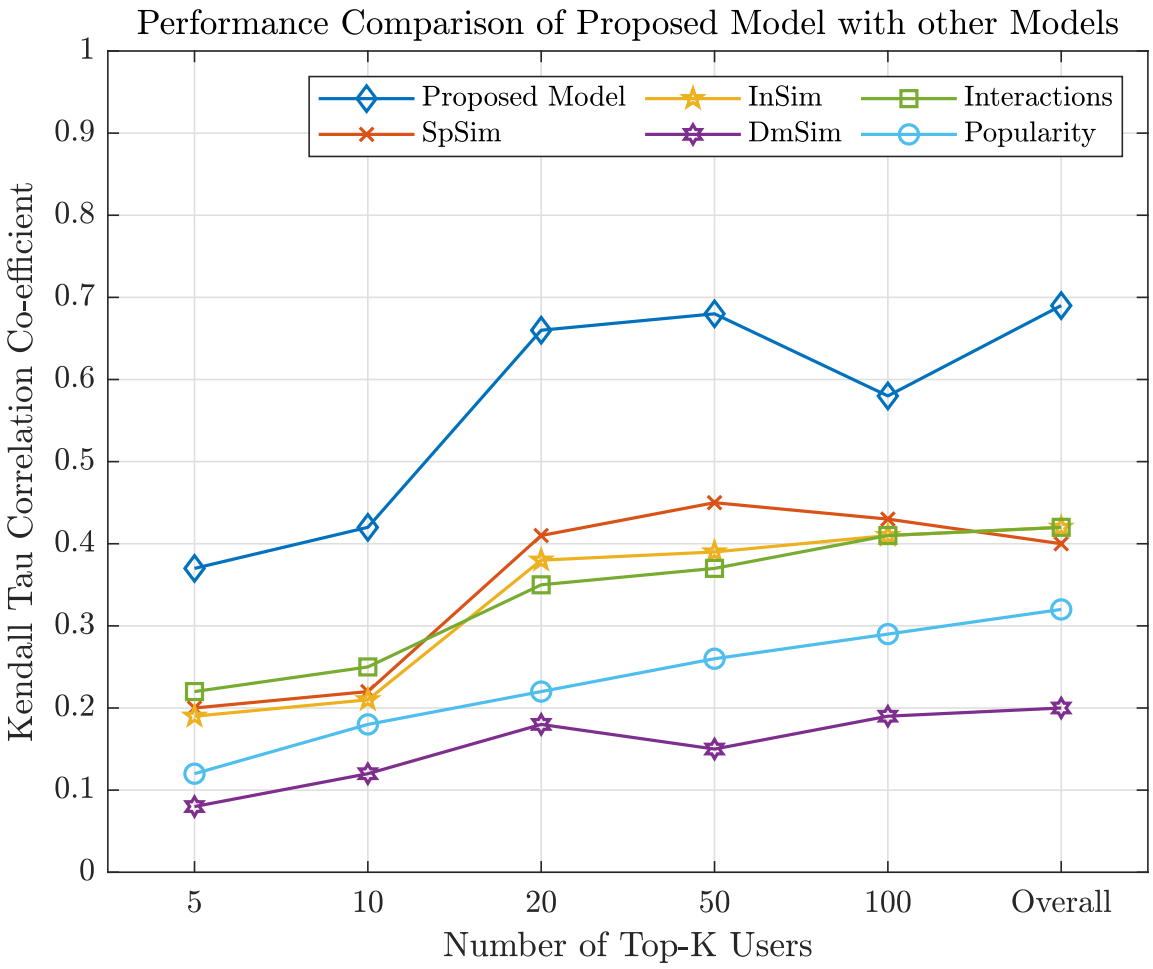

FIGURE 3. Kendall Tau Correlation Co-efficient Comparison of Proposed Model with other Models.

TABLE 11. Comparison of Top-5 friends recommendations by the proposed model with Top-5 trusted friends in the Dataset; for randomly selected users.

