A strategy for trust propagation along the more trusted paths

KIANINEJAD, Marzieh and KABIRI, Peyman <http://orcid.org/0000-0001-5143-0498>

Published version

KIANINEJAD, Marzieh and KABIRI, Peyman (2018). A strategy for trust propagation along the more trusted paths. In: 2018 3rd Conference on Swarm Intelligence and Evolutionary Computation (CSIEC). IEEE, 1-6.

Copyright and re-use policy

A strategy for trust propagation along the more trusted paths

Marzieh Kianinejad
Department of Computer Engineering
Iran University of Science and Technology
Tehran, Iran
Marzieh9.kiani@gmail.com

Peyman Kabiri
Department of Computer Engineering
Iran University of Science and Technology
Tehran, Iran
Peyman.kabiri@iust.ac.ir

***Abstract*— The main goal of social networks are sharing and exchanging information among users. With the rapid growth of social networks on the Web, the most of interactions are conducted among unknown individuals. On the other hand, with increasing the biased behaviors in online communities, ability to assess the level of trustworthiness of a person before interacting with him has an important influence on users' decisions. Trust inference is a method used for this purpose. This paper studies propagating trust values along trust relationships in order to estimate the reliability of an anonymous person from the point of view of the user who intends to trust him/her. It describes a new approach for predicting trust values in social networks. The proposed method selects the most reliable trust paths from a source node to a destination node. In order to select the optimal paths, a new relation for calculating trustable coefficient based on previous performance of users in the social network is proposed. In ciao dataset there is a column called helpfulness. Helpfulness values represent previous performance of users in the social network. Advantages of this algorithm is its simplicity in trust calculation, using a new entity in dataset and its improvement in accuracy. The results of the experiments on Ciao dataset indicate that accuracy of the proposed method in evaluating trust values is higher than well-known methods in this area including TidalTrust, MoleTrust methods.**

Keywords- Trust, Trust propagation, Trust network, Trust chain, Trust paths

I. INTRODUCTION

Nowadays, popularity of social networks such as Facebook, Twitter and YouTube is undeniable. Each of these networks has different applications and provides distinct services to its users. These networks with dozens or hundreds of millions of members are considered as powerful tools for managing the flow of information. Unlimited number of users, the widespread impact of changing structure of the social communications within communities and facilitating the spread of news and information are partial reasons behind the importance of these networks. Although these networks provide researchers and users with a wealth of information, sometimes they can confuse their users. There are many types of networks, especially social networks that are largely dynamic. These networks grow rapidly adding new links that each represent new interactions between the network nodes. The most of interactions taking place within a social network are between anonymous people. Individuals in virtual societies may express their trust to other members, which, is referred to by explicit trust. These statements create a network, called trust network, and is used to predict the trust relationships within the network. In other words, the trust network is a network, in which, each person is allowed to give credence to the people whom he/she interacts with them. This trust can be complete trust, partial trust, distrust or ignorance. The trust value is mainly based on the history and the background of interactions between individuals. The trust network is represented as a directed graph $G(V, E)$, in which, V is a set of nodes and E is a series of edges, such that each edge $e = (u, v) \in E$ represents the trust between the nodes u and v. Weight of the edges are used to show the trust score between the nodes. Trust networks can be divided into three categories: binary networks, signed networks, and valued networks. Dependent on the type of trust network, the intention can be predicting trust values between its users. Trust prediction methods can be divided into three general categories as follows: trust prediction methods in the binary trust networks, trust prediction methods in signed trust networks and trust prediction methods in valued networks [1]. In approaches that are based on machine learning [2-4], they first extract some of the features that are structurally meaningful and then use a classifier to solve a prediction problem. Two algorithms are suggested for trust prediction in social networks [5], in one of them trust values in all steps of the algorithm is 0 or 1. In the other algorithm, the trust values during the execution of the algorithm belong to the continuous interval [0-1] but in the last step, predicting the final trust value between the two users, the predicted value will be one of the values of 0 or 1. In the trust prediction approach in valued trust networks [5-7], continuous values of [0-1] are used to predict the trust and to express the level of trust of users to each other. Therefore, trust worthier is a user, the more reliable the user will be. On the other hand, as the trust value is close to 0, the user will be less reliable. Now, with trust network, trust can be deduced from anonymous users by propagating trust throughout existing trust relationships. Propagation of trust is one of the main mechanisms for trust inference. This approach consider the trust paths between the source node to the target node in order to predict level of trust with minimum error [8]. In spite of its importance, there are no determined rules to follow that may lead to selection of the most reliable trust paths. As a result, researchers in their works use different strategies to find optimal trust paths from the source user to the target user that guarantees trust prediction with highest precision.

