REGULAR PAPER

A propagation trust model in social networks based on the A* algorithm and multi-criteria decision making

Nasrin Hamzelou[1] · Mehrdad Ashtiani[1] · Raha Sadeghi[1]

Abstract
The rapid growth of social networks facilitates the exchange of information whereas malicious behaviors are also steadily increasing in these ecosystems. This results in a challenging situation for individuals to trust other parties. This paper studies the propagation of trust within a chain of trust relations to calculate the trust values of existing users. In this research, an approach for the precise selection of trustworthiness paths as well as the integration of indirect trust values based on the most reliable routes is introduced. The presented approach fuses the ideas from the A* algorithm and multi-criteria decision making approaches using (i.e. TOPSIS method) under fuzzy environments for finding the most reliable path. Moreover, for selecting the most appropriate middle node, a set of criteria such as topological similarity, profile similarity, Dunbar's theorem, local trust, and contextual trust are considered. The evaluation results of the proposed approach demonstrate the propagated trust distance with the different average path lengths while preserving the accuracy of the inferred trust values between each unconnected pair of nodes. The evaluations are performed using the Facebook and Twitter networks having different topologies and the results are compared to the TidalTrust and the MoleTrust algorithms.

✉ Mehrdad Ashtiani
m_ashtiani@iust.ac.ir

Nasrin Hamzelou
nasrinhamzelou@qiau.ac.ir

Raha Sadeghi
r_sadeghi@alumni.iust.ac.ir

[1] Social Computing Laboratory, School of Computer Engineering, Iran University of Science and Technology, Hengam St., Resalat Sq., 16846-13114 Tehran, Iran

1 Introduction

Social networks play a significant role in a diverse set of environments such as trade, biology, economics, marketing, and so forth [1]. With the existence of a massive amount of sensitive information on these networks, the identification and recognition of trustworthy users have turned into a big challenge. This was the initial motivation for the initiation of computational trust modeling research which aims to model the way humans trust each other in social domains and create a mapping to the computational ones.

With the rapid growth of social networks, the number of interactions among anonymous users has increased significantly [2]. On the other hand, within online environments, the ability to assess the trustworthiness level of a user which is going to be the target of interaction is of much importance [3]. Social networks currently act as an enabling ecosystem for electronic commerce in which trust is believed to be one of the vital sources of decision making for end-users [4]. Since there is a sharp rise in the number of people engrossed in using social media, as well as the excessive propagation of information through it, trustworthy social networks whereby individuals can interact with each other without much concern for a potential malicious behavior is of significant importance. Assessing the trust value of two unconnected nodes seems to be a big challenge for the researchers as there are ample problems that need to be addressed. These problems include the correct recognition of trustworthy users for reliable path selection, selecting trustworthy routes based on trustworthy middle nodes, controlling the path lengths as well as integrating the final trust values of the selected nodes when there is no direct relationship between each pair of nodes.

When there is no direct relationship between two parties, their friends or even friend of friends can facilitate the establishment of a connection between them [2]. If there is no direct relationship between two people, they cannot trust each other unless there is at least a path existing in their friends' network. It is a logical consequence of trust propagation based on the transitivity property of trust. This approach can prognosticate the trust value of the target node from the perspective of the source node based on the chain of trust relations between them [5]. Recently, researchers have applied different strategies for finding optimized trustworthy routes to guarantee the prediction of trust values with the highest accuracy. However, despite the importance of the subject and to the best of our knowledge, so far, there has never been a specific focus on finding the most trustworthy path in these networks [2]. Most of the previous studies suffer from a set of shortcomings. Using partial information from the graph structure, considering only the strong short paths which underestimates the potentially remarkable information of longer ones and vice versa, considering trust as a binary value, and finally processing all the possible routes in which even though it can increase the quality of inferred trust values but can considerably increase the time complexity of the method and lead to the exponential complexity of the algorithms [2] are a few of these shortcomings.

Social network users have multiple choices for reaching the trustworthy destination and contradictory factors can potentially make this process more

challenging. This, in turn, creates a multi-criteria decision-making situation and the quality of the routes is determined by how middle nodes are chosen. Therefore, this research takes advantage of a synthesized approach based on the A* algorithm and fuzzy TOPSIS multi-criteria decision making for selecting reliable middle nodes. This can be used for anticipating and assessing the indirect trust value amongst unconnected users within the social network. For the most part, one thing that makes the proposed approach different from the current literature is the simultaneous consideration of multiple factors for ranking the trustworthiness value of middle nodes while taking into account the propagative, dynamic, and contextual properties of trust which in turn leads to more precise results. Compared to the prior studies, we have paid more attention to the structural and interactive information of graphs which is a combination of certain quantitative and qualitative factors. This is performed by taking into account a set of criteria such as profile similarity, topological similarity, interaction rate among users, Dunbar's theorem, as well as other factors such as local trust, contextual trust, and finally, the remaining steps between the middle nodes toward the destination. On the other hand, multi-criteria decision making approaches act as a powerful tool for a better decision making process. These approaches help the decision-maker to evaluate and compare the results of the current alternative solutions [6]. Most of the time, the decision is performed by systematically taking into account three important parameters namely: (1) certainty, (2) risk, and (3) uncertainty. The ultimate goal is to find an optimal solution by considering these parameters [7, 8]. To this aim, fuzzy TOPSIS is used as a method for multi-criteria decision making to rank the trustworthiness level of each middle node between two unconnected nodes. The power of fuzzy approaches especially in information aggregation with contradictory data having an associated uncertainty is demonstrated in a wide range of studies [8, 9]. Then, the A* heuristic algorithm is used for opting the most trustworthy path. It is worth mentioning that the A* algorithm is a well-known algorithm that has been widely used in graph traversal and pathfinding due to its high performance and accuracy. It is a graph search algorithm based on an evaluation function used to sort nodes [10]. Another advantage of the proposed approach is the consideration of both certain and uncertain parameters for finding the proper route and the sequence of middle nodes within the path. In other words, as opposed to other approaches, there is no restriction for choosing only short or long paths. However, we have given more weights to shorter paths in equal certainties, to control the path length. But, we have also given more weights to longer routes with higher certainty compared to shorter but uncertain ones.

The rest of the paper is organized as follows. The related work of this research is reviewed in Sect. 2. Section 3, gives a general introduction of the contextual trust and the trust properties as well as an introduction to the fuzzy TOPSIS multi-criteria decision-making followed by some preliminary description of the A* algorithm. Section 4 describes the proposed approach in more detail. This is followed by a numerical example in Sect. 5. Section 6 contains the simulation results of the proposed algorithm in which a comparison is drawn amongst the proposed approach and other existing models named TidalTrust and MoleTrust

evaluated with two different datasets namely the Facebook and Twitter datasets. Finally, Section 7 concludes the paper.

2 Related work

Trust has always been a vital part of online social networks. Therefore, trust models play a significant role in the context of social trust. Inferring the trust level between two unknown parties is a challenging task. First, it is necessary to find the best possible routes between two target nodes. Second, integrating the trust value of nodes within each path should be performed. Many researchers have worked on the subject of propagation and the combination of trust values in social networks. Generally, trust has been categorized from three different perspectives: (1) trust information gathering, (2) trust value evaluation, and (3) trust value propagation. Various techniques such as statistical methods, machine learning, heuristic methods, and behavior-based approaches have been presented. Machine learning and statistical approaches focus on mathematical trust models whereas heuristic methods aim at creating practical models for implementing core trust systems. Behavioral trust models study the behavior of users in society [11]. Many different sets of criteria have been suggested so far for calculating propagated trust values in social networks. They can be divided into two general categories: (1) the ones for assessing global trust and (2) the ones for assessing local trust. Regarding the global trust calculation, all relations in the trust graph are taken into account. Then, global trust value is assigned to each node which is called reputation. Reputation refers to the trust level of the whole network towards a specific user. This inferential approach for expressing one's belief regarding other users may not be the best and the most accurate technique. On the other hand, local trust metrics takes advantage of structural information defined by the individual's trust network. These approaches define a specific node as the source and compute the trust value among that node and the destination. To put it another way, in terms of local trust calculation, the trust value of each node is computed based on its neighbors' trust and the nodes connected to it, whereas, in the case of global trust evaluation, the complete graph information is required and this type of evaluation could be potentially more complicated. Therefore, the local approach most of the time is the preferred choice amongst researchers [12, 13].

In a well-known research, Golbeck has suggested the TidalTrust algorithm for inferring trust in social networks. In this algorithm, the shorter the path, the more reliable it will be considered if there is no direct relation between the source and destination nodes. This approach looks for the shortest paths from the source node to the target node. If there is more than one similar short path, it will opt the strongest one. The results indicate that the longer the selected path, the weaker the approximation of the trust values will be. It solely works fine in terms of short paths. Even though this approach has low time complexity, the biggest downside of this algorithm is that it is not always efficient, as it overlooks the long routes which may contain valuable information that should not be omitted [2, 5, 12, 14]. Another downside of the approach is the consideration of trust as a binary concept [11]. Also, the

approach is highly influenced by the density of the trust network. Hence, if there is a scattered trust network, finding possible routes could be a tough task [5]. In the proposed approach we have improved these shortcomings by also considering longer paths according to their certainty, considering the interaction rate of middle nodes along with the topological information, and using conflicting and fuzzy values containing uncertainty instead of assuming binary values. Also, because of using TOPSIS as the decision-making mechanism, more optimized trust paths are found in social networks with a scattered structure such as Facebook.

Golbeck and Kuter in another interesting research took advantage of the SUNNY algorithm for anticipating trust values. The main goal of this algorithm is to predict trust between unconnected nodes. Certainty criterion is used in this algorithm for assessing the trust level. It presents an explicit probabilistic definition for trust in social networks and utilizes a probabilistic approach for finding reliable sources. The results show higher performance compared to TidalTrust. This approach is amongst the first researches where certainty criterion has been used for inferring trust values. This algorithm uses a kind of probabilistic sampling for assessing certainty and calculates the certainty based on highly trustworthy information sources. In other words, it implements the trust inferring approach in a more reliable sub-network [2, 15]. A similar set of shortcomings to Tidal trust such as not considering longer paths, taking into account only the topological data, as well as lower accuracy in low-density graph structures, still exist in this work.

Avesani et al. suggested an inferring trust algorithm called MoleTrust. In this algorithm and as the first step, all the shortest paths from the source node to the destination are identified. Then, the trust level of a path is calculated based on the weighted sum of the nodes. As an improvement, the users with trust levels less than 0.6 are not considered. In this algorithm, the importance of the user's beliefs and suggestions is dependent on the certainty of the source node. It consists of two steps. First, the trust graph should be converted to a DAG (directed acyclic graph). Next, this DAG should be traversed while calculating the trust value of each visited node. Although this algorithm is also based on the shortest path, it solely takes into account nodes with trust level less or equal to the propagation trust horizon. Propagation trust horizon/distance is the maximum distance from the source node in which the trust would be propagated to independent of any users. One of the side effects of using this approach is its high time complexity [4, 14, 15]. In the proposed work, many of these shortcomings are addressed. For example, bi-directed graph structures are assumed with different trust values for each direction. Also, the length of the trust paths are controlled by simultaneously considering several criteria in the fuzzy-TOPSIS and A* algorithm, Using the fuzzy-TOPSIS approach will be effective in reducing the time-complexity of the approach as well. This is because many of the branches will be removed from the A* algorithm's open list due to high uncertainty.

Ziegler et al. suggested the Appleseed algorithm for social networks. In this approach a node is selected as a core and energy E is surged into it. Then this energy is freed and is distributed between nodes close to the core. More energy will be transmitted to adjacent neighbors with higher weights. This approach works with partial graph information. In other words, in each step, it only needs the information

for nodes that preserve energy and not the whole graph. This is a well-known algorithm in the trust research that assigns the global trust value to the members of a community. However, the biggest drawback is that global trust is not as accurate as local trust value [3, 12, 16]. A similarity existing between this work and the proposed approach in this paper is the consideration of middle short paths. But the main differences between these two approaches are first, taking into account the longer paths with high certainty when shorter paths have low certainty values and second, the calculation of local trust values for the middle nodes which will result in more accurate calculations and hence the selection of more optimized paths.