User	Top-5 Trusted friends					Top-5 trustworthy Friends Recommendation by our Proposed Model				
	FR1	FR2	FR3	FR4	FR5	TFR1	TFR2	TFR3	TFR4	TFR5
1	933	97	438	550	536	438	97	651	933	728
30	592	275	569	965	66	592	66	965	153	565
123	761	937	221	897	438	536	221	246	352	761
390	781	576	770	607	889	889	770	781	576	782
450	714	618	757	565	553	590	565	965	492	714

of attributes for similarities computation. We conducted our experiments in MATLAB 2019 (Mathworks, Inc., United States) on LG system (LG Electronics Nanjing Displays Company Ltd., China), with *Corei5*, 2.3 *GHz* processor, 8 *GHz* memory, and Windows 10 *Pro*64 − *bit* (Microsoft, United States). The algorithm elapsed 58.34 *sec*. for trust matrix computation. The execution time of the proposed algorithm was reduced to 12.21 *sec* (almost 80%) when we provide pre-computed similarity matrix to the algorithm.

C. TRUST BASED FRIENDS RECOMMENDATION

We applied our proposed user trustworthiness computation model to friends-recommendation (FR) system. We compared the results of our proposed model with the existing

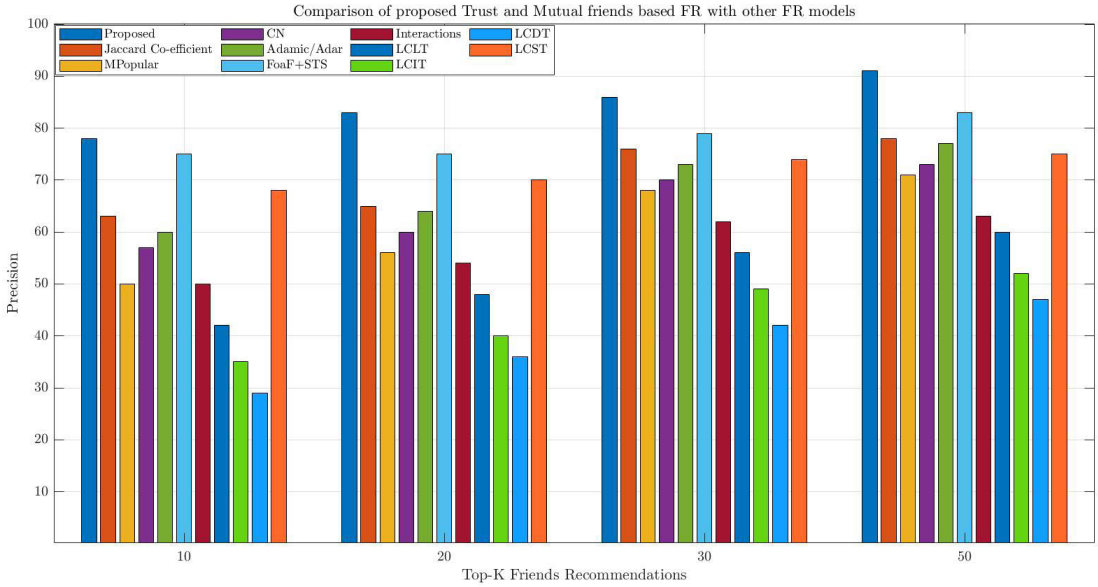

FIGURE 4. Comparison of Precision of our Proposed Model based FR with other FR Models.

