

Research Article

Consumers Team Detection Model Based on Trust for Multi-Level

Xiaoming Li,^{1,2} Guangquan Xu ,¹ Sandhya Armoogum ,³ and Honghao Gao ⁴

¹School of Computer Science and Technology, Tianjin University, Tianjin, China

²School of Information Science and Engineering, Zaozhuang University, Shandong, China

³School of Innovative Technologies and Engineering, University of Technology Mauritius, Pointe-Aux-Sables, Mauritius

⁴Computing Center, Shanghai University, Shanghai, China

Correspondence should be addressed to Honghao Gao; gaohonghao@shu.edu.cn

Received 20 July 2018; Revised 11 November 2018; Accepted 5 December 2018; Published 3 February 2019

Guest Editor: Jaegeol Yim

Copyright © 2019 XiaoMing Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Due to rapid advances in technology, social networks have become important platforms for daily communication, product marketing, and information dissemination. Targeted delivery of social network advertisement can considerably improve the efficacy of the advertisement and maximize the profits from it. In this context, managing the specific audience of a social network advertisement and achieving targeted advertisement delivery have been the ultimate goals of the social network advertising sector. Identifying user groups with similar properties is critical to increasing targeted sales. When both the scale of mobile social network and the complexity of social network user behaviors grow, similar groups are hidden in user behaviors. In order to analyze community structure with user trust relationship more appropriately in the large-scale multilevel social network environment, a novel local community detection model E-MLCD is proposed in this paper. It is jointly based on the multilevel properties and the strength of similarity of multilevel social interaction among communities. By studying three real-world multilevel social networks and specific QQ Zone marketing data, the model defines a new metric of community trust based on similarity. Comparison between other state-of-the-art detection methods demonstrate E-MLCD's ability to detect communities more effectively.

1. Introduction

Due to the growing diversity of customer requirements and the rapid advances in mobile social network, how to identify exact requirements of customers by distinguishing between customer groups is an important aspect of core competitiveness for enterprises. Identifying specific customer groups from the large-scale mobile social network refers to the issue of community detection.

Since Newman et al. proposed the Girvan–Newman (GN) algorithm in 2002 [1], a lot of attention has been paid across the globe to community detection. In recent years, community detection and community detection algorithms have become a focus of research on complex networks [2–5]. However, the multifaceted correlation between entity and property means that many complex systems are interrelated rather than independent. Consequently, the traditional single-layer network model cannot describe the system very accurately. In this context, a new type called multilevel complex network has not only emerged an extension of

existing network model but also represented a breakthrough in the entire network theory. In the multilayer complex network, the network structure is not completely flat. Instead, interlayer interaction is introduced. As a consequence, most of the traditional single-layer network evaluation methods are no longer suited for multilayer network [6, 7]. Figure 1 shows that the different consumer groups are found on multilayer networks according to geographical location.

The rapid progress in mobile network and electronic communications as well as the growing scale of network, the social networks (e.g., Facebook, Twitter, QQ, Zhihu, and Sina Weibo) have tens of millions of nodes and links [8, 9], making it very difficult to collect data across the entire network. Displaying advertisements and recommending products and friends to users can be more effective if specific groups identified through community detection. However, considering the colossal scale of these social networks, even if the parallel distributed platform is used to analyse the entire networks, the space and time consumption is too high to be acceptable. To make matters even worse, many

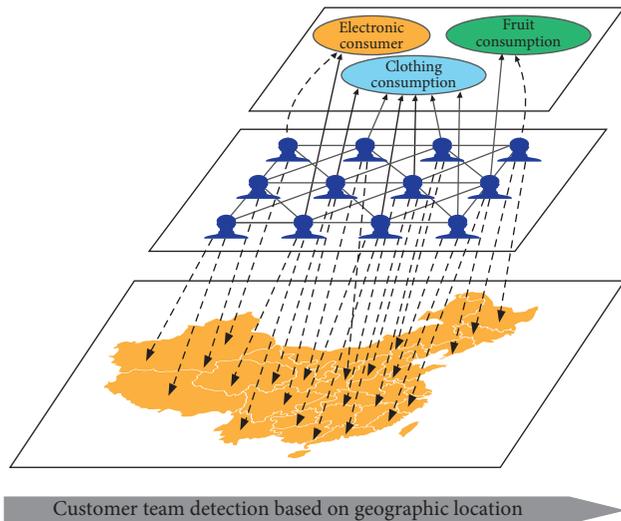


FIGURE 1: Different consumer groups are found on multilayer networks according to the geographical location.

real-world social networks have a stratified structure, which cannot be detected globally without observing network community at different scales. How to detect community very accurately and efficiently using all information tables obtained from the large-scale diverse social networks is a significant challenge. In this context, some works began identifying certain nodes or a local community of several nodes quickly and accurately, given a limited amount of network information. In this way, the prohibitive space and time overhead associated with global calculation can be avoided.

Each user of the social networks has different characteristics, such as the families, circle of friends and classmates, hobbies, and interests. These characteristics are called node trust relationship. The network based on attribute similarity between nodes is called correlation network. The network with node or edge trust relationship is called the attribute network.

In this paper, we jointly consider the trust relationship of all network nodes and the strengths of social similarity. A new metric is defined to describe node attribute and social similarity strength. An algorithm is proposed to detect local community through effective use of the new metric. Experiments are performed to test the proposed algorithm. Unlike the traditional methods which cluster nodes based on topological structure, the proposed algorithm determines node similarity using certain trust relationship of nodes in real life. The relationship between these nodes is then used to obtain the integrated level of similarity. The representative seed node of the community is identified. The community correlated with the attribute is merged by analyzing the similarity between the seed and the community. The strength of social similarity between the attribute and the boundary node is used as a criterion to terminate community expansion. Finally, a scheme based on attribute relationship and social similarity strength is proposed to detect local community within a multilevel network.

The remainder of this paper is organized as follows: in Section 2, the fundamental theory on the detection of

community within multilevel social networks is reviewed. In Section 3, our trust-based method for detection of local community from multilevel network is proposed. In Section 4, the proposed method is evaluated and compared with three other models in real-world multilevel networks and specific QQ Zone marketing data. In Section 5, the conclusion is presented and the future work is recommended.

2. Related Work

The aim of community detection is to segment social network, gain more understanding of social network, identify community members, and find groups with similar ideas, motivations, shopping preferences, and other common characteristics.