Distrust is one of the controversial issues in the trust prediction. Indeed, distrust and trust are the opposite of each other, and distrust is defined as a contradiction to the trust [9]. Research conducted in this area shows that distrust in the social networks is as important as the trust itself. Reported work by Gans et al. [10] was one of the first research tried to recognize the importance of distrust and considered clear distinction between trust and distrust. This research has pointed out that distrust in social networks can be very important. However, in many works, only trust has been considered and distrust has been completely ignored [9, 11, 12]. Our proposed method for predicting trust is in the category of trust prediction in valued networks, which considered trust values. In this paper, in order to implement the proposed method to predict trust values, the first step is to select the strongest and the shortest paths. Thus, users should be ranked according to their previous performance in the social network. For this purpose, information provided in the dataset is used. Finally, the new trust prediction method based on trust propagation is proposed. Basis of the local trust metrics is propagation of trust values along trust paths. For this purpose, the source user asks its neighbours to assess the trust value of an unknown user. If its neighbours directly know this unknown user, they will return their information to the source user. Otherwise, they will ask their neighbours about it. Hence, trust information propagates along the trust paths. In this research, propagation of trust information is limited to the shortest path among the most trusted paths since trust values obtained from these paths are more reliable. Additionally, research suggests a trust model based on the importance of users in social network. Therefore, nodes are ranked based on their previous performance in social network in order to predicted trust values. To this end, using the structural data of the network, firstly, users are ranked in terms of their significance and their impact on trust relationships, then using these ratings, a new relationship is proposed to predict the trust value.

The rest of this paper is organized as follows: section 2 briefly describes previous works in measuring trust that are related to the proposed method. In section 3 the proposed method is explained. In section 4, the proposed method is evaluated and in the last section, concluding remarks are presented.