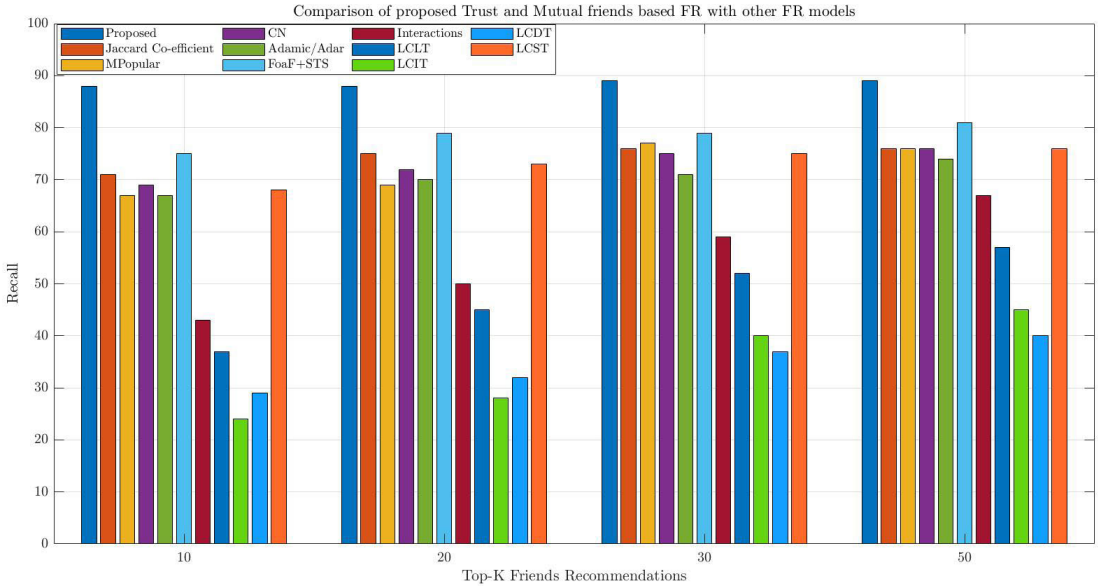

FIGURE 5. Comparison of Recall of our Proposed Model based FR with other FR Models.

state of the art models for experimental evaluation. The existing FR models we considered are common neighbors (CN), Mpopular (mutual friendship weighted by user's popularity), trust from friend of a friend (FoaF+STS) [24], interactions (neighbor of neighbors who are highly interactive), Jaccard coefficient [63], and Adamic/Adar [67], and LCIT (Linear combination of interest and trust). The trust in LCIT is estimated from tie-strength with FoaF. The idea of LCIT is derived from [68]. We also considered variant of the LC models such as LCLT, LCDT and LCST, which are linear combination of trust with location, demography and average similarity, respectively. We considered *precision@K*, *recall@K* and $F - scores$ as the performance metrics.

From the experimental analysis, we observed that the precision of our proposed model is better than other existing models in top-k trusted friends recommendation, as shown in Figure 4. Hence the probability of false trusted friends recommendations by our proposed model is lower than other algorithms.

From Figure 5 it is evident that the coverage of our model in trusted friends recommendation is higher than other existing algorithms. As we increase K the precision and recall are increased.