2.1. Local Community Detection. The social network is increasingly large, and it contains more information. The method based on local information detects community from a local perspective. By eliminating the need for a global analysis of the network structure, this type of method has become more popular in recent years [10]. In [11], Bagrow and Bollt proposed to start with source node and keep adding continuous shells as the node. In [12], Lancichinetti et al. proposed the F metric for the fitness function to measure the connectivity difference between the inside and outside of community. Their method is easy to implement, but the random selection of the original node usually causes instability in community detection. Moreover, certain parameters of the fitness metric have to be determined in advance. In [13], Chen et al. proposed a method to detect local community using the local-degree central node. In their method, the local community is not identified from the given start node. Instead, it is identified from the local central node correlated with the given start node. In [14], Tabarзад and Hamzeh proposed a heuristic method to detect community by investigating local information. Comparison with other state-of-the-art algorithms demonstrated the ability of their method to detect community and member more effectively and accurately. In [15], Chen et al. proposed a metric called semilocal centrality to find a balance between centrality with a low level of correlation and other time-consuming metrics. In [16, 17], the authors proposed a multiagent algorithm for autonomous community detection from the distributed environment.

The members of a social community usually have multilevel relationships. Most of the traditional community detection algorithms are based on the information of network structure. User behavioral trust relationship is not taken into account in the local community detection methods described above.

2.2. Multilevel Network Community Detection. Detecting community within a multilevel network has drawn a lot of attention in recent years [18–21]. Due to the considerable complexity of the real world, the single-level network is no longer able to describe the community very effectively. In [22], Berlingerio proposed a multilevel network model to

analyze the complex system of the real world and defined the multilevel network relationships. In [23], Gayel et al. proposed a general framework of network quality function, which allows the community in any multilevel network to be studied. In this framework, the network is a combination of coupled links which connects each node of a network slice with each node of other slices. This framework enables us to study the community structure of as many slice networks as we want. In [18], Domenico revealed that any complex system can be represented by a multilevel network. For example, organism genes and the interaction between them can be represented by 7 network layers. In [24, 25], the authors proposed a community detection algorithm based on multilevel network modularity. The concept of modularity was introduced to a broad variety of dynamic and connected networks. In [26, 27], a community detection algorithm based on multilevel network clustering was presented, where the original multilevel networks were first merged into a single-level network under certain strategy before the community detection method for the single-layer network was implemented. In [28–32], a community detection method based on consensus clustering was described. The community detection algorithm for the single-layer network was first implemented on each layer of network. The detection result was then converted into a characteristic matrix of nodes, which was subsequently processed using the traditional clustering algorithm.

The methods described above focus on community detection from a global perspective. It is a new challenge to detect community from the complex and large-scale real-world network systems.

2.3. Similarity Trust. In their paper published in 1998, Buskens [33] derived the assumption of the influence of density, degree centrality, and centralization on the trust degree of client, and according to the test, higher trust can be generated with higher density and higher degree. Sherchan et al. [34] pointed out that the network structure may affect the trust level of social network, and in the network, high density and high interaction between members can generate high-level trust. By utilizing the social network structure and its dynamics, Trifunovic et al. [35] proposed two supplementary methods to build social trust: dominant social trust and implicit social trust. In their article on transactions, Wang et al. [36] obtained the similarity between the same group of neighbours based on the interest of one pair of partners, and they used the two points P_i and P_j to represent two neighbours and define the Jaccard similarity trust relation between two nodes through Jaccard measurement. Jin et al. [37] proposed a trust model based on the evaluation of group similarity; this model can compute the global trust value with similarity between groups as the recommendation reliability, the local trust value, and global trust value can be provided by this model based on the recommended reliability, and they evaluated the model from the perspective of security. In their research, Ziegler and Lausen [38] used trust to supplement or even replace the filtering mechanism, but in order to obtain meaningful result, they

assumed that trust could reflect the user similarity in a certain degree, and they obtained the following conclusion: that is, when the trust network of community is closely integrated with certain application program, there is correlation between trust and user similarity. Boratto et al. [39] proposed the setting to realize trust parameter by computing the average value of top similarity; this parameter represents the minimum value in which the average value of similarity must reach to be recognized as reliable, and the trust parameter obtained from prediction rating must be smaller than 0.85. Among most current works, in [40, 41], the trust degree is generally represented by the user dominance; however, in most online social networks (such as Facebook, Epinions, and Flixster) of real world, there is no specific value to measure the degree of user trust relation, so their paper defined the direct similarity relation between them based on the cosine similarity, and direct trust is represented through similarity relation. Applying existing algorithms to local community detection within a multilevel network has many limitations. For example, global information is not easily available; the node attribute cannot be selected appropriately and utilized effectively; and the model is not as accurate and stable as expected. To address these problems, we propose a model for local community detection within a multilevel network based on trust.

3. Community Detection from Local Network

In this section, we first describe the multilevel network graph model. Next, we propose the E-MLCD model for local community detection from multilevel networks. Finally, the algorithm framework is presented based on the input seed node.

3.1. Problem Formulation. In real life, there is a frequent need to recommend users with appropriate products by identifying their preferences. To achieve this, a core user interested in a certain product is determined and then expanded to find more users who are also interested in the same product. This process is referred to as local community detection, i.e., obtaining a community structure centred on one or more seed users, given limited information on the network. Obviously, the local community detection algorithm is able to effectively identify a community of interest by accessing and manipulating a relatively small part of the network.

After reviewing existing community detection methods, it is learned that no work has been done in the past to jointly consider node similarity strength and node attribute during community detection. The limitations of existing methods can be summarized as follows: (1) detecting community within a multilevel network is very sensitive to the location of start node; (2) information on structural similarity between communities and the topological information are underused; and (3) the expansion direction and breadth within the multilevel network cannot be controlled accurately. In existing methods, the community is expanded outward at a step length of one node. Effective guidance on

expansion is lacking. Therefore, the start node location and community similarity strength should be fully taken into account. The attribute difference between users in different network layers can be used to formulate a new strategy for local community detection within a multilevel network.

3.2. Local Seed Node. In the social network, there is a broad variety of trust, such as age, gender, residence place, shopping preference, and behavioral description. Each of these trust relationship can be seen as an attribute layer. The attribute of the core node in each layer is critical to the node feature description. Therefore, in addition to structure-based social similarity, similarity between nodes in the same layer, which arises from node attribute similarity, should also be taken into account.