II. RELATED WORKS

Many scholars have worked in the field of propagation and composition of trust. TidalTrust algorithm [12] is one of the most important methods for trust prediction. In this method, trust calculations are reciprocal, in the sense that initially, trust value moves from the source node towards the target node and stores a series of necessary information. In the next step, it returns from the destination node to the source node where it calculates the value of trust. The general process of this approach begins with the source node searching for the destination node and calculating the trust value of each one of its neighbours to the destination. The path from the source node to the current depth, as well as the strength of the path is stored. Each one of the neighbours keeps the most trusted path. Whenever a path from the source node to the destination node is found, that depth is considered as the maximum depth. Since the search of the corresponding graph is breath first search, the first found path has the lowest depth. However, search continues to find other paths at this minimum depth. Ziegler et al. [6] proposed Apleseed algorithm for valued trust networks. In this algorithm, a node is considered as the nucleus node, and the energy E is injected to it. This energy is released and is distributed between nodes that are near the nucleus. The higher the edge weight between the core nodes and its neighbouring nodes, the more energy that node will receive. This algorithm works with partial information from the graph, that is, at any one time, only the information of the nodes that hold energy are needed and the graph information is not required. Golbeck and Kuter [13] have proposed Sunny algorithm for trust prediction. Since the goal is to estimate the amount of trust among users who are not directly related to each other, an assurance measure can be used to estimate the trust value. This algorithm suggests an explicit probabilistic interpretation to express confidence in social networks, and actually uses a probabilistic approach to find the most reliable sources. This algorithm has been compared against the TidalTrust algorithm. Results indicate that the SUNNY algorithm is more accurate than the TidalTrust algorithm. Researchers who proposed the SUNNY algorithm claimed that their algorithm is the first algorithm, which considers a confidence criteria for predicting trust values. In MoleTrust algorithm [14] the importance of a user's opinion about trust to others is dependent on the trustworthiness of source node. The trust calculation for the source node who trusts to target node starts from the source node and is propagated through the edges to the destination node. The algorithm consists of two stages. After determining the source user and the destination user, the first step is converting the trust graph to a directional graph without any cycle. The second step is to scroll through this directed graph and calculate the trust values for the nodes that are visited. Kim and Sung [8] have presented a trust-based deduction model based on the propagation of trust and reinforcement learning. In other words, in this paper they have shown how the length of trust paths and aggregation methods can effect accuracy of the results. Ghaemi and Shakeri [15] have proposed a method to improve accuracy of the multiplication strategy. In this method, they introduced the recommendation trust between the two nodes based on the similarity of the two nodes' opinions in assessing trust to others. Initially, propagated trust is estimated based on recommendation trust and to achieve more precision. Then, the level of trust is calculated by multiplying trust values along the trust paths. Use of fuzzy logic and its operators for inferring trust has been suggested in many works [16-18]. Lesani and Montazeri [19] presented a fuzzy trust inference model. If the weighted average method is used to overcome the problem of conflicting information in the network, the trustor (a node who trusts another node) and trustee (a node who is trusted by another node [20]), will remain unaware of the existence of this contradictory information on the network. Therefore, in order to determine the trust level of users, Persian language vocabulary has been used and an algorithm based on TidalTrust has been presented. Results show the superiority of the fuzzy trust composition method to the method of explicit trust composition, especially, in the presence of contradictory information. Hossteinzadeh aghdam et al. [21] use trust information to produce more reliable recommendations in a trust-aware recommender system. In this research, they use resistive circuits and their physical formulas for prediction.

Guha et al. [22] use matrix multiplication operations. With matrix multiplication, the trust propagates in one-step. Finally, a $p\hat{}k$ matrix is generated, in which, (i, j) represents the amount of propagated trust between i and j after the k atomic propagation. One of the issues that exists in this model is not considering the most reliable paths.

III. PROPOSED METHOD

Where the source node does not have direct interaction with the target node, calculating value of the trust between the two entities is one of the challenges in the area of trust management. If the goal is prediction level of trust according to path discovery, the first step is to choose the optimal paths and then propagate trust values along them. At this point, the main problem is the length of these optimal paths. Trust prediction models are highly influenced by length of the paths. Jøsang et al. [23] stated that longer the trust path is the weaker the predicted trust will be. Any combination of path length and combination methods can be used to select the best paths of trust in order to predict trust value. In this research, goal is to find the shortest and the most trusted path between the source and the destination node to calculate the trust value. For this purpose, the strongest paths among shortest paths are selected. One of the advantage of this strategy is that when the strategy considers the shortest trust paths from the source node to the target node, it doesn't need to use the information available on all the trust paths (various length paths) to estimate the value of trust of the target node. Therefore, this strategy decreases the computational complexity. In other words, the more trusted shortest paths are more likely to have an accurate trust prediction. One way to understand the importance of nodes and their reliability in a trust path is that the user ranks other users based on their previous performance in the social network. In trust networks with node rankings, using the structural information of the network, the impact of each node in the relationship of trust is calculated. In this paper, background of the users in the social network is used and users are ranked based on their previous performance. For this purpose, information that is provided in the dataset are used. The Ciao dataset contains a column called helpfulness. The helpfulness values present the usefulness of opinions of a user from the perspective of other users in the network. Helpfulness values determine which nodes are more reliable for calculating trust values. Using helpfulness values to find the most trusted paths makes the algorithm significantly resistant against malicious nodes. Therefore, helpfulness values can be used to identify the most trusted paths to propagate trust values. Thus, a coefficient is assigned to each user along the trust path according to (1):