Lesani and Montazeri have argued in their research that the inferential trust value for longer trust chains with more trustworthy median nodes could represent more accurate trust information. They take advantage of the fuzzy algorithm and cumulative fuzzy techniques. The comparison of the cumulative fuzzy technique with existing cumulative approaches show the predominance of the fuzzy approach, especially in contradictory information gathering. However, this approach has a high computation cost as well as high time complexity which makes it a non-practical algorithm in real-world social networks [9]. The work of Lesani and Montazeri has a similarity with the proposed approach in this paper where both of these approaches consider qualitative and uncertain values for trust representation instead of a binary viewpoint. But, the proposed approach has the advantage of considering both the shorter and longer paths according to their certainty value as well as taking into account a wide range of criteria including topological similarity, profile similarity, interaction rate, remaining nodes toward the destination, and so on for a more accurate trust calculation process. Also, using fuzzy-TOPSIS will result in the filtration of many paths with lower certainty values starting from the source node which in turn will reduce the time-complexity of the proposed approach.

Ghaemi and Shakeri have proposed a new method to improve the accuracy of trust propagation and calculation along different paths. In this approach, recommended trust amongst nodes is approximated based on belief similarity. Then, the trust level will be calculated by the multiplication of trust value along the trust path. Although this approach could decrease uncertainty along the trust chain, it cannot solve the dispersion of the trust network [2, 17]. In another study, Hassan et al., have evaluated the performance and efficiency of repetitive multiplication in the trust chain whereby trust of a path is calculated by multiplication of trust values within the chain. The trust value for small intervals would be relatively small since the trust value of each node is a real number in the range of [0, 1]. That is why the efficiency of this approach is not considerable. Another weakness of this approach is that it does not work for trust values in other ranges such as (1, 5). It is shown that there is a strong positive linear correlation between direct trust values and trust propagation [3, 18].

Hamdi et al., have introduced the TISoN algorithm for inferring trust in online social networks. Additionally, they have suggested a trust search algorithm called TPS and a trust inferring method named TIM. TPS algorithm identifies high priority adjacent nodes based on direct trust and trustworthy path while controlling the path length. In this algorithm, path selection is performed based on two criteria: (1) the maximum allowed path length and (2) the minimum threshold level for trust edges

within the path. Afterward, inferential trust is calculated based on the most trustworthy path, the average of trusted edges along the path, the variance of path edges, and the weight of the path according to its length. TISoN computations have high exponential complexity for massive data in the real world and the dynamic nature of trust is not considered as well [19]. This work is similar to the proposed approach concerning the consideration of longer paths as well as inferring trust values between pairs of unconnected nodes. But, due to the filtration mechanism of fuzzy-TOPSIS, the time-complexity of the proposed approach would be better. Also, the consideration of a wide range of parameters including the interaction rate of users instead of just considering the static topological info will result in more accurate calculations.

Ghavipour et al., have proposed a local inference and propagation algorithm called DLatrust based on distributed learning automata for finding the most trustworthy paths and calculating the final trust value of a chain of nodes in a trust path. They studied three important areas namely (1) finding trustworthy paths, (2) trust path length calculation, and (3) an integration method for calculated trust values**.** The proposed method allocates a learning automaton to each node in the trust network. A new cumulative function based on standard cooperation filtering for integrating trust values generated from different paths is also suggested. Experimental results based on real network data illustrate that the suggested algorithm can effectively be used for trust inference compared to other trust propagation algorithms [15]. The proposed approach in this paper also focuses on these three aspects but the consideration of a larger and potentially conflicting set of criteria would result in more accurate trust calculations and thus more optimized paths.

A significant portion of the studied approaches have utilized partial trust information and only a few factors for trust evaluation. Table 1 summarizes some of the well-known trust propagation methods currently existing in the social networks field.

3 Background knowledge

Trust is a fundamental component of decision making. It has a significant role in daily interactions of people from interpersonal communications to vital decision making in e-commerce, as well as development, preserving, and keeping relations in society [20]. Trust is defined as a mental state of an individual in psychology [21] and is a dynamic, contextual, propagative, asymmetric, self-reinforcing concept [16, 22]. Approaches for computing and evaluation of trust can be divided into three categories: (1) topological trust models, (2) trust models based on users' interactions, and (3) hybrid trust models which is a combination of two aforementioned models [16]. One of the apparent characteristics of social networks is their dynamic nature. As time passes by, new relations are established [16, 21]. On the other hand, users in a social network do not have the entire information from their surrounding environment. Therefore, deciding on trusting/not trusting other users has turned into a big challenge. For the ones who have never been in direct contact, trust is inferred based on its propagative attribute. Such propagation will occur on at least one possible path which connects the source

Table 1 The comparison of well-known methods in trust propagation models

Algorithm name	Data	Characteristics	Drawbacks
Tidal Trust	Web social networks	1. Trust of the shortest paths 2. The topological graph information 3. Performs a modified breadth-first search using local trust	1. Longer but possibly valuable paths might be overlooked 2. It is highly influenced by the density of the trust web 3. Trust is defined as a binary concept
MoleTrust	Epinions.com	1. Using local trust values 2. The importance of a suggestion or belief is dependent on the source node	1. Considering only the shortest paths 2. Users with trust values less than 0.6 are removed 3. High time complexity
Appleseed	Advogato community web site	1. Defining a core node, surging energy to it, and distributing it to adjacent neighbors 2. More energy to closer nodes with higher weights 3. Using global trust value	1. Utilizing global trust value for each peer which is not as accurate as local trust 2. Partial graph information is required 3. The graph is not weighted
EigenTrust	Peer-to-peer sharing files	Global trust calculation of each peer based on the upload date of the file	Utilizing global trust value for each peer which is not as accurate as local trust Specifically designed for peer-to-peer networks
Iterative multiplication Algorithms for trust propagation	Web social networks	1. Propagated trust is equal to the average trust values of all paths 2. Trust of a path is calculated by multiplication of trust values within the chain	1. The more the hops between the source and the destination nodes, the less the propagated trust value 2. If the trust level of one of the nodes in the path is low, even though all other nodes highly trust each other, the final trust value will be negligible
Fuzzy trust aggregation	Virtual marketplaces on the Web	Employing linguistic terms for trust specification. Corresponding fuzzy sets are defined for trust linguistic terms and a fuzzy trust aggregation method is also proposed	This approach has a high computation cost as well as high time complexity which makes it a non-practical algorithm in real-world social networks

and destination nodes [15]. Another advantage of this propagative nature is that users can find new friends and establish new connections [16].

It is argued in some studies that trust is a pseudo-transitive subject [22]. It is not completely transitive, and its value would decrease in a chain of relations [12, 16]. According to Fig. 1, if A trusts B and B trusts C, it can be concluded that A would trust node C up to a threshold based on its trust toward node B. This is called trust inference. Apparently, as there is no direct connection between A and C, A uses the trust information of B toward the node C. However, when node A is directly connected to B, there is no need for any intermediate information to decide about the trustworthiness level of B. This is called direct trust, while the former situation is named referral trust [23]. In massive social networks, due to the existence of a huge volume of users, as well as the fact that not all of them are in close connection with each other, the propagative and referral nature of trust is a facilitating factor for computing the trust value of unconnected nodes.

In Fig. 1, if node B is the only middle node in a path between node A and C, and node A trusts node B, then the trust level of A toward C can be easily inferred. However, when there are multiple chains of routes between nodes A and C, which in turn will increase the number of different middle nodes for deciding about trusting/distrusting the target node, decision making can get complicated. In this case, different parameters can have an impact on the final trust value. Some of these parameters are: (1) trust chain length, (2) local trust value, (3) context-based trust value of middle users, (4) topological structure, (5) interaction rate among users, and (6) different strategies for integrating middle values. This final attribute refers to the trust aggregation mechanism which is one of the core components of every computational trust model [16]. As mentioned earlier, investigating all the possible routes between two nodes would lead to exponential time complexity [2] which is not preferable. Therefore, selecting one or a few number of trustworthy paths with the least number of middle nodes seems to be an indispensable problem that needs to be addressed.

3.1 Fuzzy TOPSIS multi-criteria decision making

In the proposed approach we have used the fuzzy TOPSIS method one of the well-known techniques in the multi-criteria decision-making domain for facilitating the trust decision making [24]. Such usage is beneficial to the approach as a result of the existence of multiple routes and various sets of criteria influencing the final decision. Generally, multi-criteria decision-making models are used for selecting the best alternative amongst all the existing ones. In these techniques, all available options are ranked and categorized. The fuzzy TOPSIS method is one of the most efficient and well-known methods available for this aim. The general idea behind this technique is that considering all possible solutions, two

Fig. 1 Direct and referral trust [23]
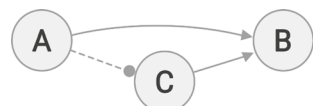
hypothetical options are selected. One of them is the set of all the best values in the decision matrix which is called the positive ideal solution. On the other hand, we would have a negative ideal solution, comprising all the negative or worse options. Each criterion could have a positive or negative influence. Also, their measurement could be different. In this approach, options are categorized based on their similarity, so that the more an option is closer to the ideal solution, the more its rank would be [24, 25]. This decision-making technique is based on a sound mathematical framework and is comprised of 7 steps:

1. Decision matrix creation and assigning weights to each criterion.
2. Normalizing the matrix.
3. Generating a weighted fuzzy matrix.
4. Determining positive ideal and negative ideal fuzzy solutions.
5. Estimating how each option is far from the positive/negative ideal solutions.
6. Determining the similarity coefficient for each node and assessing the similarity index.
7. Finally, ranking options are based on the similarity ratio [24–26].

3.2 The A* algorithm

In graph theory, finding the shortest path problem deals with finding a path between two nodes so that the sum of the weights of the paths between these nodes is minimized. For example, consider finding the quickest route to be transported from a point on a map to another. In this situation, edges symbolize the path between two nodes which are labeled based on the time required to pass through the path. Consider a weighted graph consisting of a set of nodes named *V*, as well as *E* edges and a weighting function F: E→R. Besides, assume two unconnected nodes such as *C* and *A* from the set V. The goal is to find a path like *P* from the node *A* to node *C* as the shortest path amongst all possible ones. The A* algorithm finds the shortest path by assigning a score to each square which is called the scoring path. In Eq. 1, $g(n)$ is the real cost of reaching node *N* from the first node. In other words, this is the cost of finding the optimized path. The $h(n)$ is the aggregated cost of reaching the target node through node *N* which requires heuristic information from various situations. How *G* and *H* are selected can be changed based on different circumstances.

$$f(n) = g(n) + h(n) \tag{1}$$

Two lists are required in the A* algorithm. One, for listing all possible squares that are presumed to be the shortest path, which is called the open list. While the closed list is used for listing all the squares which are not required to be evaluated

anymore [10]. Indeed, this algorithm is an extension of the Dijkstra's algorithm [27]. This algorithm takes advantage of the best-first search and finds the shortest path between two given nodes based on middle nodes. The A* algorithm is represented as Algorithm 1 below:

```
Algorithm 1: The A* path finding algorithm
─────────────────────────────────────────────────
Input:
  (1) A graph G(V, E) With source node (Node_start) and goal node (Node_goal).
Output:
  (1) The least cost path from the start node to end
─────────────────────────────────────────────────
     PROCEDURE:
1:   | Put Node_start in the OPEN list with f(Node_start) = h(Node_start)
2:   | WHILE the OPEN list is not empty
3:   |     Take from the open list the node Node_current with the lowest
4:   |     f(Node_current) = g(Node_current) + h(Node_current)
5:   |     IF Node_current == Node_goal THEN we have found the solution
6:   |     Break;
7:   |     Generate each state Node_successor which comes after Node_current
8:   |     FOR each Node_successor of Node_current
9:   |        Set successor_current_cost =
10:  |           g(Node_current) + w(Node_current, Node_successor)
11:  |         IF Node_successor is in the OPEN list
12:  |          IF g(Node_successor) ≤ successor_current_cost THEN
13:  |                Continue to line 24;
14:  |         ELSE IF Node_successor is in the CLOSED list
15:  |            IF g(Node_successor) ≤ successor_current_cost
16:  |                Continue to line 24;
17:  |          Move node_successor from the CLOSED list to the OPEN list
18:  |          ELSE
19:  |            Add Node_successor to the OPEN list
20:  |            Set h(Node_successor) to be the heuristic distance to node_goal
21:  |
22:  |        Set g(Node_successor) = successor_current_cost
23:  |        Set the parent of the Node_successor to Node_current
24:  |     ENDFOR
25:  |  Add Node_current to the CLOSED list
26:  | ENDWHILE
27:  | IF (Node_current != Node_goal) exit with error! (the OPEN list is empty)
28:  | END-PROCEDURE
```

4 Proposed approach

In social networks, when a source node is not directly connected to the destination node, one of the existing challenges is to calculate the trust between them and also find the most trustable path connecting them. One of the shortcomings of most of the existing approaches is that the accuracy of trust calculation significantly decreases when the path lengths become larger. Because of this issue, many of the existing approaches limit their calculation only to shorter paths and completely ignore the longer ones which will reduce their accuracy to a great extent. In the few cases where the trust information of all possible paths is considered, the time-complexity of the approach becomes a major issue.