Attribute-based node classification can be achieved by classifying nodes with the same attribute into the same category. During local community detection of a network layer, the attribute similarity between node and its neighbouring system in the corresponding structure is defined as the criterion for selecting the seed node. The set of all seeds fulfilling the criterion is determined and then expanded to provide a community.

Let $G_{\mathcal{L}} = (V_{\mathcal{L}}, E_{\mathcal{L}}, V, \mathcal{L})$ denotes the trust-based network model, where $V_{\mathcal{L}}$ denotes the set of nodes in layer \mathcal{L} , V is, $\{V_1, V_2, \dots, V_n\}$ N denotes the number of nodes in the network, and $E_{\mathcal{L}}$ denotes the edges connecting nodes in layer \mathcal{L} .

As the seed node algorithm of the single-layer network cannot be applied to the multilayer network, we describe the core node in the network by defining the local centre degree of the multilayer network. The larger the degree of the node indicates that the more the node is located in the centre of the network, the more important it is in the network and the more it can affect other nodes of the network.

Consider a multitier network with an M layer and N nodes per layer $u = (y, t)$. The node $i \in X (i = 1, \dots, N)$ is expressed as the connection degree or the vector of degree:

$$k_i = (k_i^{[1]}, \dots, k_i^{[M]}), \quad (1)$$

where $k_i^{[1]}$ be the degree of node i in the a layer α . The degree of $k_i^{[\alpha]} = \sum \alpha_j^{[\alpha]}$ is

$$D_i = \sum_{\alpha=1}^M k_i^{[\alpha]}. \quad (2)$$

3.3. Measurement of Intralayer Similarity Trust. During the detection of local community in the multilayer network, it is very important to control the community expansion of network based on the intralayer trust relation of multilayer network and the interlayer trust relation. In this section, we only consider the similarity trust relation when nodes are on the same layer. In order to better express the similarity trust between nodes on the same layer, in this section, the cosine similarity is used to represent the trust relation between nodes on the same layer of network. For an \mathcal{L} -layer network

$G_{\mathcal{L}} = (V_{\mathcal{L}}, E_{\mathcal{L}}, \mathcal{V}, \mathcal{L})$ consisting of the different network layer set \mathcal{L} and node set V , the intralayer trust relation of multilayer network has the following definition.

Definition 1 (intralayer similarity trust relation of multilayer network). For two nodes with cosine similarity on the same layer of multilayer network, according to literatures [42, 43], it is easier for similar nodes to form a community. Obtain the trust measurement of candidate node and target node on the same layer, and the computation method for similarity trust on the same later can be expressed with the following formula:

$$\begin{aligned} T_{\text{int}}^i(C_u, S_v) &= \sum_{u \in C} \sum_{v \in S} \text{sim}_i(C_u, S_v) \\ &= \sum_{u \in C} \sum_{v \in S_v} \frac{|N_i(C_u) \cap N_i(S_v)|}{\sqrt{|N_i(C_u)| |N_i(S_v)|}} \end{aligned} \quad (3)$$

$$\begin{aligned} T_{\text{int}}^i(S_u, S_v) &= \sum_{u \in S} \sum_{v \in S} \text{sim}_i(S_u, S_v) \\ &= \sum_{u \in S} \sum_{v \in S_v} \frac{|N_i(S_u) \cap N_i(S_v)|}{\sqrt{|N_i(S_u)| |N_i(S_v)|}} \end{aligned} \quad (4)$$

In formula (3), $T_{\text{int}}^i(C_u, S_v)$ refers to the trust measurement of node u in local community and node v on the same layer of neighbourhood, in which, C_u is the node in local community C and S_v is the node in neighbourhood set S connected to local community C . $\text{sim}_i(C_u, S_v)$ represents the cosine similarity of two nodes; $N_i(C_u)$ is the neighbour set of node C_u on layer L_i and $N_i(S_v)$ is the neighbour set of node v in set S . In formula (4), $T_{\text{int}}^i(C_u, S_v)$ refers to the trust measurement between nodes S_u, S_v in neighbourhood set S ; $\text{sim}_i(S_u, S_v)$ is the cosine similarity of nodes S_u, S_v in set S .

Because there is no extraattribute, the value of overall trust relation on the same layer is the higher value between $T_{\text{int}}^i(C_u, S_v)$ and $T_{\text{int}}(S_u, S_v)$:

$$T_{\text{int}}(u, v) = \max(T_{\text{int}}^i(C_u, S_v), T_{\text{int}}^i(S_u, S_v)). \quad (5)$$

3.4. Measurement of Interlayer Similarity Trust. In the last section, the similarity trust between nodes on the same layer of multilayer network was defined. However, the nodes on different layers of multilayer network also have different trust relations, and we consider the similarity trust between nodes on different layers as beneficial supplementation to the community detection of the multilayer network.

Definition 2 (interlayer similarity trust relation of multilayer network). For two nodes with cosine similarity on different layers of multilayer network, according to literatures [42, 43], it is easier for similar nodes to form a community. The trust measurement of candidate node and target node on different layers is obtained, and the computation method for interlayer similarity trust can be expressed with the following formula.

According to formula (3), obtain the similarity between nodes u and v on layer L_i and layer L_j . Similar to previous

definition, $\text{sim}_{i,j}(u, v)$ corresponds to the topologic similarity measurement between adjacent sets of u and v , but in this example, they are on different layers:

$$\begin{aligned} T_{\text{out}}^{i,j}(C_u, S_v) &= \sum_{u \in C} \sum_{v \in S} \text{sim}_{i,j}(C_u, S_v) \\ &= \sum_{u \in C} \sum_{v \in S} \frac{|N_i(C_u) \cap N_j(S_v)|}{\sqrt{|N_i(C_u)| |N_j(S_v)|}} \end{aligned} \quad (6)$$

$$\begin{aligned} T_{\text{out}}^{i,j}(S_u, S_v) &= \sum_{u \in S} \sum_{v \in S} \text{sim}_{i,j}(S_u, S_v) \\ &= \sum_{u \in S} \sum_{v \in S} \frac{|N_i(S_u) \cap N_j(S_v)|}{\sqrt{|N_i(S_u)| |N_j(S_v)|}} \end{aligned} \quad (7)$$