$$a_i = \frac{\sum_{j=1}^{k} h_j}{N}. \quad (1)$$

In (1), h_j refers to the helpfulness of user J and N is number of all users in trust network. In next step, the strength of trust paths is calculated according to (2):

$$strength(1.i) = \frac{\sum_{j \in outlink\ nodes(1)} trust(1,j) * rank(j) * strength(j,i)}{\sum_{j \in outlink\ nodes(1)} trust(1,j) * rank(j} \quad (2)$$

In (2), *outlink nodes(1)* denotes immediate neighboring nodes whom the first node trusts to them in any path and rank(j) refers to rank of user(j) which has been calculated according to (1). The final stage is the use of a weighted average to predict the trust between the two nodes in the network according to (3).

$$trust(1,n) = \frac{\sum_{k \in inlink\ nodes(n)} strength(1,k) * rank(k) * trust(k,n)}{\sum_{k \in inlink\ nodes(n)} strength(1,k) * rank(k)} \quad (3)$$

where *inlink nodes(n)* refers to nodes that are immediate neighboring nodes that trust to node(n) and rank(k) refers to rank of user(k) which has been calculated according to (1).

IV. RESULTS

To evaluate accuracy of the proposed algorithm and compare it against TidalTrust and MoleTrust, all three methods are evaluated using the Ciao[1] dataset that is a real-world dataset. The Ciao website is an e-commerce website where users write their opinions about the items, announce their trust to other users and, most importantly, they can rate other users' opinions. Trust information has two key uses. First, many users are looking at a particular category rather than a specific product. Therefore, a series of items should be displayed to them. To select the right items, they can use trust information. For the second key use, whenever a specific product is displayed to the user, a number of comments from other users about that product should also be displayed to him. Since the total number of related comments is often too many and study of them is time consuming, considering that quality of these comments will vary as well, hence, trust information can be used. Thus, with the help of trust relationships between the user and those who gave the ratings, and users who have expressed their views, it would be easy to choose the specific comments that will be useful for the user. The dataset consists of two subsets: a trust network dataset and a dataset for rating items. The statistical information of this dataset is presented in the table I.

Fig. 1 displays the number of trustors and trustees for each user in trust network. As fig. 1 shows, trust matrix is usually sparse since usually most people trust a small number of people. Fig. 2 displays the distribution of the input and output of the graph within the dataset. As fig. 2 shows, the distribution of input and output of the graph of the Ciao dataset has a power low distribution [20] that is common in social networks.

Due to the execution time of the algorithms over the entire Ciao dataset, two subsets are extracted from the Ciao dataset. One includes 1000 nodes with 26799 edges and the other contains 1400 nodes and 39154 edges. Subset extraction has been performed using the Snowball sampling algorithm [24]. The node extraction process using this algorithm is performed

[1] www.Ciao.org

as follows: At first, one node is randomly selected, then 10 nodes among the neighbours of this node are randomly selected. For each of these 10 nodes, the previous steps are repeated again. This will continue until the number of nodes obtained reaches the number of desired nodes. Moreover, the trust values of the Ciao dataset include discrete numbers, and the proposed method requires use of a continuous dataset with trust values within [0-1] interval. Therefore, extracted datasets are converted to the continuous dataset using (4).

$$\begin{cases} T_1 = rand(0,1) \\ T_2 = (1 - T_1) * rand(0,1) \\ T_3 = T_1 + T_2 * \dfrac{\ln(input_{degree(B)})}{\ln(\max\_degree)} \end{cases} \quad (4)$$

In (4), $input\_degree\ (B)$ is the number of nods who trust the node B, max_degree is the maximum degree between all of the nodes within the trust network. In this strategy, trust values are assigned to the edges so that nodes with higher inputs are more likely to be trusted.