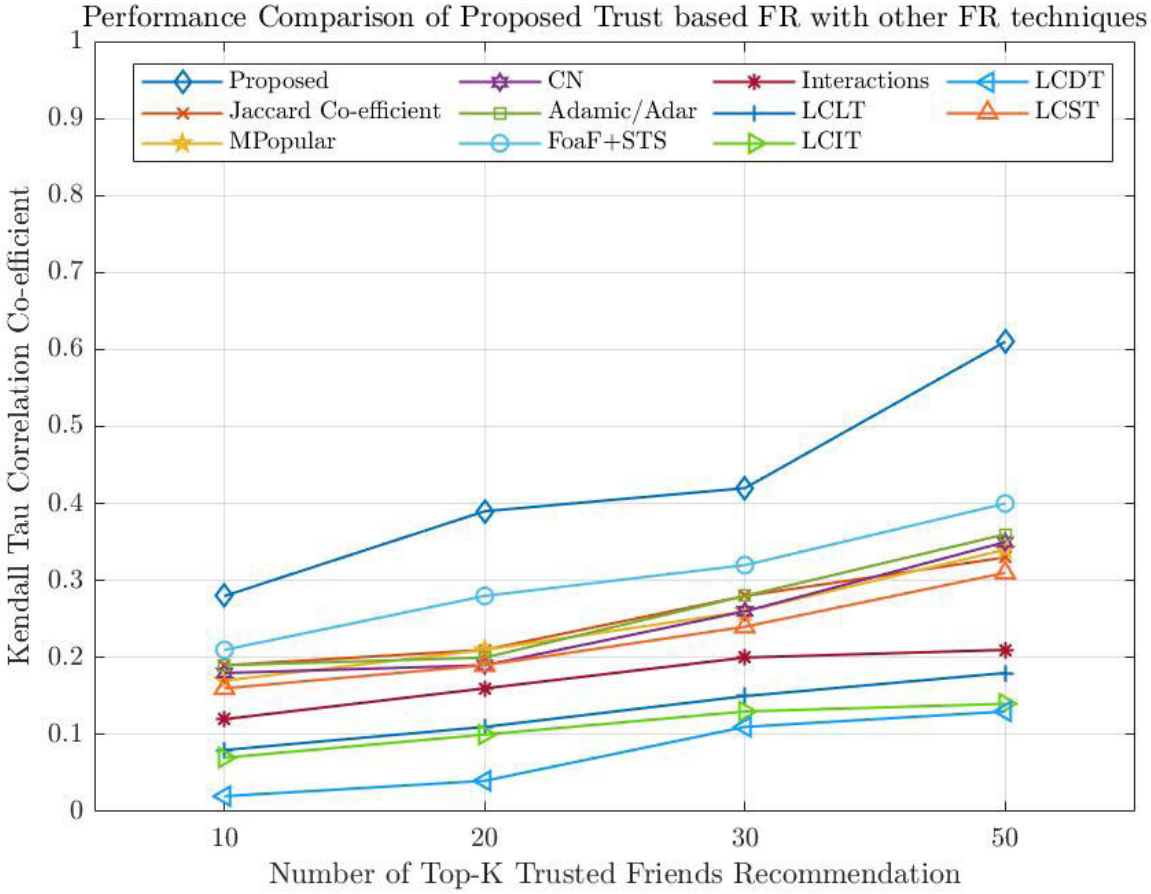

FIGURE 6. Kendall Tau correlation coefficient based comparison of proposed Trust-aware friends recommendation model with other friends recommendation models.

We compared the top-5 friends recommended by our proposed algorithm with the top-5 trusted friends present in the dataset, as shown in Table 11, to show the matching correlation between the top-5 recommendation lists. We found a high similarity between the two lists, which shows the feasibility of our proposed model in FR.

Figure 6 illustrates the comparison of our proposed model with other models in FR based on the Kendall Tau correlation coefficient, to measure the recommendation quality. The recommendation list which has higher Kendall Tau is more trusted and relevant to the target user with most trusted friends on the top of the list. Our proposed model outperforms other methods in FR, as shown in Figure 6, on the mentioned metrics.

V. CONCLUSION

We proposed a generic user's trustworthiness estimation model, which can be applied to many SN applications (trusted news sharing, trusted FR, trusted contents recommendations, etc.) by using proper weights. These weights can be estimated/assigned based on their importance in the target domain, and the user's behavior provided that the data is available. The proposed model can be extended to context-aware trust modeling in SN by introducing the user's competence in a specific domain. The focus of our paper is user's trustworthiness estimation in SN; however, this model can also be utilized to determine user-user (relational trust) by modification in weights and trust parameters user-user reliability and credibility.

We leveraged the SN user's attributes, interaction behaviors, and relationships for SN user's trustworthiness computation. We derived a set of SN intermediate trust parameters by mapping SN raw parameters to their relevant intermediate trust parameters. We mapped the intermediate trust parameters to the SN trust parameters, i.e., social ties strength, credibility, and reliability, to compute the user's implicit trustworthiness score. The objective of intermediate trust parameters derivation is to map the SN raw parameters into SN trust parameters in a systematic manner. The SN raw parameters are user-user interactions and relationships, user's spatial information, demographic information, and interests.

We evaluated the feasibility of our proposed model using synthetic data. Our model outperformed the existing methods in terms of precision, recall, accuracy, and F1-scores, as illustrated in the results. We observed that the MRR, MAP, NDCG and Kendall Tau correlation coefficient of

top-k users predicted by our model is higher than other methods, which illustrate better prediction quality in the top-k trusted users. For verification of our results, we used Twitter interactions data and publicly available datasets, i.e., movielens and filmtrust.

The proposed model is multi-facet, which is dependent on multiple SN parameters for trust calculation. These factors are activeness, popularity, and similarities. Based on the simulation results, we can conclude that activeness, popularity, and similarities alone cannot determine the user's trustworthiness. Some users may be active, but they are not necessarily trustworthy and vice versa. Similarly, the above is valid for popularity and similarities as well.