In formula (6), $T_{\text{out}}^{i,j}(C_u, S_v)$ refers to the trust measurement of node u in local community and node v in the neighbourhood on different layers, in which C_u is the node in local community C and S_v is the node in neighbourhood set S on different layers connected to local community C . $\text{sim}_i(C_u, S_v)$ represents the cosine similarity of two nodes; $N_i(C_u)$ is the neighbour set of node C_u on layer L_i ; $N_j(S_v)$ is the neighbour set of node v in set S on layer L_j . In formula (7), $T_{\text{out}}^{i,j}(S_u, S_v)$ refers to the trust measurement between nodes S_u, S_v in neighbourhood set S ; $\text{sim}_i(S_u, S_v)$ is the cosine similarity of nodes S_u, S_v in set S . According to formula (6), the value of overall trust relation between the target node and the node to be added on different layers is the higher value between $T_{\text{out}}^{i,j}(C_u, S_v)$ and $T_{\text{out}}^{i,j}(S_u, S_v)$:

$$T_{\text{out}}(u, v) = \max(T_{\text{out}}^{i,j}(C_u, S_v), T_{\text{out}}^{i,j}(S_u, S_v)). \quad (8)$$

Therefore, the seed node of the multilayer network is the node with the largest degree, which is expressed as follows:

$$V_{\downarrow 0} = \max(s_{\downarrow} i^{\uparrow} [1], L, s_{\downarrow} i^{\uparrow} [\mathcal{L}]). \quad (9)$$

3.5. Trust-Based Algorithm for Multilevel Local Community Detection. In this subsection, we propose E-MCLD and a trust-based algorithm for multilevel local community detection. It is valid for any graph with two or more layers. The pseudocode of the general method for trust-based multilevel local community detection is given in Algorithm 1.

The convergence of the proposed algorithm guarantees the dominant relationship between individuals and the upper bound of the two objective functions. According to their definitions, there is an upper bound for the multilevel network to be expanded using the seed node. The community to which the node belongs is determined by the level of similarity of this node with two sets of attribute. Therefore, the proposed algorithm is able to converge. Figure 2 shows the algorithm's trust-based community detection.

4. Experimental Results

The performance of E-MLCD is evaluated through extensive experimentation on single- and multilevel real-world

networks. The experiments are performed on the computer with Windows 7, 3.10 GHz, and 32.00 GB RAM.

4.1. Experimental Datasets. As we know, datasets on social networks with explicit behavioral trust relationship are very rare, and classification trust relationship almost shares the same characteristics as the behavioral trust relationship. Therefore, three real-world networks with classification trust relationship and mobile QQ Zone blog datasets with behavioral trust relationship were used in the experiments to evaluate E-MLCD.

World Championships in Athletics 2013: based on different behavioral trust relationship of this dataset, the network is divided into three layers, i.e., forwarding, mentioning, and replying.

Airline data: different airlines in Europe are defined as attribute. Each layer corresponds to an airline, yielding a network of 37 layers.

Dataset of staff in Aarhus University: five online and offline relationships between university staff are defined as trust relationship, and each attribute is regarded as a relationship. In this way, the network is divided into five layers.

QQ Zone dataset: this dataset was collected from actual QQ Zones of sales personnel. It contains classification of products, as well as gender and age of postissuers, repliers, and praises. The network is divided into three layers, i.e., posting, replying, and praising.

Table 1 shows the main characteristics of the four data we need to use. We use #Nodes to represent the number of multilayer nodes, #Edges to represent the number of edges of multilayer networks, #Layers to represent the number of layers, A_{deg} to represent the average number of nodes considering the average degree of multilateral nodes, and A_{layer} to represent the average number of layers of nodes.

4.2. Evaluation Methods and Metrics

4.2.1. Evaluation Methods. For the purpose of performance evaluation, E-MLCD was compared with three community detection algorithms, inferring multilayer global community structure. In order to compare our E-MLCD method of global community detection, after iterating the algorithm cycle, we finally get a global community structure.

(1) *Louvain (a Modularity-Based Algorithm That Can Detect Community Very Efficiently and Effectively).* In addition, it can detect layered communities. Its optimization objective is to maximize the modularity of the entire graph. On the contrary, Louvain is a nonoverlapping community detection method. By optimizing the value of the modularity function, it allocates each node to the "optimal" cluster, allowing it to process large-scale data efficiently. However, trust relationship and overlapped layers of the network are ignored.

(2) *LCD (a Local Expansion-Based Community Detection Algorithm).* It involves selection of original expansion

Input: Multilayer graph $G_{\mathcal{L}} = (V_{\mathcal{L}}, E_{\mathcal{L}}, V, \mathcal{L})$
 $V_0 \in V_{\mathcal{L}}, E_0 \in E_{\mathcal{L}}, L, N \in G_{\mathcal{L}}$
Output: Local community C for V_0

- 1: $G_{\mathcal{L}} = (V_{\mathcal{L}}, E_{\mathcal{L}}, V, \mathcal{L})$
- 2: $S = V$
- 3: while node $S_i \neq \emptyset$
- 4: v_0 ; // Utilize formula (9) to obtain the seed node
- 5: $S = S - \{v_0\}; C = \{v_0\}$
- 6: for node $v_j \in S$ do
- 7: if $T_{\text{int}}(u, v) \geq T_{\text{out}}(u, v)$ then // Utilize formulas (5) and (8) to determine whether it is intralayer or interlayer expansion
- 8: if $T_{\text{int}}(C_u, S_v) \geq T_{\text{int}}(S_u, S_v)$ then // Utilize formulas (3) and (4) to determine the intralayer nodes to be detected
- 9: $C = C \cup \{v_j\}$ // Combine the nodes to be detected to the local community C
- 10: else $S = S - \{v_j\}$ // Remove the nodes to be detected to set S
- 11: end if
- 12: else // Otherwise, it is interlayer expansion
- 13: if $T_{\text{out}}(C_u, S_v) \geq T_{\text{out}}(S_u, S_v)$ // Utilize formulas (6) and (7) to determine the interlayer nodes to be detected
- 14: $C = C \cup \{v_j\}$ // Combine the nodes to be detected to the local community C
- 15: else $S = S - \{v_j\}$ // Remove the nodes to be detected to set S
- 16: end if
- 17: end if
- 18: $j = j + 1$
- 19: : End for
- 20: $i = i + 1$
- 21: end while
- 22: return C
- 23: end

ALGORITHM 1: Trust-based algorithm for multilevel local community detection.