TABLE I. STATIC INFORMATION OF CIAO DATASET

	ciao
Users	*7375*
Items	*9974*
Date of the first rating	*30 may 2000*
Date of the last rating	*12 may 2010*
Trust relationships	*111781*
Ratings	*27483*
Density of network	*0.0379*

In order to evaluate accuracy of the proposed method and compare it versus the TidalTrust and MoleTrust algorithms, the Leave-One-Out method that is commonly used in trust researches has been used. In this method, for nodes v_i and v_j, where the direct trust v_i and v_j is available, the indirect trust (propagated trust) v_i to v_j is calculated using the proposed algorithm then according to (5) difference between direct trust and propagated trust will be calculated. The mean absolute error value and the second root mean square error (The difference between direct trust and indirect trust of the two nodes) are considered as the criteria for evaluating accuracy of the algorithm.

$$MAE = \frac{|trust_{ij} - \widetilde{trust}_{ij}|}{N} \quad . \quad RMSE = \sqrt{\frac{trust_{ij} - \widetilde{trust}_{ij}}{N}} \quad (5)$$

In (5) $trust_{ij}$ refers to the value of direct trust between users i and j and $\widetilde{trust}_{ij}$ refers to predicted trust value by proposed algorithm.

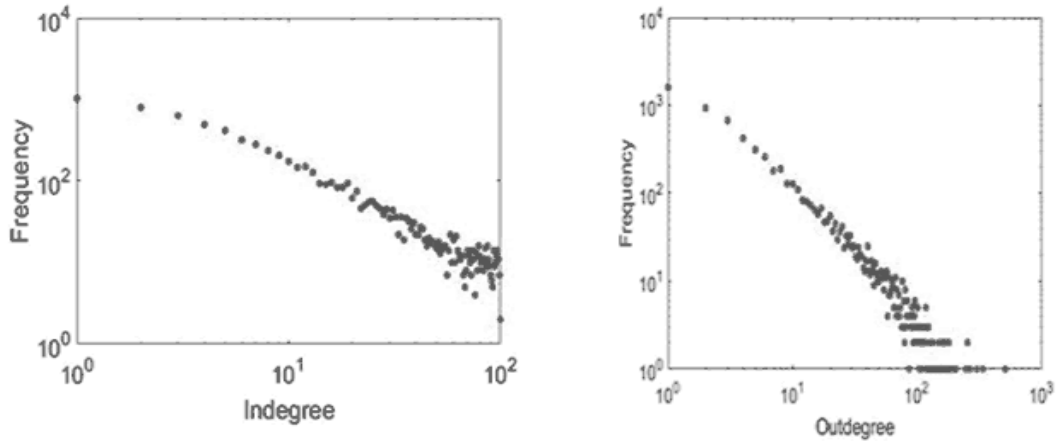

Figure 1. Distribution of number of indegree and outdegree for each user in dataset

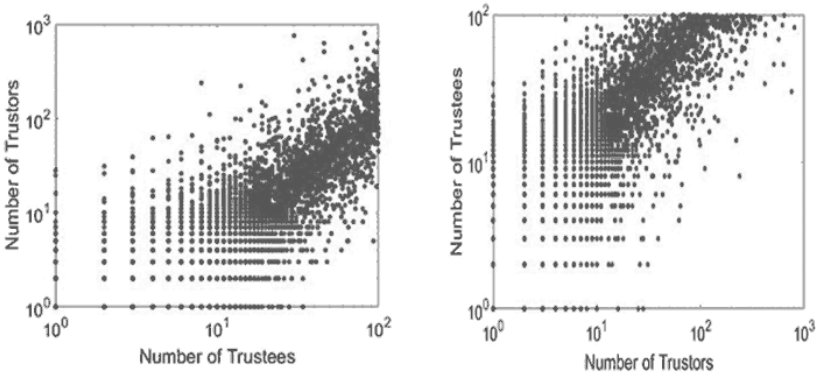

Figure 2. Correlation between number of trustors and trustees

Fig. 3 and fig. 4 show results of this method on the first and second subsets of Ciao dataset.