The existing models mostly focus on relational trust. The relational trust also referred to as social ties strength. These models are for SN trust-aware applications such as information spreading, contents recommendations, and FR, to improve their performance, but they do not ensure trustworthy in SN. Relational trust is a trust between two neighbor users, these users can have a mutual-trust/intimacy/strong relationship, but it is not necessary that they both are trustworthy. Such users can be active, influential, and popular, but not necessarily trustworthy. Our model integrates credibility and reliability along with STS for trust computation, to depict a better representation of user trustworthiness in SN.

We applied our trust model for trusted FR and compared the result with existing FR models. The FR algorithm ranks the candidate friends based on trustworthiness and friendship similarity. The user trustworthiness in recommender systems prevents recommendations from/of malicious/untrustworthy users by minimizing their recommendation scores. Our proposed trust model is independent of explicit trust, i.e., user-user/items ratings; therefore, it is resistant to the problems of trust sparsity and cold start users. We measured the performance based on various performance metrics such as precision@K, recall@K, and Kendall Tau. We found, from experiments that the performance of our trust model for FR was superior to the existing FR models based on the mentioned metrics. We found that the top friends recommended by our system are more similar, to the target users, demographically, spatially, have high interest similarity score and have stronger social ties with target neighbors.

REFERENCES

[1] G. Donnelly. (Aug. 16, 2018). *The WordStream: 75 Super-Useful Facebook Statistics for 2018*. Accessed: Aug. 30, 2018. [Online]. Available: https://www.wordstream.com/blog/ws/2017/11/07/facebook-statistics

[2] Y. A. Kim and R. Phalak, "A trust prediction framework in rating-based experience sharing social networks without a Web of trust," *Inf. Sci.*, vol. 191, pp. 128–145, May 2012.

[3] V. Podobnik, D. Striga, A. Jandras, and I. Lovrek, "How to calculate trust between social network users?" in *Proc. 20th Int. Conf. Softw., Telecommun. Comput. Netw.*, Sep. 2012, pp. 1–6.

[4] S. Nepal, W. Sherchan, and C. Paris, "STrust: A trust model for social networks," in *Proc. IEEE 10th Int. Conf. Trust, Secur. Privacy Comput. Commun.*, Nov. 2011, pp. 841–846.

[5] C.-N. Ziegler and J. Golbeck, "Investigating interactions of trust and interest similarity," *Decis. Support Syst.*, vol. 43, no. 2, pp. 460–475, 2007.

[6] P. Massa and P. Avesani, "Trust-aware recommender systems," in *Proc. ACM Conf. Rec. Syst.*, 2007, pp. 17–24.

[7] M. Jamali and M. Ester, "TrustWalker: A random walk model for combining trust-based and item-based recommendation," in *Proc. 15th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2009, pp. 397–406.

[8] J. B. Rotter, "A new scale for the measurement of interpersonal trust," *J. Personality*, vol. 35, no. 4, pp. 651–665, 1967.

[9] J. B. Rotter, "Generalized expectancies for interpersonal trust," *Amer. Psychol.*, vol. 26, no. 5, p. 443, 1971.

[10] R. Zhou and K. Hwang, "PowerTrust: A robust and scalable reputation system for trusted peer-to-peer computing," *IEEE Trans. Parallel Distrib. Syst.*, vol. 18, no. 4, pp. 460–473, Apr. 2007.

[11] L. Xiong and L. Liu, "PeerTrust: Supporting reputation-based trust for peer-to-peer electronic communities," *IEEE Trans. Knowl. Data Eng.*, vol. 16, no. 7, pp. 843–857, Jul. 2004.

[12] S. Paradesi, P. Doshi, and S. Swaika, "Integrating behavioral trust in Web service compositions," in *Proc. IEEE Int. Conf. Web Services*, Jul. 2009, pp. 453–460.

[13] R. B. Zajonc, "Feeling and thinking: Preferences need no inferences," *Amer. Psychol.*, vol. 35, no. 2, p. 151, 1980.

[14] P. M. Romer, "Thinking and feeling," *Amer. Econ. Rev.*, vol. 90, no. 2, pp. 439–443, 2000.

[15] D. Helbing, "A mathematical model for the behavior of individuals in a social field," *J. Math. Sociol.*, vol. 19, no. 3, pp. 189–219, 1994.

