The Computer Journal Advance Access published June 5, 2014

doi:10.1093/comjnl/bxu035

A Formal Analytic Approach to Credible Potential Path and Mining Algorithms for Multimedia Social Networks

ZHIYONG ZHANG[1,2,*] AND KANLIANG WANG[3]

[1]*Department of Computer, Information Engineering College, Henan University of Science and Technology, Luoyang 471023, People's Republic of China*
[2]*School of Management, Xi'an Jiaotong University, Xi'an 710049, People's Republic of China*
[3]*School of Business, Renmin University of China, Beijing 100872, People's Republic of China*
[*]*Corresponding author: z.zhang@ieee.org*

Multimedia social networks (MSNs) services and tools provide a convenient platform for users to share multimedia contents, such as electronic book, digital image, audio and video, with each other. However, in an open network, uncontrolled sharing and transmission mode of digital content between users create considerable problems regarding digital rights management (DRM). This paper aims to explore potential paths on the propagation of copyrighted contents. An approach to mining credible potential paths is proposed for MSNs. The formal descriptions were primarily based on rough set theory for mining potential paths. Trust was also measured to find credible potential paths. We presented related algorithms for mining two kinds of paths between any two nodes. Finally, we conducted an experiment based on three non-overlapped sharing communities multiplied by 150 nodes. In the communities found by using a representative real-world MSN YouTube dataset, we further mine the general and credible potential paths based on the simulated trust assessment values. The proposed method could effectively and accurately mine two kinds of potential paths of copyrighted digital content distribution and sharing, which can help to resolve critical DRM issues.

Keywords: multimedia social networks; digital rights management; rough set; security; algorithm

Received 14 September 2013; revised 4 April 2014
Handling editor: Dell Zhang

1. INTRODUCTION

With the rapid development of network socialization, multimedia social networks (MSNs) have increasingly emerged. These MSNs offer network tools, services and applications for multimedia contents (e.g. electronic book, digital image, audio and video) that can be shared among users in the same group or between different groups within social networks. Many MSNs are popular nowadays, such as video sharing networks YouTube and audio sharing network SongTaste, etc. These networks facilitate the sharing of multimedia content between users because they offer direct, rapid and flexible transmission of such contents. For example, the music video of the song 'Gangnam Style' became the most popular video on YouTube in 2012. The music video has been viewed more than 900 million times and has received more than 3 million 'likes'. However, easy digital content reproduction and the convenient distribution and sharing of such contents facilitated by open MSN environments have enabled people to share and distribute valuable copyrighted digital contents within social networks. Therefore, digital rights management (DRM) has become a more serious matter.

Recent studies on digital content security and copyright protection [1] have focused on cryptology, access control, trusted computing and digital watermarking [2, 3]. MSN is a new scenario characterized by a small world in which multimedia contents or rights, including copyrighted digital assets, are freely distributed much more efficiently. Moreover, the relationship between users is characterized by rapid transitivity, dynamics and the strengthening of weak ties, and immeasurably extends the range of multimedia content/rights propagation within the network. The explosive growth of user scale and content-sharing behavior in this network has made the rights management of content dissemination more difficult. Especially, direct and indirect relationships exist between users in any MSN. Direct relations can be observed easily, but

indirect ones cannot because they are relatively concealed. Unfortunately, these indirect relationships determine plenty of possible but easily ignored distribution and spread paths of digital contents among users enthusiastic about free shares. The potential spreads result in much more considerable and significant risks for copyrighted contents than the explicit distributions. These existing methods for digital contents security and usage control have not yet met the requirements of DRM in MSNs.

Therefore, we aim at finding and mining the relationship-based implicit or potential propagation path among users, and it is an important aspect of the original motivation of DRM application in the context of MSNs. Based on the structural property of MSNs, the rough set theory of soft computing is adopted to estimate the rough path or potential propagation path. This paper identifies the credible potential path by calculating the trust value of the potential path. To the best of our knowledge, no existing researches on implicit and potential path mining address the question of DRM in MSNs, and only some studies of the explicit path discovery have ever been discussed based on the relationships among the users for social networks. So, the main contributions of the paper include a formal analytic approach to (credible) potential paths based on the rough path, especially on the proposed mining algorithms.