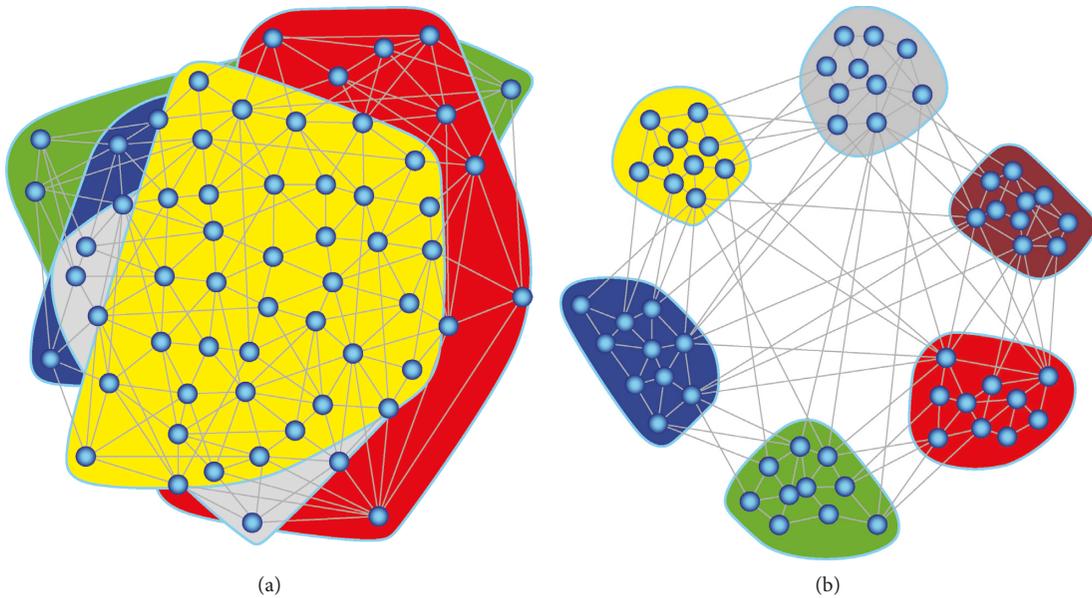


FIGURE 2: Multilevel community detection based on different node trusts.

TABLE 1: Major characteristics of multilevel datasets.

Dataset	#Nodes	#Edges	#Layers	Density	A_{deg}	A_{layer}
MoscowAthletics2013	103319	144591	3	1.68e-3	20.3	1.86
Airlines	417	3588	37	0.023	17.20	4.88
CS-AARHUS	61	620	5	0.122	20.32	3.67
QQ Zone personal selling	820484	16044572	3	2.38e-3	25.56	2.32

subgraphs, expansion strategies, and the conditions for terminating the expansion. In most of the local community detection methods, the expansion process is greedy. In other words, given a fitness function of a local community, the neighbour that can produce the most gains to the fitness is added to the community. The iterative process does not end until no neighbour can improve the local community. Note that LCD is very similar to our method, but LCD is unable to detect community within a multilevel network.

(3) *ML-LCD (a Nonsupervised Method for Community Detection within a Multilevel Network)*. After setting an internal-to-external connectivity ratio, it achieves local community detection by applying ML-LCD-lw, ML-LCD-wlsim, and ML-LCD-clsim to different layers. It supports layer-weight similarity and intralayer and interlayer similarities. This method achieves community detection by comparing connectivity but fails to take into account the level of similarity between communities, resulting in limited performance in the face of communities with many similar trust relationships.

4.2.2. *Evaluation Metrics*. The metrics used to evaluate E-MLCD include the scale of detected community, the scale of multilevel community, and the computational complexity.

(1) *Multilayer Modularity*. Algorithm performance is measured by the community scale. Table 2 shows the detection results of E-MLCD and the other algorithms.

It can be seen that E-MLCD produced the most communities for all datasets except Airlines. Also, the proposed algorithm was more effective in detecting community from dataset with behavioral trust relationship. Consider Moscow Athletics 2013 and QQ Zone datasets, which have obvious behavioral trust relationship such as replying and forwarding. The proposed algorithm achieved performance gain by categorizing these behavioral trust relationships into the same layer. In particular, a large number of attribute information such as gender, geographical location, and age were collected, and more communities were detected in the QQ area. But in the case of the Airlines dataset which has fewer trust relationship, our method was slightly inferior to ML-LCD.

(2) *Average Multilevel Module Club Evaluation*. In the multilayer network, it is impossible to use the network module degree to realize the evaluation, so the multilayer network module degree is used to evaluate the algorithm in this part. In a multitier network, the higher the Q , the better the result of the division of the community.

In the multilayer network, in addition to the largest community, each test result can also reflect the algorithm deviation. Figure 3 shows the size of the module degree obtained by the results of 50 runs by different algorithms under different datasets. Since the LART algorithm cannot display correct results in a large-scale datasets, only GL, PMM, and E-MLCD algorithms are shown in Figure 3.

According to the graphic analysis, in MoscowAthletics2013, the E-MLCD algorithm could obtain great

TABLE 2: Comparison of multilayer modularity for different algorithms.

Dataset	GL	LART	PMM	E-MLCD
MoscowAthletics2013	0.025	NA	0.630	0.680
Airlines	0.037	0.013	0.018	0.023
CS-AARHUS	0.249	0.154	0.222	0.261
QQ Zone personal selling	0.021	NA	0.580	0.620

results, and among 50 operations, most of them achieved great modularity. This is consistent with the results of the maximum modularity in Table 2. According to Table 1, we can see that A_{deg} of the MoscowAthletics2013 dataset was relatively low because the dataset was relatively sparse with large data volume. Therefore, we can see that the algorithm proposed in this section can obtain great modularity in a large-scale large network. In the BioProte dataset, MLCD could obtain similar results as the classic algorithms. In the dataset CS-AARHUS, the GL algorithm had stable performance, and its results were also higher than the results of the E-MLCD algorithm. Therefore, it can be seen that the E-MLCD algorithm does not have advantages in the analysis of small datasets. In the Airlines dataset, the results of our algorithm were lower than the results of the PMM algorithm and slightly lower than the results of the GL algorithm. This is mainly because the Airlines dataset had high A_{deg} and low A_{leyas} , and our algorithm had less trust relationship. The special case was the QQ Zone dataset. The GL and PMM algorithms obtained stable test results, and at most circumstances, the results of the GL algorithm were higher than the results of the E-MLCD algorithm. In the QQ Zone dataset, the algorithm proposed by us obtained high modularity for 6 times.