[16] S. Webb, J. Caverlee, and C. Pu, "Social honeypots: Making friends with a spammer near you," in *Proc. CEAS*, 2008, pp. 1–10.

[17] J. Caverlee and S. Webb, "A large-scale study of MySpace: Observations and implications for online social networks," in *Proc. ICWSM*, 2008, pp. 1–9.

[18] Z. Yan and R. Yan, "Formalizing trust based on usage behaviours for mobile applications," in *Proc. Int. Conf. Auton. Trusted Comput.* Berlin, Germany: Springer, 2009, pp. 194–208.

[19] Z. Yan, V. Niemi, Y. Dong, and G. Yu, "A user behavior based trust model for mobile applications," in *Proc. Int. Conf. Auton. Trusted Comput.* Berlin, Germany: Springer, 2008, pp. 455–469.

[20] S. Nepal, W. Sherchan, and A. Bouguettaya, "A behaviour-based trust model for service Web," in *Proc. IEEE Int. Conf. Service-Oriented Comput. Appl. (SOCA)*, Dec. 2010, pp. 1–4.

[21] S. Adali, R. Escriva, M. K. Goldberg, M. Hayvanovych, M. Magdon-Ismail, B. K. Szymanski, W. A. Wallace, and G. Williams, "Measuring behavioral trust in social networks," in *Proc. IEEE Int. Conf. Intell. Secur. Inform.*, May 2010, pp. 150–152.

[22] Z. Yan, P. Zhang, and R. H. Deng, "TruBeRepec: A trust-behavior-based reputation and recommender system for mobile applications," *J. Pers. Ubiquitous Comput.*, vol. 16, no. 5, pp. 485–506, Jun. 2012.

[23] V. Buskens, "The social structure of trust," *Social Netw.*, vol. 20, no. 3, pp. 265–289, 1998.

[24] J. Golbeck, B. Parsia, and J. Hendler, "Trust networks on the semantic Web," in *Proc. Int. Workshop Cooperat. Inf. Agents*. Berlin, Germany: Springer, 2003, pp. 238–249.

[25] C.-N. Ziegler and G. Lausen, "Spreading activation models for trust propagation," in *Proc. IEEE Int. Conf. e-Technol., e-Commerce e-Service*, Mar. 2004, pp. 83–97.

[26] J. A. Golbeck, "Computing and applying trust in Web-based social networks," Ph.D. dissertation, Dept. Comput. Sci., Univ. Maryland, College Park, MD, USA, 2005.

[27] Y. Zhang, H. Chen, and Z. Wu, "A social network-based trust model for the semantic Web," in *Proc. Int. Conf. Auton. Trusted Comput.* Berlin, Germany: Springer, 2006.

[28] M. Maheswaran, H. C. Tang, and A. Ghunaim, "Towards a gravity-based trust model for social networking systems," in *Proc. 27th Int. Conf. Distrib. Comput. Syst. Workshops (ICDCSW)*, Jun. 2007, p. 24.

[29] Y. A. Kim, M.-T. Le, H. W. Lauw, E.-P. Lim, H. Liu, and J. Srivastava, "Building a Web of trust without explicit trust ratings," in *Proc. IEEE 24th Int. Conf. Data Eng. Workshop*, Apr. 2008, pp. 531–536.

[30] Y. Zuo, W.-C. Hu, and T. O'Keefe, "Trust computing for social networking," in *Proc. 6th Int. Conf. Inf. Technol., New Gener.*, Apr. 2009, pp. 1534–1539.

[31] G. Liu, Q. Yang, H. Wang, and A. X. Liu, "Three-valued subjective logic: A model for trust assessment in online social networks," *IEEE Trans. Depend. Sec. Comput.*, to be published.

[32] T. Cheng, G. Liu, Q. Yang, and J. Sun, "Trust assessment in vehicular social network based on three-valued subjective logic," *IEEE Trans. Multimedia*, vol. 21, no. 3, pp. 652–663, Mar. 2019.