The remainder of this paper is organized as follows. Section 2 introduces the community structure property of social networks and covers related work on DRM for social networks. Section 3 presents and describes a rough set-based approach to mining potential paths in MSNs. Section 4 describes an approach to mining credible potential paths. Section 5 provides the algorithms for mining potential paths and credible potential paths. Section 6 presents a simulation experiment by using the proposed algorithms and analyses the results. Section 7 concludes the paper and outlines future work.

2. BACKGROUND AND RELATED WORKS

2.1. Community structure in social networks

A social network is a relatively stable relational network established by special groups based on a certain relationship. These groups can be among individuals, organizations, enterprises or even countries. Moreover, relationships in the social network can be established by real-life acquaintances (e.g. friends, classmates, colleagues or relatives) or strangers who met in the network (e.g. through friends' recommendations or common interests) [4]. Based on certain established relationships in a social network, users can exchange information and share resources in real time. To formally study the characteristics of a social network, the researchers model the social network as a diagram based on the relationship between people. Generally, a social network is computationally represented by a node-edge undirected graph. That is, each user is regarded as one node in the graph, and the relationship between users is represented by an edge. Most studies on social network analysis adopt binary relationship representation.

Social networks have a community structure [5]. Community structure is an inherent basic structure of a social network that reflects the user distribution, node source and other inherent properties. Furthermore, a social network consists of several locally dense communities. In social networks, each community represents an actual social organization formed on the basis of social relationships or interests. Thus, a more obvious division exists between a community and other nodes outside the community. That is, the node–node connection within a community is relatively dense, but connections between communities are very loose. Granovetter [6] proposed the concept of the 'strength of weak ties' and suggested that the spread of information within a group of people with strong relations might be limited to a smaller range and does not have a significant value for users in the community. In contrast, weak relations tend to transfer non-recurring information among different communities. Therefore, more alternation occurs and more information is spread between communities through weak relations, thereby making such relations 'information bridges'.

2.2. DRM for social networks

Researchers have extensively studied DRM issues in the social network platform. Barghin *et al.* [7] suggested that trust between users of online social networks (OSNs) should be a key parameter for access control decisions in OSNs, proposing a combined access control and trust negotiation framework for the protection of OSN resources. Wang and Sun [8] observed that traditional access control mechanisms are conducted in a data management system but are unsuitable for collaborative open social networks. Thus, Wang and Sun proposed a trust-based access control framework, which, unlike traditional methods, can provide full support for expressing highly complex privacy-related policies. As for content access control in MSNs, Sachan *et al.* [9] found that traditional access control mechanisms may not scale well for fine-grained access control requirement and a large number of users. Thus, Sachan *et al.* proposed a fine-grained multimedia access control model based on a bit-vector transform as well as verified the safety, storage and efficiency of the application through simulation experiments. Park *et al.* [10] proposed the development of an access control framework for OSNs based on user activity. This access control framework goes beyond traditional access control in that it enables users to control general usage activity and such special characteristics as attributes, policies, relationships and session. This framework can support relationship-based as well as general attribute-based access control. In connection with the copyright protection of media contents in MSNs and the illegal distribution control of media content over the Internet, Lian *et al.* [11] introduced a robust content distribution and copyright verification system based on a media index and watermarking technology. Chung and Ko [12] proposed a new video-matching algorithm for the

copyright protection of YouTube videos and other MSNs, and designed an intelligent copyright protection system based on this algorithm. This proposed algorithm can effectively match videos, and the system can be used in copyright protection within video sharing networks.