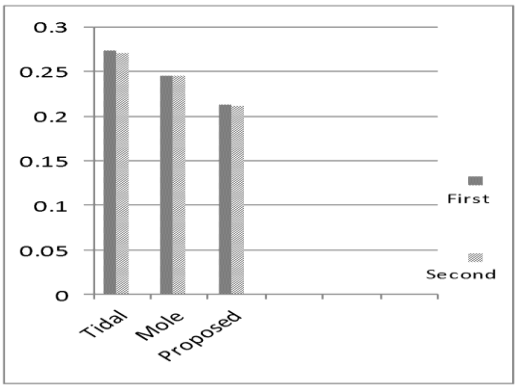

Figure 3. RMSE results on first dataset

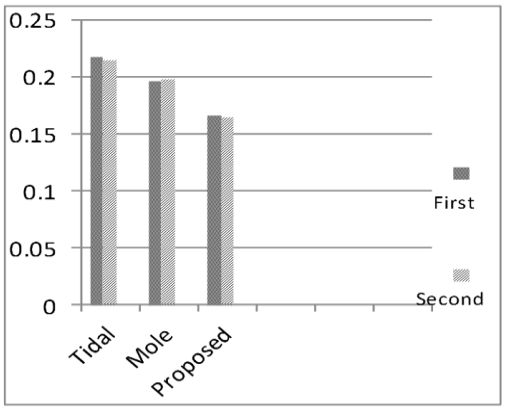

Figure 4. MAE results on second dataset

As these figures show, precision of the proposed method is more than TidalTrust and MoleTrust metrics. Results show that the proposed method outperforms the aforementioned methods applied on Ciao dataset.

V. CONCLUSION AND FUTURE WORK

Considering only the shortest paths and local prediction are main problems for TidalTrust and MoleTrust metrics. Considering only the shortest paths between any two nodes and using only limited local information leads to low accuracy in trust value prediction. It is noteworthy to say that it is possible to use a strategy where all the trusted paths considered. However, this strategy will suffer from high computational complexity and consequently it will not be applicable on large datasets.

High complexity of earlier methodologies increases the required computational power and consequently makes them unsuitable for online use. On the contrary, simplicity of the proposed method provides opportunities for online implications. In this paper, a trust prediction model to calculate trust propagation was proposed. In the future works, intention is to use other ranking methods that improve discovery of the most trusted path. For evaluating the proposed method, a small dataset has been used. In future works, using a much larger dataset to examine the proposed method will help to hammer and harden the method and to improve it. Dataset used in this research only includes trust relationships. The proposed method can be extended to use ranking methods in signed networks if a dataset with distrust information is applied.

VI. REFRENCES

[1] R. T. Mahani and M. Analoui, "Trust prediction in multiplex networks," in *Knowledge-Based Engineering and Innovation (KBEI), 2015 2nd International Conference on*, 2015, pp. 263-268: IEEE.

[2] J. Leskovec, D. Huttenlocher, and J. Kleinberg, "Signed networks in social media," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2010, pp. 1361-1370: ACM.

[3] J. Leskovec, D. Huttenlocher, and J. Kleinberg, "Predicting positive and negative links in online social networks," in *Proceedings of the 19th international conference on World wide web*, 2010, pp. 641-650: ACM.

[4] K.-Y. Chiang, N. Natarajan, A. Tewari, and I. S. Dhillon, "Exploiting longer cycles for link prediction in signed networks," in *Proceedings of the 20th ACM international conference on Information and knowledge management*, 2011, pp. 1157-1162: ACM.

[5] J. Golbeck and J. Hendler, "Inferring binary trust relationships in web-based social networks," *ACM Transactions on Internet Technology (TOIT)*, vol. 6, no. 4, pp. 497-529, 2006.

[6] C.-N. Ziegler and G. Lausen, "Propagation models for trust and distrust in social networks," *Information Systems Frontiers*, vol. 7, no. 4-5, pp. 337-358, 2005.