To improve upon these shortcomings, in the proposed approach, the utility function of each node (representing the trust value) is calculated based on a fuzzy-TOPSIS decision-making process. This will allow the model to consider a wide range of qualitative, uncertain, and potentially conflicting criteria for trust calculation (7 parameters are considered in this paper). Using fuzzy-TOPSIS helps in both reaching more accurate trust calculations for middle-nodes as well as acts as a pre-filtration mechanism for the next step where nodes with lower trust and certainty values will be discarded resulting in a lower time-complexity compared to approaches where nodes in all the existing paths are considered. Also, the approach uses an extended version of the A* algorithm which results in the selection of a combination of shorter and longer paths but with high certainty values. This in turn will increase the accuracy of the proposed approach compared to the approaches where the trust information of only the shorter paths is taken into account.

By considering the social network as a graph, each user can be considered as a node in the graph, and all nodes existing between two unconnected source and target nodes are middle nodes. In this case, the value of $g(n)$ and $h(n)$ could be calculated based on a set of criteria including topological similarity, profile similarity, Dunbar's theorem, interaction rates between the source node and middle node, contextual and local characteristics of trust, and finally the remaining steps from a middle node to the target. Final trust as a utility function is used for ranking each option. Middle nodes with the largest final trust values in each step are added to the open list. Respectively, the ones with the lowest final trust values will be inserted into the closed list. In the end, similar to the A* algorithm, traversing edges with large final trust (i.e. routes with high certainty from the source node to the destination) is the ultimate goal of the proposed approach. Figure 2, represents how the A* algorithm is extended in this approach. The direction of the edges determines the trust direction. For example, if node A trusts node B, the edge direction will be represented as A→B. On the other hand, we know that due to the asymmetric nature of trust, when A trusts B, it is not mandatory that B trusts A as well.
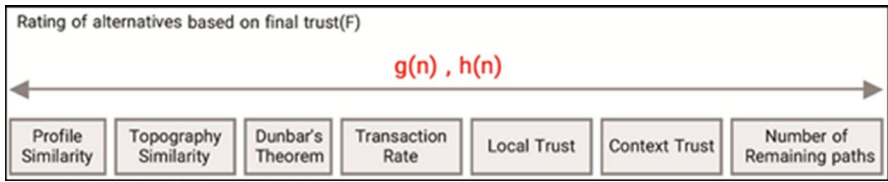
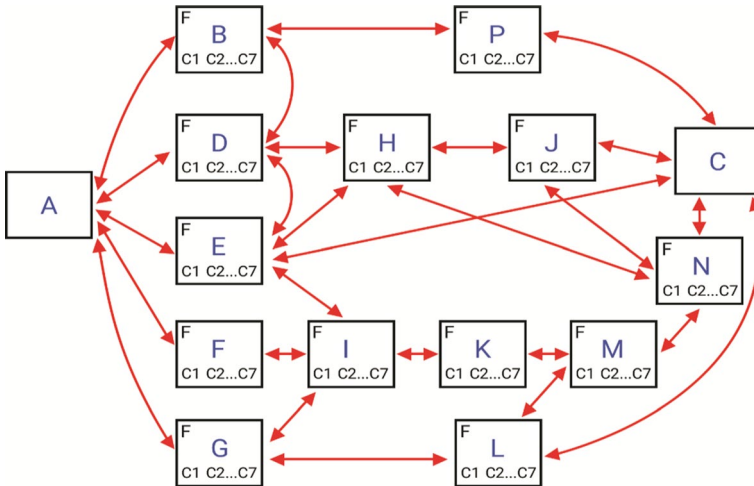Fig. 2 An extension of the A* algorithm for ranking middle nodes

Fig. 3 C1 to C7 criteria and utility function (F) for assessing A's trust toward C

When both parties trust each other (i.e. A↔B), the trust value could be different depending on their perspective and previous interactions [12, 16]. Assume that the trust relationships are established as represented in Fig. 3. The ultimate goal is to find the most trustworthy path for inferring the trustworthiness level of C from the viewpoint of node A. For this purpose, it is first required to find the reliable middle nodes connecting these two nodes. According to Fig. 4, all these nodes should be ranked based on a set of criteria. The following steps will be performed by the proposed approach:

1. For each unconnected pair of nodes, all the reliable middle nodes within the trust chain should be identified. The nodes' trustworthiness value should be calculated
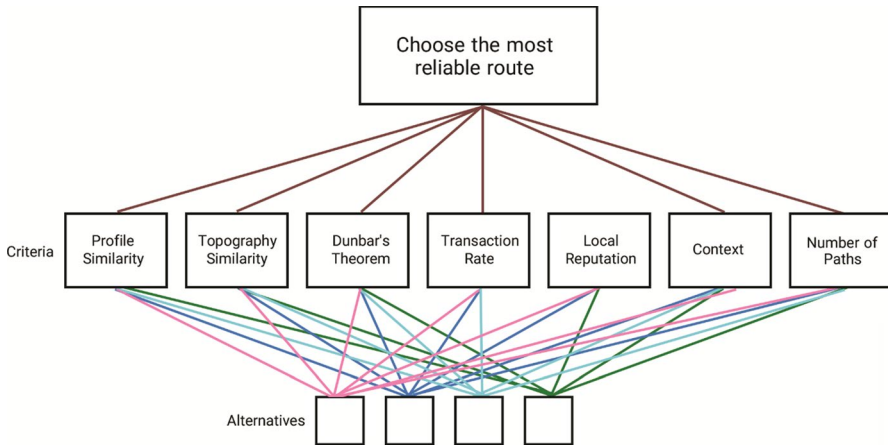
Fig. 4 Users' criteria for decision making in a social network

based on the combination of different factors such as profile similarity, topological similarity, and interaction rate among users [16].

2. Bearing in mind the contextual characteristic of trust [16, 28], trustworthy users should be selected in the trust chain and based on the specified context.

3. Since trust highly depends on an individual's mental status [16, 21], for preventing the unintended biased negative feedback, local trust value between each unconnected pair of nodes within the trust chain should be calculated and be considered while deciding on the middle users for finding the trustworthy path.

4. Dunbar says people have the ability to keep their connection with at most 150 different parties and the higher the number of connections the lower the quality of the relationships [23, 29, 30]. Therefore, another efficient factor for ranking middle nodes is giving more weights to the nodes that have fewer connections.

5. As the last deciding factor in the decision-making process, the remaining steps toward the destination is calculated. This is to select the shortest possible path amongst the available trusted nodes.

6. The final trust value would be calculated by utilizing the fuzzy TOPIS method.

7. Finally, for traversing the graph between the source and the destination, the nodes with the highest final trust values will be chosen based on the A* algorithm.

Algorithm 2: A*-Topsis trust path finding algorithm

Input:
(1) A graph G(V,E) With source node (Node$_{start}$) and goal node (Node$_{goal}$)
Output:
(1) The superlative FinalTrust path from the start node to end

---

1: **FUNCTION** Topsis
2:   **PASS IN** Node
3:   **FORM** the neighborhood-attribute matrix $X := x_{ij}; i \in alternatives, j \in attributes$
4:   **CALCULATE** normalized rating n_{ij}
5:   **CALCULATE** the weighed alternative matrix $v_{ij} := W_j . n_{ij}$
6:   **IDENTIFY** the ideal alternative $S^+ := \{v_1^+, v_2^+, \ldots, v_j^+, \ldots v_q^+\} :=$
    $\begin{cases} \max v_{ij} \mid j \in emission\ attributes \\ \max v_{ij} \mid j \in cost\ attributes \end{cases}$
7:   **IDENTIFY** the nadir alternative $S^- := \{v_1^-, v_2^-, \ldots, v_j^-, \ldots v_{mq}^-\} :=$
    $\begin{cases} \min v_{ij} \mid j \in emission\ attributes \\ \min v_{ij} \mid j \in cost\ attributes \end{cases}$
8:   **FOR** each alternative i
9:     **CALCULATE** $D^+ := distance\ between\ alternative\ i\ and\ S^+$
10:     **CALCULATE** $D^- := distance\ between\ alternative\ i\ and\ S^-$
11:     **CALCULATE** ratio $CC_i := \frac{D_i^-}{D_i^- + D_i^+}$
12:   **ENDFOR**
13: **OUTPUT** the maximum ratio
14: **ENDFUNCTION**

PROCEDURE:
1:   Put Node$_{start}$ in the OPEN list with f(Node$_{start}$)
2:   **WHILE** the OPEN list is not empty
3:     **CALCULATE** finalTrust using Topsis $f(Node_{current}) := Topsis(Node_{current})$
4:     **EXTRACT** from the open list the node Node$_{current}$ with the highest $f(Node_{current})$
5:     **IF** Node$_{current}$= Node$_{goal}$ **THEN** we have found the solution
6:     **Break**;
7:     Generate each state Node$_{successor}$ which comes after Node$_{current}$
8:       Set successor_current_cost := Topsis(Node$_{start}$)
9:     **FOR** each Node$_{successor}$ of Node$_{current}$
10:       **IF** $f(Node_{current}) \geq limit_{finaltrust}$ then
11:         **IF** Topsis(Node$_{successor}$) >= successor_current_cost **THEN**
12:           **IF** Node$_{successor}$ is in the CLOSED list
13:           **Move** Node$_{successor}$ from the CLOSED list to the OPEN list
14:         **ENDIF**
14:         **ELSE**
15:         Add Node$_{successor}$ to the OPEN list
16:         **Continue** to line 23;
17:       **ENDIF**
18:       **Else**
19:         Add Node$_{successor}$ to the closed list
20:         Set successor_current_cost:=
21:         Topsis(Node$_{current}$) + w(Node$_{current}$, Node$_{successor}$)
22:         **Set** the parent of the Node$_{successor}$ to Node$_{current}$
23:     **ENDFOR**
24:     Add Node$_{current}$ to the CLOSED list
25:   **ENDWHILE**
26: **IF** (Node$_{current}$!= Node$_{goal}$) exit with error! (the OPEN list is empty)
27: **END-PROCEDURE**

Because in each traversed path a large set of viable candidates may exist and a large set of criteria should be analyzed, the decision-making process is not a trivial task. In addition, a combination of concrete and fuzzy numbers, as well as positive and negative factors with different measurement units may exist in various contexts. Forming the decision matrix is considered to be the first step in all multi-criteria decision making approaches. This is a matrix for the evaluation of multiple options based on a set of established criteria. In other words, in this matrix, each option is ranked based on different factors. A decision matrix is a matrix for appraising *m* options based on *n* criteria. So, as an initial step, a decision matrix from the perspective of node *A* in the sample social network's graph presented in Fig. 3 should be created. This matrix will be used for ranking options *B, D, E, F*, and *G* based on the criteria represented in Fig. 4.

What is crucially important during this matrix formation is the weight or type (i.e. positive/negative) of each criterion. Also, if there were some uncertain values amongst the ones that form the decision matrix, an appropriate range and definition should be used so that the fuzzy TOPSIS technique could be applied. The overall algorithm of the proposed approach is provided as Algorithm 2. In the following, each criterion will be discussed in more detail.

4.1 Profile similarity criterion

The first criterion we have considered for the trust calculation of middle nodes in the paths existing between source and destination nodes, is the profile similarity. Researches proved that those who trust each other are more likely to be similar [1]. In social networks, users can find their classmates, colleagues, old friends, and so forth based on their profile similarity. Figure 3 is a slice of a massive social network. Suppose that the trust values labeled on each edge and the ranking of each middle node (B, D, F, E, G) are computed solely based on the profile similarity of users. Now if another set of criteria such as the topological similarity, or the number of remaining nodes from a middle node toward the target node would be considered, each edge's trust value and therefore, the node *A*'s decision for choosing the middle nodes would differ entirely.