Aside from the above-mentioned studies, the risk of content information dissemination and sharing in social networks has been investigated. To address the problem of unauthorized propagation of information in social networks, Carminati *et al.* [13] introduced a probability-based approach to evaluate unauthorized access risk as well as accurately quantified the probability that user content may become accessible to another user of the OSN. The practicability of the approach was verified through experiments. As for information leakage and access control risks in social networks, Wang *et al.* [14] proposed a statistical risk assessment method. This approach quantifies a threat in the network and assesses information flows between two users in a social information network.

2.3. Path discovery and control in social networks

The social network relationship graph enables two individuals to find the relationship paths that connect them. The relationship path discovery mechanisms can provide a basis for access control mechanisms suitable for social networks, where users determine and control the authorized users based on their distance to themselves in the social network.

For instance, Mezzour *et al.* [15] provided techniques to discover paths between individuals for social network, proposing a privacy-preserving multi-top relationship path discovery mechanism. This proposed algorithm can operate offline during the path discovery phase. Finally, the algorithm has been simulated on real social network topologies. In addition, Xue *et al.* [16] address relationship privacy of path discovery by proposing a Privacy-Preserving Path Discovery protocol (P^3D) for social networks, and present the security analysis of P^3D, in combination with showing its robustness against the main security threats. BrÓdka *et al.* [17] focus on analyzing the shortest paths discovery in multi-layered social networks. Based on Dijkstra and FloydWarshall algorithms, BrÓdka proposed two separate algorithms for evaluation of shortest paths.

To more effectively solve the DRM problem in MSNs, a novel study should be made from a new perspective on seeking potential propagation paths among users and finding a way to control the malicious dissemination of digital contents in such paths of MSNs.

3. ROUGH SET-BASED APPROACH FOR MINING POTENTIAL PATHS IN MSNs

The topological structure of real social networks shows community features; as previously mentioned, a social network consists of several communities. The nodes within a community are very closely connected, whereas connections between communities are very loose. Therefore, potential relationships and paths between different users can be identified according to the weaker connections between communities and based on the binary relation rough method in the rough set.

3.1. Rough set

The rough set theory was first proposed in 1982 by Zdzisław Pawlak, a Polish scientist. Based on set theory, rough set is a mathematical tool for processing incomplete information. By using the rough set theory, we can find implicit knowledge and uncover potential rules. Some basic definitions related to the rough set are given below.

DEFINITION 1 (Approximation space) [18]. *In this definition, U indicates a non-empty finite set of study objects, called the universe of discourse; $R \subseteq U \times U$ is the aggregation of equivalence relations on U, and $G = (U, R)$ is called the approximation space. If $S \subseteq U \times U$ is a binary relation different from R on U, where S may be a common binary, fuzzy, compatibility or partial order relation, then S is called a binary relation in approximation space $G = (U, R)$.*

DEFINITION 2 (Indiscernibility relation) [18]. *The equivalence relation R divides U into a series of disjoint subsets, represented by $U/R = \{[U_1], [U_2], \ldots, [U_n]\}$, where $[U_1]$ denotes an equivalence class of U. If two objects u, v belong to the same equivalence class $[U_i]$, u and v are said to have an indiscernibility relation.*

3.2. Formal description of potential paths in MSNs

The MSN is a dynamic network. New entities and links continuously join MSNs, while the old ones disappear over time. This paper provides a snapshot of a certain period in the growth of MSNs. First, we divide users into different communities based on community structure. And then, a formal description of potential paths is presented by using the binary relation rough method in the rough set.