[33] H. Liu, E.-P. Lim, H. W. Lauw, M.-T. Le, A. Sun, J. Srivastava, and Y. A. Kim, "Predicting trusts among users of online communities: An Epinions case study," in *Proc. 9th ACM Conf. Electron. Commerce*, 2008, pp. 310–319.
[34] D. E. Byrne, *The Attraction Paradigm*, vol. 11. New York, NY, USA: Academic, 1971.
[35] C.-N. Ziegler and G. Lausen, "Analyzing correlation between trust and user similarity in online communities," in *Trust Management—iTrust* (Lecture Notes in Computer Science), vol. 2995, C. Jensen, S. Poslad, and T. Dimitrakos, Eds. Berlin, Germany: Springer, 2004.
[36] F. Menczer, "Evolution of document networks," *Proc. Nat. Acad. Sci. USA*, vol. 101, no. 1, pp. 5261–5265, 2004.
[37] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks," *Annu. Rev. Sociol.*, vol. 27, no. 1, pp. 415–444, 2001.
[38] R. Sinha and K. Swearingen, "Comparing recommendations made by online systems and friends," in *Proc. DELOS-NSF Workshop Personalization Rec. Syst. Digit. Libraries*, 2001, pp. 1–6.
[39] K. Swearingen and R. Sinha, "Beyond algorithms: An HCI perspective on recommender systems," in *Proc. ACM SIGIR Workshop Rec. Syst.*, 2001, vol. 13, nos. 5–6, pp. 1–11.
[40] C.-N. Ziegler and J. Golbeck, "Investigating correlations of trust and interest similarity-do birds of a feather really flock together," *Decis. Support Syst.*, vol. 43, no. 2, pp. 1–34, 2005.
[41] K. Akilal, H. Slimani, and M. Omar, "A robust trust inference algorithm in weighted signed social networks based on collaborative filtering and agreement as a similarity metric," *J. Netw. Comput. Appl.*, vol. 126, pp. 123–132, Jan. 2019.
[42] J. O'Donovan and B. Smyth, "Trust in recommender systems," in *Proc. 10th Int. Conf. Intell. User Interfaces*, 2005, pp. 167–174.
[43] F. Ullah and S. Lee, "Social content recommendation based on spatial-temporal aware diffusion modeling in social networks," *Symmetry*, vol. 8, no. 9, p. 89, 2016.
[44] M. L. Ginsberg, "Multivalued logics: A uniform approach to reasoning in artificial intelligence," *Comput. Intell.*, vol. 4, no. 3, pp. 265–316, 1988.
[45] J. Golbeck, "Generating predictive movie recommendations from trust in social networks," in *Proc. Int. Conf. Trust Manage.* Berlin, Germany: Springer, 2006, pp. 93–104.
[46] J. Golbeck, "Introduction to computing with social trust," in *Computing with Social Trust*. London, U.K.: Springer, 2009, pp. 1–5.
[47] C. Hess and C. Schlieder, "Trust-based recommendations for documents," *AI Commun.*, vol. 21, nos. 2–3, pp. 145–153, 2008.
[48] P. Massa and P. Avesani, "Trust metrics on controversial users: Balancing between tyranny of the majority," *Int. J. Semantic Web Inf. Syst.*, vol. 3, no. 1, pp. 39–64, 2007.
[49] M. D. Cock and P. P. da Silva, "A many valued representation and propagation of trust and distrust," in *Proc. Int. Workshop Fuzzy Logic Appl.* Berlin, Germany: Springer, 2005, pp. 114–120.
[50] G. Gans, M. Jarke, S. Kethers, and G. Lakemeyer, "Modeling the impact of trust and distrust in agent networks," in *Proc. AOIS*, 2001, pp. 45–58.
[51] J. Golbeck and A. Mannes, "Using trust and provenance for content filtering on the semantic Web," in *Proc. MTW*, 2006, pp. 3–4.
[52] P. Massa and B. Bhattacharjee, "Using trust in recommender systems: An experimental analysis," in *Proc. Int. Conf. Trust Manage.* Berlin, Germany: Springer, 2004, pp. 221–235.
[53] R. C. Mayer, J. H. Davis, and F. D. Schoorman, "An integrative model of organizational trust," *Acad. Manage. Rev.*, vol. 20, no. 3, pp. 709–734, 1995.
[54] D. J. McAllister, "Affect- and cognition-based trust as foundations for interpersonal cooperation in organizations," *Acad. Manage. J.*, vol. 38, no. 1, pp. 24–59, 1995.
[55] L. Mui, M. Mohtashemi, and A. Halberstadt, "A computational model of trust and reputation," in *Proc. 35th Annu. Hawaii Int. Conf. Syst. Sci.*, Jan. 2002, pp. 2431–2439.
[56] S. Noh, "Calculating trust using aggregation rules in social networks," in *Proc. Int. Conf. Auton. Trusted Comput.* Berlin, Germany: Springer, 2007, pp. 361–371.
[57] J. O'Donovan, "Capturing trust in social Web applications," in *Computing With Social Trust*. London, U.K.: Springer, 2009, pp. 213–257.
[58] S. Cheng, B. Zhang, G. Zou, M. Huang, and Z. Zhang, "Friend recommendation in social networks based on multi-source information fusion," *Int. J. Mach. Learn.*, vol. 10, no. 5, pp. 1003–1024, Feb. 2018.
[59] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[60] M. Dillon, *Introduction to Modern Information Retrieval*, G. Salton and M. McGill, Eds. New York, NY, USA: McGraw-Hill, 1983, pp. 402–403.
[61] R. W. Sinnott, "Virtues of the haversine," *Sky Telescope*, vol. 68, no. 2, p. 159, 1984.
[62] M. C. Palmer, "Calculation of distance traveled by fishing vessels using GPS positional data: A theoretical evaluation of the sources of error," *Fisheries Res.*, vol. 89, no. 1, pp. 57–64, 2008.
[63] P. Jaccard, "Distribution de la Flore Alpine dans le Bassin des Dranses et dans quelques Régions voisines," *Bull. Soc. Vaudoise Sci. Naturelles*, vol. 37, pp. 241–272, Jan. 1901.
[64] J. Khan and S. Lee, "Online social networks (OSN) evolution model based on homophily and preferential attachment," *Symmetry*, vol. 10, no. 11, p. 654, 2018. [Online]. Available: https://www.mdpi.com/2073-8994/10/11/654
[65] *GroupLens Dataset (ml-100k)*. Accessed: Jan. 2, 2019. [Online]. Available: http://grouplens.org/datasets/movielens/
[66] G. Guo, J. Zhang, and N. Yorke-Smith, "A novel Bayesian similarity measure for recommender systems," in *Proc. 23rd Int. Joint Conf. Artif. Intell. (IJCAI)*, 2013, pp. 2619–2625.
[67] L. A. Adamic and E. Adar, "Friends and neighbors on the Web," *Soc. Netw.*, vol. 25, no. 3, pp. 211–230, 2003.
[68] J. Chen, W. Geyer, C. Dugan, M. Muller, and I. Guy, "Make new friends, but keep the old: Recommending people on social networking sites," in *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, 2009, pp. 201–210.