According to Fig. 3, *A* should find its most trustworthy adjacent neighbors next to the trustworthy path to reach node *C* (i.e. target). The main idea behind calculating trust based on similarity lies in the fact that trust is built based on social similarities [31]. People tend to create groups based on their similar interests [32]. This is the main reason that trust based on profile similarity is chosen as an initial criterion of the fuzzy TOPSIS technique in this paper. However, people reveal their interests at various levels depending on the activity type in social networks. For example, they show their movie or music tastes much more than their political orientation. Here, Eq. 2 shows how more weights are assigned to more specific interests based on profile similarity definition of trust.

$$Criteria(\text{Pr}\,ofile - similarity)_1 = \begin{cases} Trust_{PS}\left(SourceNode_A, MiddleNode_B\right) = \frac{\sum_{i \in I_{AB}} w_{(i)}}{\sqrt{I_A} * \sqrt{I_B}} \\ 0 < Trust_{PS}(A, B) \le 1 \end{cases}$$

(2)

where $w_{(i)}$ is equal to $\frac{1}{\log(N)}$ and N, is the number of users who are interested in i. In other words, the more people are interested in an activity, the less its weight would be. While I_A and I_B are the number of interests of user A and B respectively, I_{AB} is the number of similar interests. If $I_B=0$ or $I_A=0$, this means that no interest is cited in their profile. It is worth mentioning that this criterion will be considered as a positive criterion in the source node A 's decision matrix.

4.2 Topological similarity criterion

Since users may not reveal all their interests thoroughly and explicitly, estimating trust based on interest's similarity is not always applicable. Hence, as it is shown in Eq. 3, another method for approximating trust based on the topological similarity of node A with each middle node B, D, E, F, and G using the Jaccard index [33] is suggested. The calculated value using this method is added to the decision matrix as a second criterion (C_{TS}).

$$Criteria(Topoghraphy - similarity)_2 = \begin{cases} trust_{TS}\left(SourceNode_A, MiddleNode_B\right) = \frac{|f_A \cap f_B|}{|f_A \cup f_B|} \\ 0 < Trust_{TS} \le 1 \end{cases}$$

(3)

Here, f_A and f_B refer to the number of friends and acquaintances of A and B respectively.

4.3 Users' interaction rate criterion

Although studies have shown that homophily is formed amongst people who trust each other, this does not mean that homophily would necessarily lead to trust [34]. That's why the interaction rate amongst users is added as a third criterion to the two aforementioned criteria in the decision matrix according to Eq. 4.

$$Criteria(TranscationRate)_3 = \frac{\sum_{N \in Activityset(SourceNode_A)} w_N * N\left(SourceNode_A, MiddleNode_B\right)}{\sum_{N \in Activityset(SourceNode_A)} W_N}$$

(4)

where $N\left(user_A, MiddleNode\right)$ is the number of interactions between A and the *MiddleNode* like B, and W_N is the weight of each transaction between a directed pair of nodes like *AF, AG,* and so on. For example, transactions such as liking a post, sending a message, commenting on a post, each can have different weights in a social network.

Fig. 5 The impact of the *B* and *D*'s local trust on how *A* would find node *C* trustworthy
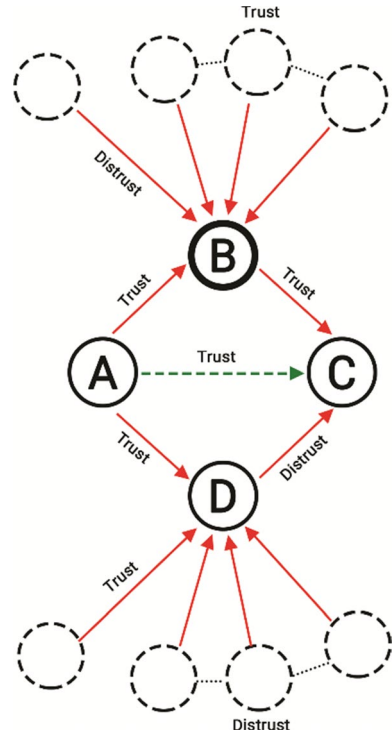
4.4 Local trust criterion

Figure 5 represents a small trust network. As there is no direct connection between node *A* as the source and node *C* as the target node, *A* needs the trust value of middle nodes *B* and *D* for inferring the trust value of the target node. Node *A* trusts both nodes *B* and *D* equally. On the other hand, *B* trusts *C*, so it is capable of transferring its trust value to *A*. However, node *D* does not trust *C*, and so it would send negative feedback regarding *C* to *A* which creates a complex situation. In this case, we check the local trust level of recommenders and see how they are trustable from the perspective of their neighbors based on previous interactions they had with each other. Suppose node *D* has been recognized as an untrustworthy node in its local trust group while *B* has developed a good reputation in its locality. In this case, the information provided by node *B* is more reliable and authentic. Now, if the trusted network gets expanded and the number of middle nodes between *A* and *C* increases, the path crossing node *D* will not be a proper option for *A* to decide whether to trust *C* or not. Therefore, in the suggested approach, each middle user will be ranked based on its reputation and the history of its former performance in the social network.

Some algorithms such as TidalTrust use binary values [0, 1] for expressing the trust level of users towards each other. However, in complex social networks as a

Table 2 Triangular 5-level fuzzy numbers

Linguistic variable	Attribute grade	Corresponding triangular fuzzy number
Very poor (VL)	1	(0, 1, 3)
Poor (L)	2	(1, 3, 5)
Fair (M)	3	(3, 5, 7)
Good (H)	4	(5, 7, 9)
Very Good (VH)	5	(7, 9, 10)

result of an inadequate amount of certain and accurate information, this form of trust definition is not suggested [11]. By this definition, there is no way to distinguish completely untrustworthy nodes from partially untrustworthy ones. Fuzzy logic is a multipurpose logic. Therefore, the uncertainty and ambiguity in different subjects define the necessity of applying fuzzy techniques for modeling qualitative aspects of the problem. Having a set A, the membership function of x is defined according to Eq. 5 [35, 36]:

$$\mu_A(X) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases} \tag{5}$$

If x does not belong to A, then $\mu_A(X) = 0$. If x partially exists in A, then $\mu_A(X) : X \to [0, 1]$ and x is a fuzzy member. In this case, a membership function is a real number and an object could partially be an instance of a set.

In fuzzy logic, the membership of set members is defined based on the $U(X)$ function where X represents a specific member and U is a fuzzy function that determines the membership degree of X in the set and its value is a real number between 0 to 1. A fuzzy number can be defined either in the triangular or trapezoidal form [25, 36]. In triangular form, the desired number is represented as $M = (a, b, c)$, where a, b, and c parameters are referring to the minimum, most likely, and the maximum amount of the target value. This value can be changed within the scope of *(a, c)*. In the proposed approach, the local trustworthiness level of each middle node in its trust group is considered to be a fuzzy set. For defining the membership degree of each middle node, 5-level triangular fuzzy numbers according to Table 2 are used.

Equivalent fuzzy numbers for each linguistic variable defined in Table 2 are formulated by Eq. 6 and it is set as the fourth criterion in the local trust decision matrix. Therefore, the local trust value of the middle nodes *B, D, E, F,* and *G* should be defined based on this equation. For calculating this value, its interaction rate with its connected nodes in its trust group, as well as its interaction rates in previous time windows should be used.

$$Criteria(LocalTrust)_4 = \begin{cases} TransctionIntensity^a_{MiddleNode_B} = \sum_{i=1}^{a} \left(\frac{N_{TR_i}}{N_{TR_i}-1} \right) > 1 \\ TransctionIntensity^a_{MiddleNode_B} = \sum_{i=1}^{a} \left(\frac{N_{TR_i}}{N_{TR_i}-1} \right) = 1 \\ TransctionIntensity^a_{MiddleNode_B} = \sum_{i=1}^{a} \left(\frac{N_{TR_i}}{N_{TR_i}-1} \right) < 1 \end{cases} \quad (6)$$

where x_i is the number of recent interactions, and x_{i-1} is the number of interactions that occurred between a middle node like B with its connected nodes in a previous time window. Trust value can be enhanced or decreased as time passes by as a result of new experiences. Generally, new experiences are much more important than the old ones and a good trust model should pay less attention to old interactions [16]. According to Eq. 6, if the interaction rate is less than 1, *poor* or *very poor* linguistic variables are used. If there is no change in interaction rates from the previous time window, *fair* is a good choice as a linguistic variable and if this rate is more than 1, then we could use either *good* or *very good*. Then, based on Table 2, triangular fuzzy numbers equivalent to each linguistic variable can be substituted.

4.5 Context-based trust criterion

As mentioned earlier, one of the important features of trust is its contextual property [16]. In social networks, individuals within a trust chain could be trustworthy in one or multiple contexts, but this does not mean that they are trustworthy in all domains. This could impact the trust propagation in the trust chain. As trust is highly dependent on context, it could result in uncertainty in some situations and so it is crucial to consider the desired context as a significant criterion when calculating the trust level. Here, another 5-level triangular fuzzy number based on Table 3 is used for demonstrating how a middle user is evaluated as an expert in a specific context.

Table 3 5-level fuzzy numbers for evaluating the expertise of a node in a given context

Linguistic variable	Attribute grades	Corresponding triangular fuzzy number
Very Low (VL)	1	(0, 0, 0.25)
Low (L)	2	(0, 0.25, 0.5)
Medium (M)	3	(0.25, 0.5, 0.75)
High (H)	4	(0.5, 0.75, 1)
Very High (VH)	5	(0.75, 1, 1)

Fig. 6 A large number of *B*'s friends in a sample trust graph
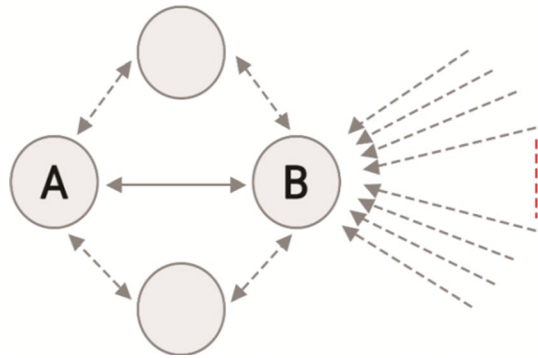
4.6 Dunbar criterion

Based on the Dunbar theorem, individuals cannot keep a large set of social relations in mind and their cognitive capacity will limit the number of social connections they have at each moment in time. Each individual has the intrinsic capability of categorizing her/his social relations. Dunbar stated that people can manage and preserve between 100 to 250 relations with an average of 150 [23, 29, 30]. The less the number of friends of a node, the more selective they would be based on Dunbar's theorem. The reason we have considered this criterion in the decision matrix is to consider the maximum number of effective relations for the middle nodes and give more weight to nodes with fewer but more focused and effective relations as opposed to nodes with large but weak and ineffective connections. This criterion is formulated based on Eq. 7. According to Fig. 6, the larger the degree of each node, the smaller score it would have based on this Equation. This is considered as the sixth criterion and as a positive ideal one in the decision matrix.

$$Criteria(Dunbar-theorem_{(SourceNode_A, MiddleNode_B)})_6 = \frac{|f_A \cap f_B|}{\max\{k_A, k_B\}} \qquad (7)$$

In the above Equation, the numerator shows the number of mutual friends, while the denominator is the maximum number of friends of *A* and *B*. The more the number of friends, the less the final score will be.

4.7 Remained path calculation criterion

Golbeck and many other scientists have argued that short paths can provide better and more accurate trust information between the source node and the destination one [2, 3, 15]. In this paper, according to their argument, we have tried to control the path length and have an estimate on the remaining hops. So, the remaining hops are considered as a negative ideal factor along with other mentioned criteria in the decision matrix. It is considered a negative criterion as it will result in a higher cost in longer paths.

Table 4 Triangular fuzzy numbers for the criteria's importance

Linguistic variable	Attribute grade	Corresponding triangular fuzzy number	Assumed attribute grade for the criteria
Very Low (VL)	1	(0, 0.1, 0.3)	C_{DT}
Low (L)	2	(0.1, 0.3, 0.5)	C_{LT}
Medium (M)	3	(0.3, 0.5, 0.7)	C_{PS}, C_{TS}
High (H)	4	(0.5, 0.7, 0.9)	C_{CT}, C_{RP}
Very High (VH)	5	(0.7, 0.9, 1)	C_{TR}

Therefore, for each middle node, the remaining path is another factor in the decision matrix. For example, in Fig. 3, the remaining path from node *B* to *C* is 2 hops, while it is 3, 1, 5, and finally 2 hops in the case of nodes *D, E, F,* and *G* respectively. It is worth mentioning that in assessing the remaining path, the nodes counted in previous rounds are not taken into account again. This is the same parameter used in Eq. 1 and by the A* algorithm.