According to the comparison analysis, our algorithm could obtain good detection results in large-scale community and network with the sparse dataset. However, in closely connected network with small dataset, the detection result of our algorithm was poorer than that of the GL algorithm, and better than that of the PMM algorithm. This is mainly because the PMM algorithm needs to set corresponding prediction parameter in advance, and it cannot accurately detect community in large-scale unknown multilayer network. In the dataset with close network connection, the E-MLCD algorithm had lower or equal performance as the GL algorithm, and this is mainly because our algorithm mainly depends on attribute similarity and social intensity similarity to find community. In the dataset with close connection, its performance is insufficient compared to the GL algorithm.

(3) *Computation Efficiency*. In order to test the time complexity of different algorithms, in this chapter, we choose the dataset in which the E-MLCD algorithm could achieve good results: the QQ Zone dataset. Through comparison, we found that the QQ Zone dataset had a large number of nodes and low A_{deg} , so the time performance of various methods in this dataset was several orders of magnitude higher than that in other datasets, and the average running time was 500 minutes. The running time of other datasets was much

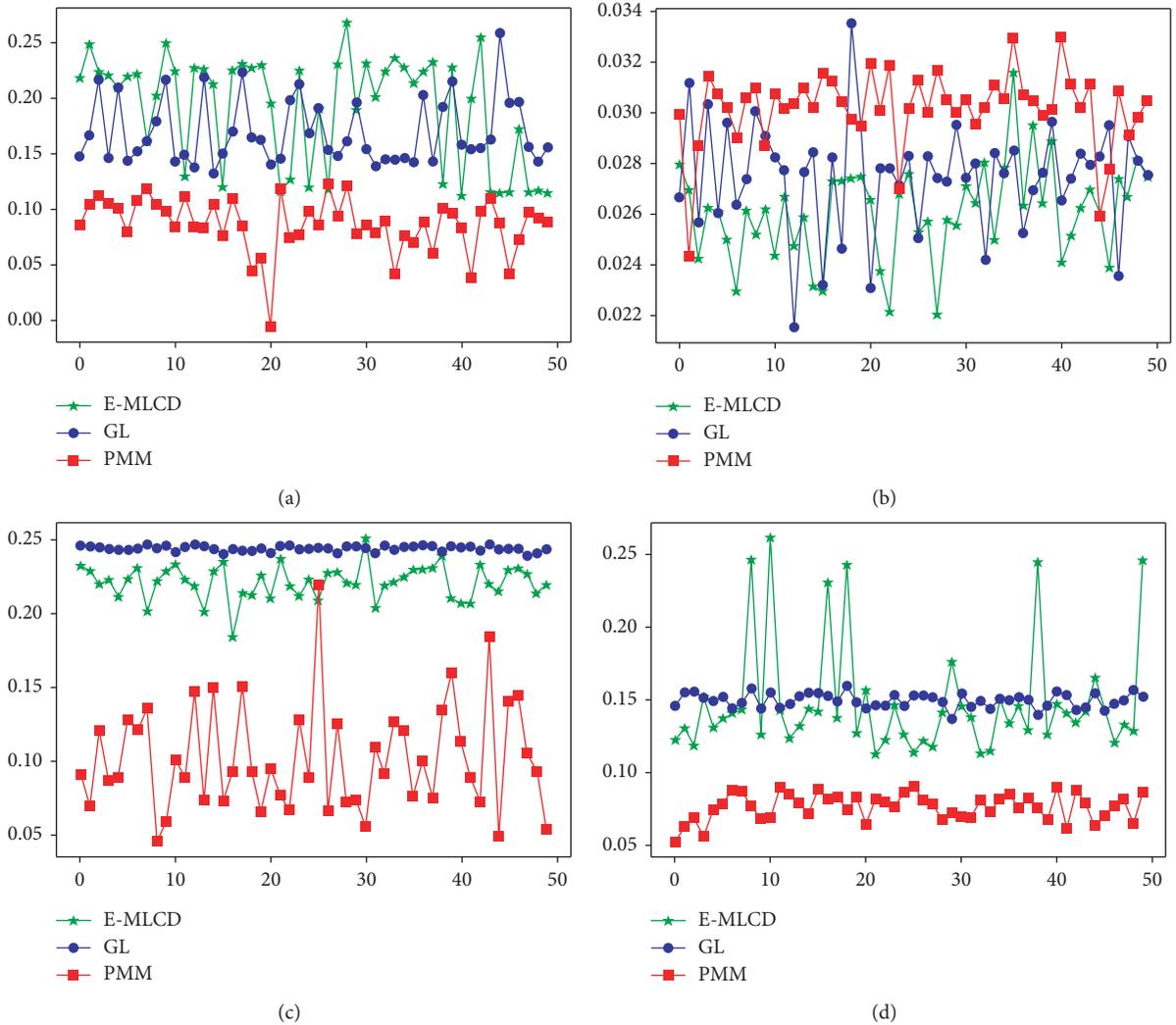


FIGURE 3: Average modularity of different algorithms. (a) MoscowAthletics2013. (b) Airlines. (c) CS-AARHUS. (d) QQ Zone.

shorter, especially the CS-AARHUS dataset, which was measured in second, and it could not well reflect the time performance of algorithm. Therefore, during the comparison of various algorithms, the operating efficiency on sparse dataset QQ Zone was used as the benchmark for comparison. In this paper, different algorithms were used for analysis in the QQ Zone dataset, and Figure 4 shows the final test results. We can see that, for all 3 algorithms, the running time presented direct-proportion distribution with the network scale; however, the running time of E-MLCD was significantly shorter than that of the other 2 algorithms, the running time of PMM algorithm was shorter than that of the GL algorithm, and during several operations, the running time of PMM was also shorter than that of our algorithm.

As shown in Figure 4, the E-MLCD algorithm can be realized in short time in most operations, it has better time performance than the GL and PMM algorithms, and its stability was also better than that of the GL and PMM algorithms. The overall stability of PMM algorithm was better than that of the GL algorithm, and it had similar time performance to our algorithm. The GL algorithm had the

longest operating time, and it has poorer stability than the E-MLCD and PMM algorithms. According to Figure 4, in the QQ Zone dataset, our algorithm could obtain better modularity than the GL and PMM algorithms. Therefore, our algorithm can efficiently process the large-scale sparse network of multiple layers and trust relationship.