JEBRAN KHAN received the B.Sc. and M.Sc. degrees in computer systems engineering from the University of Engineering and Technology at Peshawar, Peshawar, Pakistan. He is currently pursuing the Ph.D. degree in electronics and information engineering with Korea Aerospace University, Goyang, South Korea. His research interests include social networks analysis, modeling, frameworks, and its applications.

SUNGCHANG LEE received the B.S. degree from Kyungpook National University, in 1983, the M.S. degree in electrical engineering from the Korea Advanced Institute of Science and Technology (KAIST), in 1985, and the Ph.D. degree in electrical engineering from Texas A&M University, in 1991. From 1985 to 1987, he was with KAIST, as a Researcher, where he worked on Image Processing and Pattern Recognition Projects. From 1992 to 1993, he was a Senior Researcher with the Electronics and Telecommunications Research Institute (ETRI), South Korea. From 2004 to 2009, he was the Director of Government Project on Intelligent Smart Home Security & Automation Service Technology. In 2009, he was the Vice President of the Institute of Electronics and Information Engineers (IEIE), South Korea, and also the Director of the Telecommunications Society, South Korea. Since 1993, he has been a Faculty with Korea Aerospace University, Goyang, South Korea, where he is currently a Professor with the School of Electronics, Telecommunication & Computer Engineering.