4.8 Calculating the final trust value based on the fuzzy TOPSIS

The first step in fuzzy TOPSIS is forming the decision matrix. Then, steps such as finding all possible alternatives, defining a set of criteria for a special source node such as *A*, determining positivity or negativity of each criterion as well as representing it as concrete or qualitative comes next. Each element of the decision matrix for node *A* is represented by X_{ij}. For instance, x_{11} is the evaluation of node *B* based on the C_1 criterion which can be either a concrete number or a qualitative one.

$$((SourceNode_A) = \begin{cases} Criteria(MiddleNodes) \\ MN_B \\ MN_D \\ MN_E \\ MN_F \\ MN_G \end{cases} \begin{bmatrix} c_{PS} & c_{TP} & c_{TR} & c_{LT} & c_{CT} & c_{DT} & c_{RP} \\ x_{11} & x_{12} & x_{13} & x_{14} & x_{15} & x_{16} & x_{17} \\ x_{21} & x_{22} & x_{23} & x_{24} & x_{25} & x_{26} & x_{27} \\ x_{31} & x_{32} & x_{33} & x_{34} & x_{35} & x_{36} & x_{37} \\ x_{41} & x_{42} & x_{43} & x_{44} & x_{45} & x_{46} & x_{47} \\ x_{51} & x_{52} & x_{53} & x_{54} & x_{55} & x_{56} & x_{57} \end{bmatrix}.$$

Every alternative is assigned a score in a triangular fuzzy number based on all the existing criteria in the fuzzy TOPSIS decision matrix. A triangular fuzzy number is represented in the form of $\tilde{X} = \left(a_{ij,b_{ij}}, c_{ij} \right)$ where $\tilde{X}$ indicates the performance of alternative *i* where $i = (1, 2, \dots, m)$ relative to the criterion *j* where $j = (1, 2, \dots, n)$. The next step is defining the importance weight of each criterion represented by $\widetilde{W} = \left[\tilde{W}_1, \tilde{W}_2, \tilde{W}_3 \right]$.

Each criterion can have a different priority from the viewpoint of each node based on that node's preferences. In this approach, for the sake of simplicity, default values according to Table 4 are assigned to the selected set of criteria. The second phase of fuzzy TOPSIS is the normalization of the decision matrix. If *x* shows each element of the decision matrix and *n* indicates a normal matrix element, then the normalization could be performed based on Eq. 8. It is worth

mentioning that all criteria except the remaining path are positive factors. While Eq. 8 must be applied to the positive factors and negative ones.

$$
\begin{cases}
\tilde{n}_{(sourceNodeA)_{ij}} = \left(\frac{a_{ij}}{C_j^+}, \frac{b_{ij}}{C_j^+}, \frac{c_{ij}}{C_j^+} \right), & C_j^+ = \overset{max}{\underset{i}{}}\{C_{ij}\}; \quad \text{for benefit criteria} \\
\tilde{n}_{(sourceNodeA)_{ij}} = \left(\frac{a_j^-}{C_{ij}}, \frac{a_j^-}{b_{ij}}, \frac{a_j^-}{a_{ij}} \right), & \tilde{a}_j^- = \overset{min}{\underset{i}{}}\{\tilde{a}_{ij}\}; \quad \text{for cost criteria}
\end{cases}
\tag{8}
$$

where $\tilde{C}_j^+ = \overset{max}{\underset{i}{}}\{C_{ij}\}$ is for the *benefitcriteria* and $\tilde{a}_j^- = \overset{min}{\underset{i}{}}\{\tilde{a}_{ij}\}$ is for the *costcriteria*. In the third step, a free scale harmonic fuzzy matrix should be formed. Equations 9 and 10 transforms matrix N to a scale-free harmonic fuzzy matrix $\tilde{V}$.

$$
\tilde{V}_{(sourceNodeA)} = \left[\tilde{v}_{ij}\right]_{mn}, \quad i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n
\tag{9}
$$

$$
\tilde{V}_{(sourceNodeA)ij} = \tilde{n}_{(sourceNodeA)ij} * \tilde{W}_j
\tag{10}
$$

where $\tilde{W}_j$ is the importance weight of C_j criterion. A fuzzy weighted decision matrix is generated by multiplying the importance weight of each criterion to the scale-free fuzzy matrix. In the fourth step, positive ideal (A+) and negative ideal (A−) must be calculated by Eqs. 11 and 12 respectively. In this step, the positive and negative ideals are the biggest and smallest elements in each criterion column respectively.

$$
A^+ = \left(\tilde{v}_1^+, \tilde{v}_2^+, \ldots, \tilde{v}_n^+ \right) \quad \text{where } \tilde{v}_j^+ = \left(\tilde{C}_{MNj}^+, \tilde{C}_{MNj}^+, \tilde{C}_{MNj}^+ \right) \text{ and } \tilde{C}_j^+ = \overset{max}{\underset{i}{}}\{C_{MNij}\}
\tag{11}
$$

$$
A^- = \left(\tilde{v}_1^-, \tilde{v}_2^-, \ldots, \tilde{v}_n^- \right) \quad \text{where } \tilde{v}_j^- = \left(\tilde{a}_{MNj}^-, \tilde{a}_{MNj}^-, \tilde{a}_{MNj}^- \right) \text{ and } \tilde{a}_j^- = \overset{min}{\underset{i}{}}\{\tilde{a}_{MNij}\}
\tag{12}
$$

In the above equations, the A^+ and A^- options denote the best and the worst cases. The sum of the distances of each element from the fuzzy positive and negative ideals is calculated based on Eqs. 13 and 14 in the next step. This distance is calculated based on the Euclidean distance [37] as below:

$$
d_i^+(SourceNodeA) = \sum_{j=1}^{n} d\left(\tilde{v}_{ij}, \tilde{v}_j^+ \right), \quad i = 1, 2, \ldots, m \quad j = 1, 2, \ldots, n
\tag{13}
$$

$$
d_i^-(SourceNodeA) = \sum_{j=1}^{n} d\left(\tilde{v}_{ij}, \tilde{v}_j^- \right), \quad i = 1, 2, \ldots, m \quad j = 1, 2, \ldots, n
\tag{14}
$$

where D is the distance between two fuzzy numbers. If $M_1 = (a_1, b_1, c_1)$ and $M_2 = (a_2, b_2, c_2)$ are two triangular fuzzy numbers, then their distance could be calculated based on Eq. 15. The closer the distance to 1, the more ideal the result would be.

$$
D(M_1, M_2)_{(SourceNodeÃ)} = \sqrt{ \frac{1}{3} \left[(a_1 - a_2)^2 + (b_1 - b_2)^2 + (c_1 - c_2)^2 \right] }
\tag{15}
$$

As the next step, the similarity index to CC_i should be calculated based on Eq. 16.

$$CC_{(FinaTrust-MiddleNodes)i} = \frac{d_i^-(SourceNodeA)}{d_i^-(SourceNodeA) + d_i^+(SourceNodeA)} \quad (16)$$

Finally, nodes will be sorted based on CC_i which will be the indicator of the final trust value and utility function. The node with the highest trust value will be selected as the next node. Afterward, the decision matrix and the ranking of middle nodes will be repeated again. In each step, after ranking all available alternatives based on the fuzzy TOPSIS method, the nodes with the highest utility function will be inserted in the open list according to the A* algorithm. On the other hand, middle nodes with the least utility function will be inserted in the closed list. Now, in each step from the source node to the destination, the nodes with the highest utility function for traversing the most trustworthy path should be selected. The final step in the proposed approach is inferring the trust value between each unconnected pair of nodes in the trust network based on the selected trustworthy paths. This aim can be fulfilled by taking advantage of Eq. 17.

$$FinalTrust_{A,C} = \begin{cases} \frac{FT_{u1,u2}+FT_{u2,u3}\cdots+FT_{u(n_{path}-1),un}}{n_{path}} \\ 0 < FinalTrust_{A,C} \leq 1 \end{cases} \quad (17)$$

It is worth noting that in the A* algorithm the F function is evaluated using Eq. 1, while here, F as the final trust value is calculated by using the fuzzy TOPSIS approach.

5 Numerical example

Consider Fig. 3 as a slice of a trust graph network. Node *A* as the source node should decide on trusting or not trusting node *C*. As there is no direct connection between *A* and *C*, the most trustworthy path between these two nodes should be selected based on a combination of the A* algorithm and fuzzy TOPSIS multi-criteria decision making to infer the trust value. Node *A* is connected to nodes *B*, *D, E, F,* and *G* and hence it should infer the trust level of node *C* based on the information generated by its neighbors. However, for selecting the most trustworthy middle node, it should take advantage of the defined set of criteria. Table 5 illustrates the default value of each of these sets of criteria. As qualitative evaluations are estimates of our subjective opinions [25], in this work, we have considered qualitative variables using triangular fuzzy numbers to express such subjectivity toward middle users.

As mentioned previously, the second phase of the fuzzy TOPSIS method is the decision matrix normalization based on Eq. 8 for positive and negative criteria. Table 6 shows the results for the second phase of the fuzzy TOPSIS.

In the third step, based on Eqs. 9 and 10, the weighted fuzzy decision matrix will be generated based on the weight of each criterion. Table 7 is the weighted fuzzy decision matrix of Table 5. In the fourth step, the positive ideal (A^+) and negative

Table 5 The decision matrix of node A

D1 (Node A)

	Positive	Positive	Positive	Positive	Positive	Positive	Negative
Weight of criteria	(0.3,0.5,0.7) M	(0.3,0.5,0.7) M	(0.7,0.9,1) VH	(0.1,0.3,0.5) L	(0.5,0.7,0.9) H	(0,0.1,0.3) VL	(0.5,0.7,0.9) H
Alternatives	C1 (Similarity Profile)	C2 (Topography Similarity)	C3 (Transaction Rate)	C4 (Local Trust)	C5 (Context Trust)	C6 (Dunbar Theory)	C7 (Remaining Path)
B	0.87	0.89	0.86	(0, 1, 3)	(0, 0.25, 0.5)	0.58	1
D	0.66	0.95	0.97	(1, 3, 5)	(0.75, 1, 1)	0.3	3
E	0.33	0.08	0.05	(3, 5, 7)	(0.25, 0.5, 0.75)	0.001	1
F	1	0.99	1	(7, 9, 10)	(0.75, 1, 1)	0.99	5
G	0.62	0.04	0.001	(5, 7, 9)	(0, 0, 0.25)	0.66	2

Table 6 Decision matrix normalization

Free scale matrix	C1-SP			C2-ST			C3-TR			C4-LT			C5-CT			C6-DU			C7-RP		
Limit->	L	M	U	L	M	U	L	M	U	L	M	U	L	M	U	L	M	U	L	M	U
B	0.87	0.87	0.87	0.9	0.9	0.9	0.86	0.86	0.86	0	0.1	0.3	0	0.25	0.5	0.59	0.59	0.59	1	1	1
D	0.66	0.66	0.66	0.96	0.96	0.96	0.97	0.97	0.97	0.1	0.3	0.5	0.75	1	1	0.3	0.3	0.3	0.33	0.33	0.33
E	0.33	0.33	0.33	0.08	0.08	0.08	0.05	0.05	0.05	0.3	0.5	0.7	0.25	0.5	0.75	0	0	0	1	1	1
F	1	1	1	1	1	1	1	1	1	0.7	0.9	1	0.75	1	1	1	1	1	0.2	0.2	0.2
G	0.62	0.62	0.62	0.04	0.04	0.04	0	0	0	0.5	0.7	0.9	0	0	0.25	0.67	0.67	0.67	0.5	0.5	0.5

Table 7 The weighted fuzzy decision matrix

Weighted matrix	C1-SP			C2-ST			C3-TR			C4-LT			C5-CT			C6-DU			C7-RP		
Limit->	L	M	U	L	M	U	L	M	U	L	M	U	L	M	U	L	M	U	L	M	U
B	0.26	0.44	0.61	0.27	0.45	0.63	0.6	0.77	0.86	0	0.03	0.15	0	0.18	0.45	0	0.06	0.18	0.5	0.7	0.9
D	0.2	0.33	0.46	0.29	0.48	0.67	0.68	0.87	0.97	0.01	0.09	0.25	0.38	0.7	0.9	0	0.03	0.09	0.17	0.23	0.3
E	0.1	0.17	0.23	0.02	0.04	0.06	0.04	0.05	0.05	0.03	0.15	0.35	0.13	0.35	0.68	0	0	0	0.5	0.7	0.9
F	0.3	0.5	0.7	0.3	0.5	0.7	0.7	0.9	1	0.07	0.27	0.5	0.38	0.7	0.9	0	0.1	0.3	0.1	0.14	0.18
G	0.19	0.31	0.43	0.01	0.02	0.03	0	0	0	0.05	0.21	0.45	0	0	0.23	0	0.07	0.2	0.25	0.35	0.45