5. Conclusions

Detecting user groups of a particular interest from the social network has been the ultimate goal of social network advertisements. This is because targeted advertising can considerably increase its effectiveness and maximize profits. These groups, also known as communities, open up opportunities for a new marketing pattern based on acquaintance circle and six-dimensional space. By clustering users with common interests into customer groups, detecting closely interrelated network users is essential for improving advertisement prompting via social media and extracting potential customers from social network. Based on three real-world social networks and QQ Zone marketing

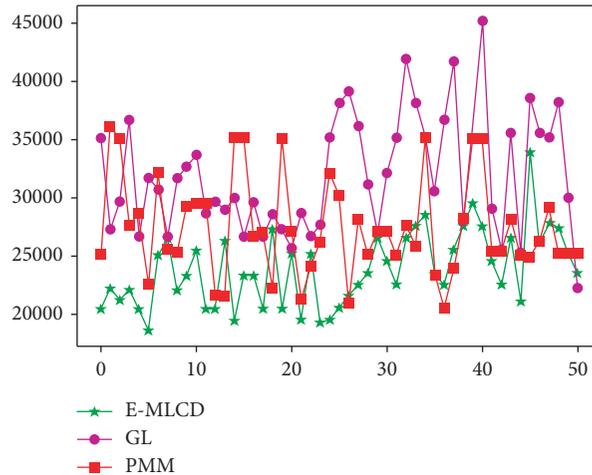


FIGURE 4: Running time of different algorithms in the QQ Zone dataset after running for 50 times.

data, we proposed a new local community detection model E-MLCD, which jointly considers multilevel trust relationship and community structure. To address the problem of expanding local community during the detection process, this chapter proposed the local community detection model based on multilayer trust relationship and community structure (E-MLCD) for the first time. For the local community detection and expansion problem, this model defined new measurement with similar community intensity based on similarity of community structure. The E-MLCD method can fully utilize the structure and attribute information to realize network partition and promote application of multilayer network, such as obtaining better detection in sparse network based on partial attribute and realizing comparison with popular and similar algorithms; the E-MLCD algorithm has advantages in the analysis of a large-scale multilayer network with sparse connection, and it can effectively identify sparse community structure in a large-scale multilayer network and obtain better time performance.

Data Availability

MoscowAthletics2013 datasets [9], Airlines datasets [40], and CS-AARHUS datasets [7] are cited at relevant places within the text as references. The QQ Zone personal selling data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

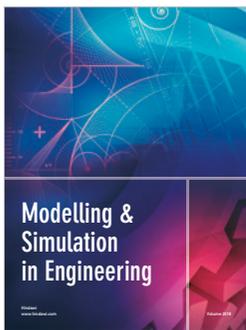
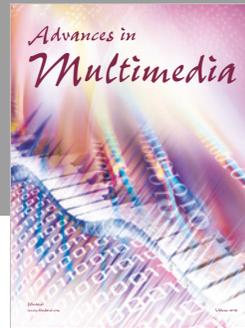
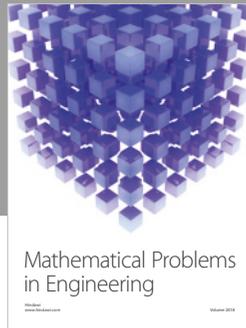
Acknowledgments

This research is partially supported by the Fundamental Research of Xinjiang Corps (2016AC015) and the Applied Basic Research Project of Qinghai Province (No. 2018-ZJ-707).

References

- [1] M. Kivela, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, "Multilayer networks," *Journal of Complex Networks*, vol. 2, no. 3, pp. 203–271, 2014.
- [2] C. W. Loe and H. J. Jensen, "Comparison of communities detection algorithms for multiplex," *Physica A: Statistical Mechanics and its Applications*, vol. 431, pp. 29–45, 2015.
- [3] J. Kim and J.-G. Lee, "Community detection in multi-layer graphs: a survey," *ACM SIGMOD Record*, vol. 44, no. 3, pp. 37–48, 2015.
- [4] G. D'Agostino and A. Scala, *Networks of Networks: the Last Frontier of Complexity*, Vol. 340, Springer, Berlin, Germany, 2014.
- [5] D. Cai, Z. Shao, X. He, X. Yan, and J. Han, "Community mining from multi-relational networks," in *European Conference on Principles of Data Mining and Knowledge Discovery*, pp. 445–452, Springer, Porto, Portugal, 2015.
- [6] M. Berlingerio, F. Pinelli, and F. Calabrese, "ABACUS: frequent pAttern mining-based community discovery in multidimensional networks," *Data Mining and Knowledge Discovery*, vol. 27, no. 3, pp. 294–320, 2013.
- [7] M. Magnani, B. Micenkova, and L. Rossi, "Combinatorial analysis of multiple networks," *Computer Science*, 2013, <https://arxiv.org/abs/1303.4986>.
- [8] M. De Domenico, V. Nicosia, A. Arenas, and V. Latora, "Structural reducibility of multilayer networks," *Nature Communications*, vol. 6, no. 1, p. 6864, 2015.
- [9] W. Wang, X. Li, P. Jiao et al., "Exploring intracity taxi mobility during the holidays for location-based marketing," *Mobile Information Systems*, vol. 2017, Article ID 6310827, 10 pages, 2017.
- [10] A. Clauset, "Finding local community structure in networks," *Physical Review E*, vol. 72, no. 2, article 026132, 2005.
- [11] J. P. Bagrow and E. M. Bollt, "Local method for detecting communities," *Physical Review E*, vol. 72, no. 4, article 046108, 2005.
- [12] A. Lancichinetti, S. Fortunato, and J. Kertész, "Detecting the overlapping and hierarchical community structure in complex networks," *New Journal of Physics*, vol. 11, no. 3, article 033015, 2009.
- [13] Q. Chen, T.-T. Wu, and M. Fang, "Detecting local community structures in complex networks based on local degree central