[7] P. S. Chakraborty and S. Karform, "Designing trust propagation algorithms based on simple multiplicative strategy for social networks," *Procedia Technology*, vol. 6, pp. 534-539, 2012.

[8] Y. A. Kim and H. S. Song, "Strategies for predicting local trust based on trust propagation in social networks," *Knowledge-Based Systems*, vol. 24, no. 8, pp. 1360-1371, 2011.

[9] A. Abdul-Rahman and S. Hailes, "Supporting trust in virtual communities," in *System Sciences, 2000. Proceedings of the 33rd Annual Hawaii International Conference on*, 2000, p. 9 pp. vol. 1: IEEE.

[10] G. Gans *et al.*, "Requirements modeling for organization networks: a (dis) trust-based approach," in *Requirements Engineering, 2001. Proceedings. Fifth IEEE International Symposium on*, 2001, pp. 154-163: IEEE.

[11] L. Mui, M. Mohtashemi, and A. Halberstadt, "A computational model of trust and reputation," in *System Sciences, 2002. HICSS. Proceedings of the 35th Annual Hawaii International Conference on*, 2002, pp. 2431-2439: IEEE.

[12] J. A. Golbeck, "Computing and applying trust in web-based social networks," 2005.

[13] U. Kuter and J. Golbeck, "Sunny: A new algorithm for trust inference in social networks using probabilistic confidence models," in *AAAI*, 2007, vol. 7, pp. 1377-1382.

[14] P. Massa and P. Avesani, "Trust metrics on controversial users: Balancing between tyranny of the majority," *International Journal on Semantic Web and Information Systems (IJSWIS)*, vol. 3, no. 1, pp. 39-64, 2007.

[15] H. Shakeri and A. G. Bafghi, "RTBIMS: Accuracy enhancement in iterative multiplication strategy for computing propagated trust," in *Information Security and Cryptology (ISCISC), 2011 8th International ISC Conference on*, 2011, pp. 9-14: IEEE.

[16] M. Lesani and S. Bagheri, "Applying and inferring fuzzy trust in semantic web social networks," *Canadian Semantic Web*, pp. 23-43, 2006.

[17] S. Schmidt, R. Steele, T. S. Dillon, and E. Chang, "Fuzzy trust evaluation and credibility development in multi-agent systems," *Applied Soft Computing*, vol. 7, no. 2, pp. 492-505, 2007.

[18] S. Nefti, F. Meziane, and K. Kasiran, "A fuzzy trust model for e-commerce," in *E-Commerce Technology, 2005. CEC 2005. Seventh IEEE International Conference on*, 2005, pp. 401-404: IEEE.

[19] M. Lesani and N. Montazeri, "Fuzzy trust aggregation and personalized trust inference in virtual social networks," *Computational Intelligence*, vol. 25, no. 2, pp. 51-83, 2009.

[20] J. Tang, H. Gao, and H. Liu, "mTrust: discerning multi-faceted trust in a connected world," in *Proceedings of the fifth ACM international conference on Web search and data mining*, 2012, pp. 93-102: ACM.

[21] M. Hosseinzadeh Aghdam, M. Analoui, and P. Kabiri, "Modelling trust networks using resistive circuits for trust-aware recommender systems," *Journal of Information Science*, vol. 43, no. 1, pp. 135-144, 2017.

[22] R. Guha, R. Kumar, P. Raghavan, and A. Tomkins, "Propagation of trust and distrust," in *Proceedings of the 13th international conference on World Wide Web*, 2004, pp. 403-412: ACM.

[23] A. Jøsang, E. Gray, and M. Kinateder, "Simplification and analysis of transitive trust networks," *Web Intelligence and Agent Systems: An International Journal*, vol. 4, no. 2, pp. 139-161, 2006.

[24] D. D. Mauá and F. G. Cozman, "Using Social Data to Predict Trust on Web Communities: A Case Study with the Epinions. com Website."