Table 8 The numerical results of the positive ideal and negative ideal

Solutions	C1-Sp			C2-ST			C3-TR			C4-LT			C5-CT			C6-DU			C7-RP		
Limit->	L	M	U	L	M	U	L	M	U	L	M	U	L	M	U	L	M	U	L	M	U
Positive ideal	0.7	0.7	0.7	0.7	0.7	0.7	1	1	1	0.5	0.5	0.5	0.9	0.9	0.9	0.3	0.3	0.3	0.9	0.9	0.9
Negative ideal	0.1	0.1	0.1	0.01	0.01	0.01	0	0	0	0	0	0	0	0	0	0	0	0	0.1	0.1	0.1

Table 9 The sum of interval for each element from the fuzzy positive ideal

Distance from Positive ideal	C1-SP	C2-ST	C3-TR	C4-LT	C5-CT	C6-DU	C7-RP	d+
B	0.3	0.3	0.3	0.4	0.7	0.2	0.3	2.5
D	0.4	0.3	0.2	0.4	0.3	0.3	0.7	2.5
E	0.5	0.7	1	0.3	0.6	0.3	0.3	3.6
F	0.3	0.3	0.2	0.3	0.3	0.2	0.8	2.3
G	0.4	0.7	1	0.3	0.8	0.2	0.6	4

Table 10 The sum of intervals for each element from the fuzzy negative ideal

Distance from Negative ideal	C1-SP	C2-ST	C3-TR	C4-LT	C5-CT	C6-DU	C7-RP	d−
B	0.4	0.5	0.8	0.1	0.3	0.1	0.6	2.7
D	0.3	0.5	0.8	0.2	0.7	0.1	0.1	2.6
E	0.1	0	0	0.2	0.4	0	0.6	1.4
F	0.4	0.5	0.9	0.3	0.7	0.2	0.1	3.1
G	0.2	0	0	0.3	0.1	0.1	0.3	1

**Table 11** The similarity index and trust score of all alternatives for the source node *A*

Result	CCi	Score
F	0.5753	1
B	0.5149	2
D	0.5131	3
E	0.2853	4
G	0.2072	5

ideal (A^-) should be calculated based on Eqs. 11 and 12. As mentioned earlier, in this step, the positive ideal is equivalent to the greatest element in each criterion column and the negative ideal is the smallest one. Table 8 demonstrates the numerically calculated results of these parameters.

As for the fifth step, the sum of intervals for each element from the positive and negative ideals based on Eqs. 13, 14, and 15 is calculated. The results are shown in Tables 9 and 10.

Similarity index to the ideal alternative (CC_i) should be assessed based on Eq. 16 as the sixth step and this should be followed by a ranking of middle nodes. The similarity index and the trustworthiness grade of all these alternatives are presented in Table 11.

According to the calculated results in Table 11, the utility function of all middle nodes from the perspective of *A* is calculated. The score of node *F* is considerably higher than other nodes (*B, D, E,* and *G*). These values are calculated based on all
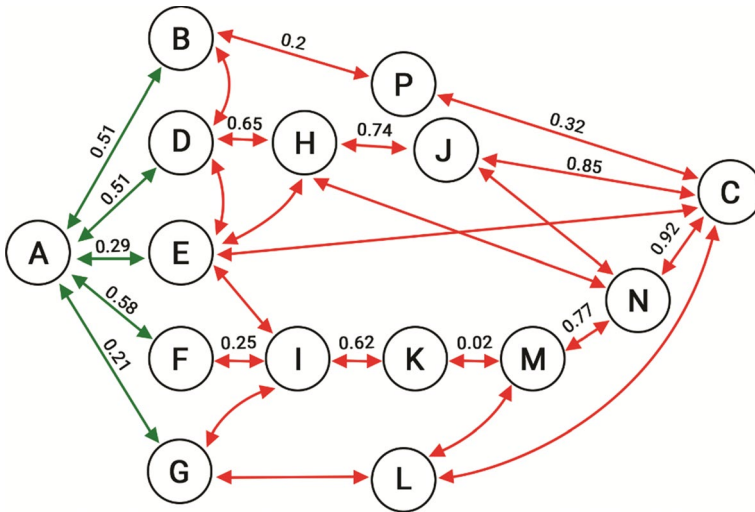
Fig. 7 Middle users ranking using the fuzzy-TOPSIS MCDM approach

the criteria stated previously and through the fuzzy TOPSIS method which is different from the case where only one criterion is considered. Node *A* can assess its trust value toward node *C* based on the results of Table 11 and the calculated final trust values in Fig. 7. Even though the remaining path length of nodes *E* and *G* toward the destination is just 1 and 2 respectively, their final trust is amongst the lowest values. Therefore, their closeness to the target node does not have any impact on *A*'s decision and both of these nodes would be omitted from the open decision list of *A*. Now, three other nodes *B, D*, and *F* are remaining. The utility function of *B* and *D* is 0.5149 and 0.5131 respectively which is roughly equal. However, the remaining path length from *B* is 2, while this is 3 in the case of *D* as a source node. Note that although the path (D-E-C) is shorter than (D-H-J-C), since the utility function of node *E* (i.e. one of the middle nodes of the former path) is the lowest value, this path is removed from the open list of *A*.

   Although node *F* has the highest utility value amongst (*B, D*, and *F*), the remaining path to the target from this node is 5, which is the longest possible path. Again, notice that the path (F-I-E-C) is removed from the open list, since *E*, one of the middle nodes has the lowest utility value. Now, *A* should choose between *B* with final trust value equal to 0.5149 and the shortest trust chain (A-B-P–C), *D* with the final trust value equal to 0.5131, or *F* with the final trust value equal to 0.58 but with the longest trust chain. In this case, a threshold of 0.6 is considered as the minimum acceptable value in the local trust chain. Of course, this is a parameter and can be adjusted in the model. Suppose that the utility function of one of these local chains of middle nodes is in this range. In this stage, choosing a middle node with a local trust value greater than the uncertainty range is the ultimate goal. So, in this step, again, all the 7 steps of the fuzzy TOPSIS should be repeated to rank the utility function of all middle nodes to find the most trustworthy trust chain for node *A*. These steps have been

performed for these 3 middle nodes and the results are shown in Fig. 7. The local trust chain of F has two hops less than the uncertainty range. The utility function of the local trust chain of B is BP$=0.2$ and PC$=0.32$, and is less than the uncertainty range. However, the utility function of the local trust chain originating from node D is greater than 0.6 and it would be selected as the most trustworthy path. Therefore, selecting middle nodes with high utility values while having control over the remaining path length is the suggested approach based on the A* algorithm. After selecting the trustworthy path, the trust between unconnected nodes should be inferred. The inferential trust value could be calculated based on Eq. 17 as follows.

$$Trust_{A,C} = \frac{0.51 + 0.65 + 0.74 + 0.85}{4} = 0.68$$

In the proposed approach, longer paths with higher utility values are preferred to shorter paths with lower utility values as they indicate the existence of more stable trust chains.

6 Evaluations and comparisons

The purpose of this section is to evaluate the performance and correctness of the proposed approach. To this aim, the proposed approach is simulated alongside well-known algorithms such as TidalTrust and MoleTrust using the Facebook and Twitter datasets. The evaluation metrics are (1) the average trustworthy path length selected by the algorithms between two unconnected nodes, and (2) the correctness of predicted inferred trust values by using the stated datasets. In what follows, first, the simulation setup is discussed and then evaluation scenarios are given.

6.1 Simulation setup

We have used Facebook and Twitter datasets [23] for the evaluation of the proposed approach. Table 12 shows the topological structure of these datasets which is

Table 12 The topological structure of the Facebook and Twitter datasets

| Type of data	Facebooks	Twitter
Number of nodes	786	1970
Number of edges	28,048	17,620
Modularity	0.571	0.269
Number of communities	9	2
Network density	0.045	0.005
Network diameter	10	5
Average path length	3.042	2.289
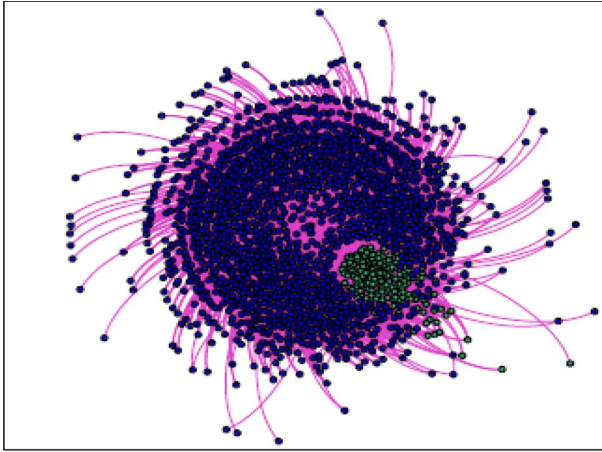
Fig. 8 Twitter dataset topology
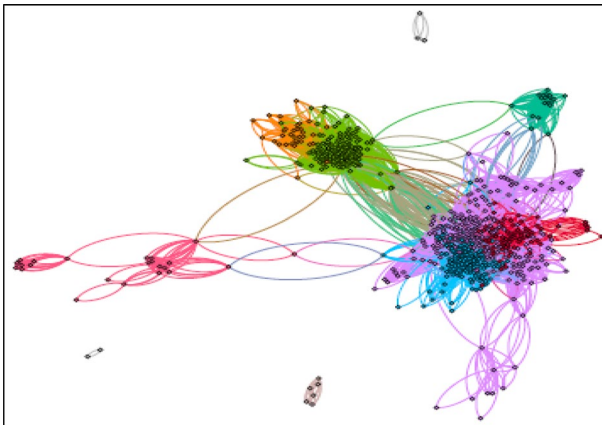
Fig. 9 Facebook dataset topology

generated by the Gephi network visualization tool [38]. Figures 8 and 9 show Twitter and Facebook's topological structures respectively.

The diameter of a network is defined as the longest of all the calculated shortest paths between nodes and shows the extent of the network. The diameter of a network is the shortest path length between two nodes and indicates how readily individuals in this network are accessible [39]. Path length refers to the efficient length of the network. Average path-length shows the distance between a pair of nodes which is a criterion of how much close the nodes are together. The average path length in many complicated networks is too short which is known as the small-world phenomenon. A network is defined as small if the path length between every two random nodes
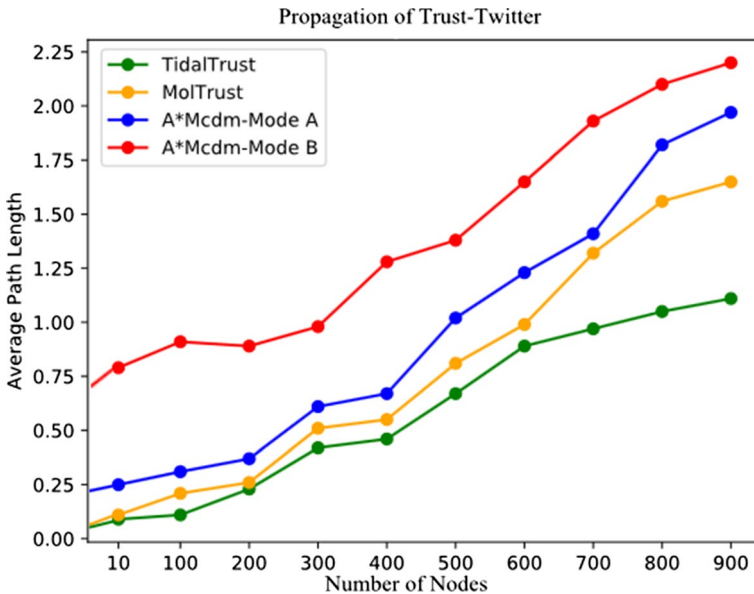
Fig. 10 Trust propagation length comparison in the Twitter network