- nodes," *Physica A: Statistical Mechanics and its Applications*, vol. 392, no. 3, pp. 529–537, 2013.
- [14] M. A. Tabarzad and A. Hamzeh, "A heuristic local community detection method (HLCD)," *Applied Intelligence*, vol. 46, no. 1, pp. 62–78, 2016.
- [15] D. Chen, L. Lü, M.-S. Lü, Y.-C. Zhang, and T. Zhou, "Identifying influential nodes in complex networks," *Physica A: Statistical Mechanics and its Applications*, vol. 391, no. 4, pp. 1777–1787, 2012.
- [16] P. Wang and J. Liu, "A multi-agent genetic algorithm for local community detection by extending the tightest nodes," in *Proceedings of 2016 IEEE Congress on Evolutionary Computation*, pp. 3215–3221, Vancouver, BC, Canada, July 2016.
- [17] Z. Bu, Z. Wu, J. Cao, and Y. Jiang, "Local community mining on distributed and dynamic networks from a multiagent perspective," *IEEE Transactions on Cybernetics*, vol. 46, no. 4, pp. 986–999, 2016.
- [18] X. Li, Q. Tian, M. Tang, X. Chen, and X. Yang, "Local community detection for multi-layer mobile network based on the trust relation," *Wireless Networks*, 2019, In press.
- [19] X. M. Li, L. Yuan, C. C. Liu et al., "An efficient critical incident propagation model for social networks based on trust factor," in *Proceedings of International Conference on Collaborative Computing: Networking, Applications and Worksharing*, pp. 416–424, Springer, Cham, Switzerland, December 2017.
- [20] M. De Domenico, A. Lancichinetti, A. Arenas et al., "Identifying modular flows on multilayer networks reveals highly overlapping organization in interconnected systems," *Physical Review X*, vol. 5, no. 1, p. 011027, 2015.
- [21] R. Lambiotte, J.-C. Delvenne, and M. Barahona, "Random walks, markov processes and the multiscale modular organization of complex networks," *IEEE Transactions on Network Science and Engineering*, vol. 1, no. 2, pp. 76–90, 2014.
- [22] V. Carchiolo, A. Longheu, M. Malgeri et al., "Communities unfolding in multislice networks," in *Complex Networks*, pp. 187–195, Springer, Berlin, Heidelberg, Germany, 2011.
- [23] I. Gaye, G. Mendy, S. Ouya, I. Diop, and D. Seck, "Multi-diffusion degree centrality measure to maximize the influence spread in the multilayer social networks," in *Proceedings of International Conference on e-Infrastructure and e-Services for Developing Countries*, pp. 53–65, Ouagadougou, Burkina Faso, West Africa, December, 2016.
- [24] Q. Han, K. Xu, and E. Airoldi, "Consistent estimation of dynamic and multi-layer block models," in *Proceedings of the 32nd International Conference on Machine Learning (ICML-15)*, pp. 1511–1520, Lille, France, July 2015.
- [25] M. Porter, "Community structure in time-dependent networks," in *Proceedings of APS March Meeting*, Portland, Oregon, March 2010.
- [26] L. Tang, X. Wang, and H. Liu, "Community detection via heterogeneous interaction analysis," *Data Mining and Knowledge Discovery*, vol. 25, no. 1, pp. 1–33, 2011.
- [27] G. Zhu and K. Li, "A unified model for community detection of multiplex networks," in *Proceedings of International Conference on Web Information Systems Engineering*, pp. 31–46, Thessaloniki, Greece, October 2014.
- [28] P. Bródka, T. Filipowski, and P. Kazienko, "An introduction to community detection in multi-layered social network," in *Proceedings of World Summit on Knowledge Society*, pp. 185–190, Mykonos, Greece, September 2011.
- [29] X. Li, G. Xu, and M. Tang, "Community detection for multi-layer social network based on local random walk," *Journal of Visual Communication and Image Representation*, vol. 57, pp. 91–98, 2018.
- [30] Z. Yu, H.-S. Wong, and H. Wang, "Graph-based consensus clustering for class discovery from gene expression data," *Bioinformatics*, vol. 23, no. 21, pp. 2888–2896, 2007.
- [31] V. Filkov and S. Skiena, "Integrating microarray data by consensus clustering," *International Journal on Artificial Intelligence Tools*, vol. 13, no. 4, pp. 863–880, 2004.
- [32] F. Saeed, N. Salim, and A. Abdo, "Combining multiple clusterings of chemical structures using cluster-based similarity partitioning algorithm," *International Journal of Computational Biology and Drug Design*, vol. 7, no. 1, pp. 31–44, 2014.
- [33] V. Buskens, "The social structure of trust," *Social Networks*, vol. 20, no. 3, pp. 265–289, 1998.
- [34] W. Sherchan, S. Nepal, and C. Paris, "A survey of trust in social networks," *ACM Computing Surveys (CSUR)*, vol. 45, no. 4, pp. 1–33, 2013.
- [35] S. Trifunovic, F. Legendre, and C. Anastasiades, "Social trust in opportunistic networks," in *Proceedings of INFOCOM IEEE Conference on Computer Communications Workshops, 2010*, pp. 1–6, San Diego, CA, USA, March 2010.
- [36] G. Wang, F. Musau, S. Guo, and M. B. Abdullahi, "Neighbor similarity trust against sybil attack in P2P e-commerce," *IEEE transactions on parallel and distributed systems*, vol. 26, no. 3, pp. 824–833, 2015.
- [37] Y. Jin, Y. Zhang, W. Qu et al., "A trust model based on similarity evaluation in P2P networks," in *Proceedings of IEEE International Symposium on Parallel and Distributed Processing with Applications*, pp. 737–742, IEEE Computer Society, Sydney, NSW, Australia, December 2008.
- [38] C. N. Ziegler and G. Lausen, "Analyzing correlation between trust and user similarity in online communities," in *Proceedings of International Conference on Trust Management*, pp. 251–265, Springer, Berlin, Heidelberg, June 2004.
- [39] L. Boratto, S. Carta, A. Chessa et al., "Group recommendation with automatic identification of users communities, web intelligence and intelligent agent technologies," in *Proceedings of WI-IAT'09 2009 IEEE/WIC/ACM International Joint Conference*, vol. 3, pp. 547–550, Milano, Italy, September 2009.
- [40] G. Xu, Z. Feng, H. Wu, and D. Zhao, "Swift trust in a virtual temporary system: a model based on the dempster-shafer theory of belief functions," *International Journal of Electronic Commerce*, vol. 12, no. 1, pp. 93–126, 2014.
- [41] S. Deng, L. Huang, G. Xu et al., "On deep learning for trust-aware recommendations in social networks," *IEEE Transactions on Neural Networks & Learning Systems*, vol. 28, no. 5, pp. 1164–1177, 2016.
- [42] X. Chen, C. Xia, and J. Wang, "A novel trust-based community detection algorithm used in social networks," *Chaos, Solitons & Fractals*, vol. 108, pp. 57–65, 2018.
- [43] V. Nicosia and V. Latora, "Measuring and modeling correlations in multiplex networks," *Physical Review E Statistical Nonlinear & Soft Matter Physics*, vol. 92, no. 3, article 032805, 2015.




Hindawi

Submit your manuscripts at
www.hindawi.com