in a graph never exceeds 6 steps [40]. Based on Figs. 10 and 11 and Table 12, it can be easily seen that the Facebook network data is more scattered and its density is much higher compared to the Twitter dataset with a few nodes and dense edges. Modularity in social networks is used to identify communities [41]. Such modules are represented in the figures with different colors. For topological visualization of these networks, Yifan Hu's algorithm is applied. This algorithm has high accuracy and low time complexity, and it is suitable for massive networks [42]. The simulation is implemented in Java and visualization of the output of the simulation is performed with Gephi. Java is selected for the implementation of simulation scenarios as Gephi is an open-source tool implemented using Java and based on the NetBeans framework [38]. As both TidalTrust and MoleTrust evaluated their algorithms based on these two datasets, here again, Facebook and Twitter datasets are selected. Also, as the aim of those aforementioned algorithms is too close to the one suggested in this paper, their results seem to be a good metric for the evaluation [2, 3, and 15]. In addition, both of these algorithms as well as the proposed algorithm in this paper are local algorithms which is another reason to be selected for comparisons. The simulations are repeated ten times for the Twitter dataset and in each step, a different number of unconnected nodes (10, 100, 200, 300, 400, 500, 600, 700, 800, and 900) is considered. The simulation repetition number for the Facebook dataset is eight and the number of unconnected nodes in each step is as follows: (10, 50, 100, 150, 200, 250, 300, and 350). The reason for considering a different number of unconnected nodes in each step of the algorithms is the difference in the maximum number of nodes in each dataset.
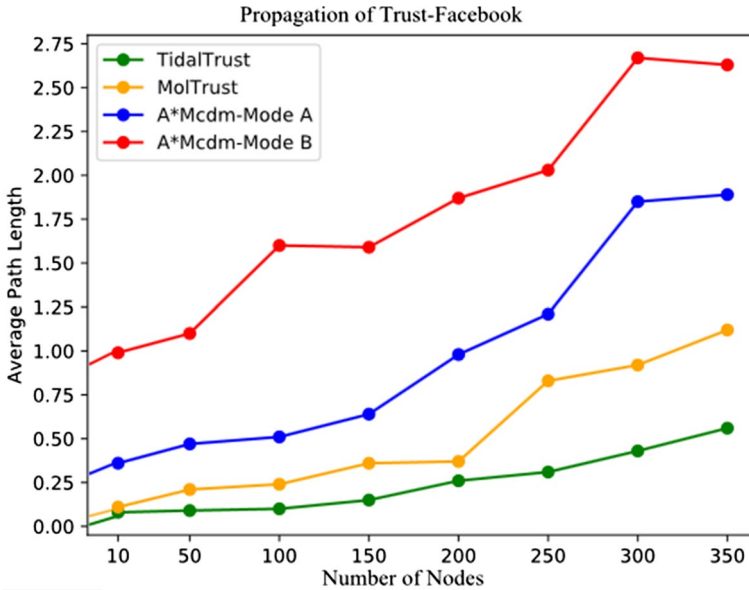
Fig. 11 Propagation level comparison in the Facebook network

Table 13 MAE results based on the Facebook and Twitter datasets

Dataset		Number of users	TidalTrust	MoleTrust	Proposed approach
	Twitter	1970	0.38	0.34	0.16
	Facebook	786	0.52	0.41	0.21

The simulation is performed based on the trust propagation distance amongst unconnected nodes for the three algorithms. The result of each simulation round for the three algorithms is visualized by Gephi and are shown in Figs. 10 and 11. The calculated average path lengths are saved and then averaged. Another evaluation factor of the simulation is the accuracy of inferential trust value anticipation. The results show that the suggested approach works much better than the other two algorithms in this case. Table 13 shows the results for the accuracy of inferential trust values using these three algorithms.

6.2 Simulation scenarios

The first important factor for the evaluation is the comparison between the lengths of the selected paths between pairs of unconnected nodes in the network. This evaluation is performed with different sets of unconnected nodes over two different datasets with different topological structures. The second factor is the importance of considering a wider range of fuzzy and potentially conflicting trustworthiness

criteria for the trust calculation of middle nodes resulting in smaller and more trust-worthy paths between the source and destination nodes. Most of the previous studies have used partial information for finding trustworthy paths and inferring trust in social networks. For example, the TidalTrust algorithm has used topological similarity, Adali et al. focused on trust behavioral information and the connection time between users [43], and Zhan et al. utilized the profile similarity, reliability of the information, and social comments [44]. None of them have considered all the important criteria suggested in the proposed approach simultaneously. Hence, our aim in this evaluation scenario is to show the accuracy of the calculated trust values and demonstrate the importance of considering the set of criteria selected in this work. The simulation results are shown based on two modes *A* and *B* to do the comparison in two different formats for the topological network structure (i.e. partial graph information) and a combination of topological graph structure and interaction rates (i.e. comprehensive graph information). Mode *A* is used for the simulation of the suggested algorithm and studying how topological network structure could make a difference by considering only three metrics for profile similarity, topological similarity, and the remaining steps toward the target node. Mode *B* is defined for the simulation of the suggested algorithm for investigating how a combination of similarity criteria between users and interaction rates by considering all the aforementioned set of criteria is compatible with Fig. 4.

When an algorithm uses partial graph information and just one or two criteria for trust evaluation, pair comparison of nodes based on the values labeled on the edges is reasonable. However, in this approach for enhancing the accuracy of trust, a set of criteria simultaneously are considered for evaluating the final trust value and ranking each middle node. Therefore, for selecting trust chains between unconnected nodes, a utility function with threshold 0.6 is used for each local chain. Also, middle nodes with high values are selected as middle trustworthy users for traversing the trust path. As the first step, by considering a different number of middle nodes for selecting the trustworthy paths, propagation trust lengths in both networks with two different scattered and wave-like topological networks are studied. These propagation trust lengths are compared with the average path length between selected unconnected nodes. Figures 10 and 11 are representing the accuracy of the inferred trust values against the Twitter and Facebook datasets compared with the TidalTrust and MoleTrust algorithms. Moreover, the *A* and *B* modes of the suggested algorithm are visualized by the Gephi software.

According to Fig. 10, as the main idea of both TidalTrust and MoralTrust algorithms is to find the shortest trustworthy paths, a large majority of reliable paths in the TidalTrust and MoleTrust algorithms are from the same society and in the center of the vicinity of those selected nodes. A large portion of these trustworthy paths is in the range of nodes with high-density edges. This is mainly because both these two algorithms focus on the topological structure. So they find the selected paths when there is a high density of edges and the number of mutual friends to the total number of friends is higher. The mode *B* results of the suggested algorithm in this figure represents a great performance improvement in trust propagation. Also, it shows a higher number of reliable longer paths with even greater utility function due to considering a set of criteria simultaneously

compared to mode *A* of this algorithm which only takes into account two criteria of profile and topological similarity. The ultimate goal of the proposed algorithm is not only selecting the short reliable paths. It is selecting middle users with high utility functions and finding local trust chains with threshold values higher than 0.6 whether in shorter paths or longer ones. Various trust properties considered in this approach such as being context-dependent and calculating local trust facilitates the propagation of trust to further distances.

It is worth mentioning that for each of the datasets, to show the better performance of the proposed approach especially in higher scales, more unconnected pairs are selected for every iteration of the simulation. In each iteration, the average path distances between pairs of unconnected nodes are calculated. The selected paths in the proposed approach are either trust paths with shorter distances or longer ones but with higher certainty values especially when the nodes are selected from the same communities. Contrary, by considering partial graph information and confining themselves to only the shortest possible paths, similar approaches such as TidalTrust and MoleTrust can only recognize the trustworthy paths in a very limited domain and scope.

As Fig. 11 illustrates, TidalTrust faces more difficulty in finding the trustworthy path and middle nodes as the scattered Facebook network grows. The primary cause might be related to just considering the shortest possible paths. Moreover, as it is highly dependent on the trust web's density, the number of detectable trustworthy paths are larger in high population density networks compared to low-density ones.

When a pair of nodes are selected in communities with larger distance, the accuracy of trust calculations significantly decreases, especially in the Facebook's sparse network structure. This is due to the reliance on just the network's topological data rather than the local trust calculation.

Albeit, the accuracy of the MoleTrust algorithm is better than the TidalTrust, its time complexity is higher than the later. Surprisingly, the simulation result for the mode *B* of the proposed algorithm demonstrates a significant accuracy improvement compared to its counterpart approaches. This is due to the consideration of various trustworthiness criteria and especially local trust information for the trust calculation of middle-nodes.

It also shows more capability in trust propagation to further distances and other societies with higher accuracy. The first phase of the evaluations focused on comparing the proposed approach with similar approaches using two different datasets. In the second phase, based on the mean absolute error metric given in Eq. 18, the accuracy of the suggested approach is evaluated according to both the Facebook and Twitter datasets.

$$MAE = \sqrt{\frac{Trust_{ij} - \widetilde{Trust_{ij}}}{N}} \qquad (18)$$

where $Trust_{ij}$ is the direct trust of nodes *i* and *j* and $\widetilde{Trust_{ij}}$ is the calculated trust of nodes *i* and *j* based on the proposed algorithm. Table 13 shows the calculated values for all of these three algorithms.

As shown, Table 13 demonstrates how the proposed approach is more accurate compared to its counterparts even with the existence of longer distances and further trust propagation distances. The distinguishing feature of the proposed approach which makes it potentially more accurate is the consideration of longer possible routes and a comprehensive set of criteria for decision making using simultaneously the quantitative and qualitative parameters based on the multi-criteria decision-making approaches. Another advantage of this approach is its simplicity together with its high performance as the result of taking advantage of the fuzzy TOPSIS technique which leads to considerable time complexity reduction.

The time complexity for the A* algorithm majorly depends on the selected heuristic function and in worst-case it has a $o(b^d)$ complexity. Here, b is the branching factor and d is the depth of the optimum solution. The number of nodes that are visited in the A* algorithm can be calculated using the following formula:

$$N + 1 = 1 + b^* + (b^*)^2 + \cdots + (b^*)^d$$

where b^* is the effective branching factor. Because in the proposed approach, fuzzy TOPSIS is used many of the branches from the source node will be assigned to the closed list due to their lower appropriateness. This will result in the reduction of the number of nodes that are placed in the open list and are thus visited. On the other hand, in this context, d is the depth of the social network. Based on the principle of six degrees of separation [45] we know that this depth on average will not exceed 6 in the domain of social networks. Both of these characteristics will result in a significant improvement in terms of time complexity compared to the classic A* algorithm.

Figures 12 and 13 represent the propagated trust distance and the correctness of the inferred trust values of the proposed approach compared with the TidalTrust and MoleTrust algorithms respectively in the Twitter and Facebook networks. The simulations are performed for the cases where the pairs of unconnected nodes are selected either from the same modularity with high density and lower path lengths or from different modularities with lower density and higher path lengths. In the
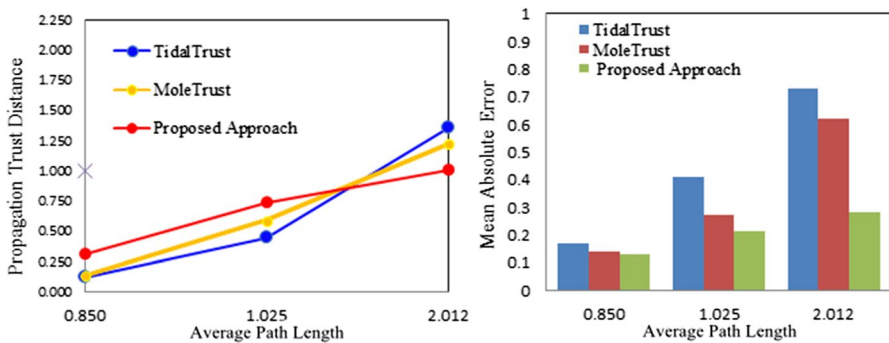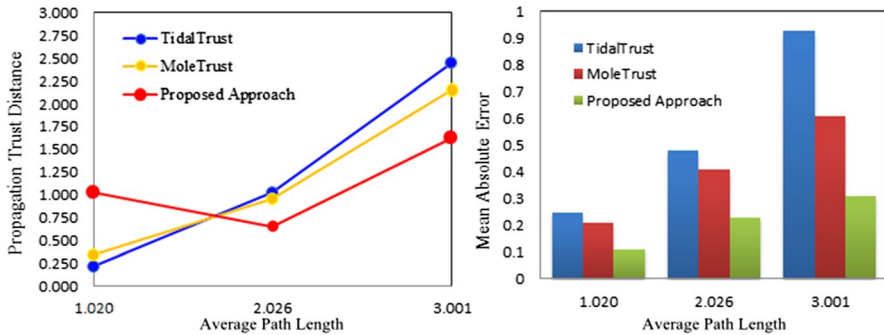

Fig. 12 Propagation trust distance of two disconnected nodes from the same or different modularities in the Twitter network

Fig. 13 Propagation trust distance of two disconnected nodes from the same or different modularities in the Facebook network

first simulation, where the pairs of unconnected nodes are selected from the same modularity with lower path lengths, because the Tidaltrust and Moletrust algorithms only consider the topological information of the network, in all the three algorithms, the focus is more on lower trust path lengths. In the Twitter network, because the density of the network is high (i.e. smaller number of nodes with larger degrees) the propagated trust distance is lower and the error in all the three algorithms is smaller compared to the Facebook network.

In the second and third iteration, where the pairs of unconnected nodes are selected from different modularities and with higher average path length, it is evident from the results that the accuracy of the TidalTrust and MoleTrust algorithms decreases. This is more noticeable in the sparser Facebook network. In these cases, the accuracy of the proposed approach is not significantly affected and the difference between the TidalTrust and MoleTrust algorithms and the proposed approach becomes more significant.

7 Conclusions

Finding efficient and trustworthy paths within massive social networks has always been a big challenge for researchers. The accuracy of trust values on directed edges and inferential trust value between two unconnected pair of nodes is highly dependent on a set of criteria used for assessing trust. In this paper, the fuzzy TOPSIS multi-criteria decision-making method along with the A* algorithm is utilized for opting the most trustworthy middle users, computing the trust level of each middle user (i.e. utility function) and finally selecting the proper trustworthy path. The ultimate goal of the suggested algorithm is to find trustworthy middle nodes and as a result of that the most trustworthy path while having control over trust paths' length. This is performed by considering a wide range of criteria in calculating the trust value and recognizing reliable routes amongst unconnected pairs of nodes with high certainty. The proposed method's performance for finding relatively longer but more certain reliable paths has been compared against

two of the well-known algorithms proposed before namely the TidalTrust and MoleTrust algorithms according to the scattered Facebook dataset and wave-like Twitter networks respectively. It is worth mentioning that the more the trust propagation distance will be while preserving the accuracy of inferential trust between unconnected pairs of nodes, the more expansion of users' relations will occur. Also, the trust information propagation in the social network will be quicker as well.

The A* algorithm with its great flexibility, accuracy and high performance can readily be implemented in complex networks and it has been used in various science majors. Its implementation could be altered based on environmental situations. For example, utilizing the A* algorithm in the game development domain could be a process-intensive task as it requires very complicated computations. However, here, the A* algorithm is used in the field of social networks. The nonexistence of either complicated parameters or sudden rotations as a result of facing insuperable obstacles in an environment which are the main contributing factors in the complexity of the pathfinding process makes it a perfect choice to be used for our purpose. In other words, the simplicity of social networks in which users are the only obstacles within the path could significantly increase the performance and simplicity of the algorithm. Notice that one of the main shortcomings of the A* algorithm is the high memory requirement as it needs to preserve the information of all available nodes. Besides, trust is defined as a fuzzy subject and the selection of one path amongst the existing ones is performed by using the fuzzy-TOPSIS multi-criteria decision-making approach. Many paths are omitted due to the lack of certainty which can decrease the complexity of the algorithm.

As for the future works of this research, it is intended to expand the suggested approach by utilizing more efficient versions of the A* which requires less memory with more massive data and larger data sets. It is worth mentioning that a disadvantage of the classic A* algorithm is the exponential memory requirement due to using a priority queue as well as the need to keep all the traversed nodes in memory. Extensions of this algorithm such as the iterative deepening A* (i.e. IDA*) as well as simplified memory bounded A* (i.e. SMA*) are introduced specifically to reduce the memory requirements of the A* algorithm [46]. Evaluating the proposed approach by using these types of extensions can be considered as an interesting future research to consider which makes the approach more suitable to denser networks and hence more real-world applicable. Additionally, due to the dynamic nature of trust, it would be a beneficial addition to also consider the time criterion amongst the set of criteria for the multi-criteria decision-making approach. In this way, the decay of trust through time and the different weights that trust values over a time period may have can be taken into account. Finally, examining other fuzzy multi-criteria decision making approaches such as VIKOR or TODIM [47] may create interesting results with their unique strengths.

Compliance with ethical standards

Conflict of interest All of the authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Ntwiga D, Weke P, Manene M, Waniki JM (2016) Modeling trust in social network. Int J Math Arch 7(2):2229–5046
2. Kianinejad M, Kabiri P (2018) A strategy for trust propagation along the more trusted paths. In: Proceedings of the IEEE 3rd conference on swarm intelligence and evolutionary computation (CSIEC), 2018, Piscataway, New Jersey, pp 1–6
3. Bhattacharya M, Nesa N (2016) An algorithm for predicting local trust based on trust propagation in online social networks. Int J Comput Appl 156(7):8887
4. Yadav A, Chakraverty S, Sibal R (2015) A survey of implicit trust on social networks. In: Proceedings of the IEEE international conference on green computing and internet of things (ICGCIoT), 2015, Greater Noida, India, pp 1511–1515
5. Kim YA, Song HS (2011) Strategies for predicting local trust based on trust propagation in social networks. Knowl-Based Syst 24(8):1360–1371
6. Qian G, Wang H, Feng X (2013) Generalized hesitant fuzzy sets and their application in decision support system. Knowl-Based Syst 37:357–365
7. Decision making under Certainty, Risk and uncertainty. http://adminscience.blogspot.com/2012/06/decision-making-under-certainty-risk.html, Accessed 4 Jul 2019.
8. Mardani A, Jusoh A, Zavadskas EK (2015) Fuzzy multiple criteria decision-making techniques and applications. Expert Syst Appl 42(8):4126–4148
9. Lesani M, Montazeri N (2009) Fuzzy trust aggregation and personalized trust inference in virtual social networks. Comput Intell 25(2):51–83
10. Yao J, Lin C, Xie X, Wang AJ, Hung CC (2010) Path planning for virtual human motion using improved A-star algorithm. In: Proceedings of the 7th IEEE international conference on information technology: New generations (ITNG), 2010, Las Vegas, Nevada, pp 1154–1158
11. Naderan M, Namjoo E, Mohammadi S (2019) Trust classification in social networks using combined machine learning algorithms and fuzzy logic. Iran J Electr Electron Eng 15(3):294
12. Al-Oufi S, Kim HN, El Saddik A (2012) A group trust metric for identifying people of trust in online social networks. Expert Syst Appl 39(18):13173–13181
13. Jiang W, Wang G, Wu J (2014) Generating trusted graphs for trust evaluation in online social networks. Future Gener Comput Syst 31:48–58
14. Chakraborty PS, Karform S (2012) Designing trust propagation algorithms based on simple multiplicative strategy for social networks. Proc Technol 6:534–539
15. Ghavipour M, Meybodi MR (2018) Trust propagation algorithm based on learning automata for inferring local trust in online social networks. Knowl-Based Syst 143:307–316
16. Sherchan W, Nepal S, Paris C (2013) A survey of trust in social networks. ACM Comput Surv 45:1–33
17. Kim YA (2015) An enhanced trust propagation approach with expertise and homophily-based trust networks. Knowl-Based Syst 82:20–28
18. Hasan O, Brunie L, Pierson JM (2009) Evaluation of the iterative multiplication strategy for trust propagation. In: Proceedings of the ACM international conference on pervasive services, 2009, London, United Kingdom, pp 49–54
19. Hamdi S, Gancarski AL, Bouzeghoub A, Yahia SB (2016) TISoN: Trust Inference in Trust-Oriented Social Networks. ACM Trans Inf Syst 34(3):17
20. Rahman FH, Au TW, Newaz SH, Suhaili WS (2017) Trustworthiness in fog: a fuzzy approach. In: Proceedings of the VI international conference on network, communication and computing, 2017, Kunming, China, pp 207–211

21. Yi Ch, Yuanyuan B, Jingchi J, Yibo X (2015) Modeling cascading failures with the crisis of trust in social networks. Phys A 436:256–271
22. Rababah O, Alqudah B (2018) Building a trust model for social network. Modern Appl Sci 12(4)
23. Hamzelou N, Ashtiani M (2019) A mitigation strategy for the prevention of cascading trust failures in social networks. Future Gener Comput Syst 94:564–586
24. Nădăban S, Dzitac S, Dzitac I (2016) Fuzzy TOPSIS: a general view. Proc Comput Sci 91:823–831
25. Sun CC, Lin GT (2009) Using fuzzy TOPSIS method for evaluating the competitive advantages of shopping websites. Expert Syst Appl 36(9):11764–11771
26. Sevkli M, Zaim S, Turkyilmaz A, Satir M (2010) An application of fuzzy TOPSIS method for supplier selection. In: Proceedings of the IEEE international conference on fuzzy systems, Jul 18, 2010, Barcelona, Spain, pp 1–7
27. Dijkstra's shortest path algorithm. https://www.geeksforgeeks.org/dijkstras-shortest-path-algorithm-greedy-algo-7/, Accessed 16 July 2019.
28. Ashtiani M, Hakimi-Rad S (2018) MA Azgomi (2018) A model of trust based on uncertainty theory. Int J Uncertain Fuzz Knowl-Based Syst 26(2):269–298
29. Dunbar RIM, Arnaboldi V, Conti M, Passarella A (2015) The structure of online social networks mirror those in the offline world. Soc Netw 43:39–47
30. Mac-Carron P, Kaski K, Dunbar R (2016) Calling Dunbar's numbers. Soc Netw 47:151–155
31. Zolfaghar K, Aghaie A (2012) A syntactical approach for interpersonal trust prediction in social web applications, combining contextual and structural data. Knowl-Based Syst 26:93–102
32. Han X, Wang L, Crespi N, Park S, Cuevas Á (2015) Alike people, alike interests? Inferring interest similarity in online social networks. Decis Support Syst 69:92–106
33. Srilatha P, Manjula R (2016) Similarity index based link prediction algorithms in social networks: a survey. J Telecommun Inf Technol 2:87–94
34. Borbora ZH, Ahmad MA, Oh J, Haigh KZ, Srivastava J, Wen Z (2013) Robust features of trust in social networks. Soc Netw Anal Min 4:981–999
35. Sun G, Guan X, Yi X, Zhou Z (2018) Grey relational analysis between hesitant fuzzy sets with applications to pattern recognition. Expert Syst Appl 92:521–532
36. Zadeh LA (2015) Fuzzy logic—a personal perspective. Fuzzy Sets Syst 281:4–20
37. How to Calculate Euclidean Distance, https://sciencing.com/how-to-calculate-euclidean-distance-12751761.html. Accessed 20 July 2019.
38. Cherven K (2015) Mastering Gephi network visualization. Packt Publishing Ltd
39. Denny M (2014) Social network analysis. Institute for Social Science Research, University of Massachusetts, Amherst
40. Golbeck J (2013) Analyzing the social web, newness; 1st edn
41. Ji X, Machiraju R, Ritter A, Yen PY (2015) Examining the distribution, modularity, and community structure in article networks for systematic reviews. In: Proceedings of the AMIA annual symposium: american medical informatics association, U.S, San Francisco, pp 1927–1936
42. Gephi Tutorial Layouts. https://gephi.org/tutorials/gephi-tutorial-layouts.pdf. Accessed 19 July 2019.
43. Adali S, Escriva R, Goldberg MK, Hayvanovych M, Magdon-Ismail M, Szymanski BK, Wallace W, Williams G (2010) Measuring behavioral trust in social networks. In: Proceedings of the IEEE international conference on intelligence and security informatics, 2010 May 23, Vancouver, BC, Canada, pp 150–152
44. Zhan J, Fang X (2011) A novel trust computing system for social. In: Proceedings of the 3rd IEEE international conference on privacy, Boston, MA, USA, pp 9–11
45. Kermani MA, Badiee A, Aliahmadi A, Ghazanfari M, Kalantari H (2016) Introducing a procedure for developing a novel centrality measure (Sociability Centrality) for social networks using TOPSIS method and genetic algorithm. Comput Hum Behav 56:295–305
46. Informed search methods. http://www.cs.cmu.edu/afs/cs/academic/class/15780-s13/www/lec/07-Informed-search.pdf, Accessed 16 July 2019.
47. Tosun Ö, Akyüz G (2015) A fuzzy TODIM approach for the supplier selection problem. Int J Comput Intell Syst 8(2):317–329