DEFINITION 3 (MSN approximation space). *In MSNs, the universe of discourse V is a set of all nodes in the small-world network, and V is divided into different communities based on the equivalence relation R (community structure feature) to form an approximate MSN space $G = (V, R)$. By $V/R = \{[V_1], [V_2], \ldots, [V_n]\}$ is denoted the division of V on R, where $[V_i](1 \leq i \leq n)$ is called the equivalence community class in the MSN approximate space. The relationship between nodes from the same equivalence community class is called an equivalence relation. The connected edge of any node in the same community is called an equivalent edge.*

DEFINITION 4 (Rough relation). *In an MSN approximate space $G = (V, R)$, $V/R = \{[V_1], [V_2], \ldots, [V_n]\}$ is the division of V on R, and S is a direct binary relation different from R. As for any two equivalence community classes $[V_i], [V_j] (1 \leq i \leq n, 1 \leq j \leq n$ and $i \neq j)$, and $\forall a \in [V_i], \forall b \in [V_j]$, if $<a, b> \in S$, then $S* = \{<[V_i], [V_j]>\}$ is considered as the rough relation of S in an MSN approximate space $G = (V, R)$, where the weak connection edge of the connection between different equivalence classes formed by relation S is called the bridge edge and the two endpoints that connect a bridge edge are called bridge nodes.*

DEFINITION 5 (Relation matrix of S^*). *In an MSN approximate space $G = (V, R)$, where $V/R = \{[V_1], [V_2], \ldots, [V_n]\}$, S^* is the rough relation of S. Thus, a rough relation matrix between equivalence community classes is established, described as an n^*n matrix $M_{S^*} = (w_{ij})$. The definition is shown in the following equation:*

$$w_{ij} = \begin{cases} 1 <[V_i], [V_j]> \in S^*, \\ 0 <[V_i], [V_j]> \notin S^*. \end{cases} \quad (1)$$

Here, M_{S^*} is the relation matrix of S^*, which reflects the existence of $S*$between any two equivalence classes.

DEFINITION 6 (S^* path). *In an MSN approximate space $G = (V, R)$, where $V/R = \{[V_1], [V_2], \ldots, [V_n]\}$, S is a direct binary relation different from R and S^* is the rough relation of S. If $S^* = \{<[V_1], [V_2]>, <[V_2], [V_3]>, \ldots, <[V_{n-1}], [V_n]>\}$, then $<[V_1], [V_2]>, <[V_2], [V_3]>, \ldots, <[V_{n-1}], [V_n]>$ is considered the S^* path from $[V_1]$ to $[V_n]$.*

DEFINITION 7 (Potential path (PP)). *In an MSN approximate space $G = (V, R)$, where $V/R = \{[V_1], [V_2], \ldots, [V_n]\}$, S^* is the rough relation of S. If $<[c], [V_1]>, <[V_1], [V_2]>, \ldots, <[V_{n-1}], [V_n]>, <[V_n], [d]>$ is the $S*$ path from $[c]$ to $[d]$, for any $a \in [c]$, $b \in [d]$ and $v_{i1}, v_{i2}, \ldots, v_{ix}, \ldots, v_{im} \in [V_i] (i = 1, 2, \ldots, n)$, $<a, v_{1x}>, <v_{1x}, v_{2x}>, \ldots, <x_{(i-1)x}, x_{ix}>, \ldots, <v_{nx}, b>$ is called the potential path on S from a to b. The potential path from a to b is expressed as $PP_{a \to b}$. The ordered pair $<x_{(i-1)x}, x_{ix}>$ in the potential path is called the rough edge ($<x_{(i-1)x}, x_{ix}> \notin S$). Thus, all edges in the potential path are composed of rough edges.*

For a social network, U represents a set of all user nodes, and the social network is divided by R into the different non-overlapped communities based on the common exclusive factors or their combinations like sex, geographic situation, affiliation and department, etc. Further, we would adopt the equivalence relation to find and mine the potential paths. For example, according to the equivalence relation, an MSN includes three equivalence community classes, namely, $[V_i]$, $[V_j]$ and $[V_k]$, and $S = \{<v_{ia}, v_{jc}>, <v_{jc}, v_{ke}>\}$ (as shown Fig. 1). According to Definition 4, $S^* = \{<[v_i], [v_j]>, <[v_j], [v_k]>\}$ is

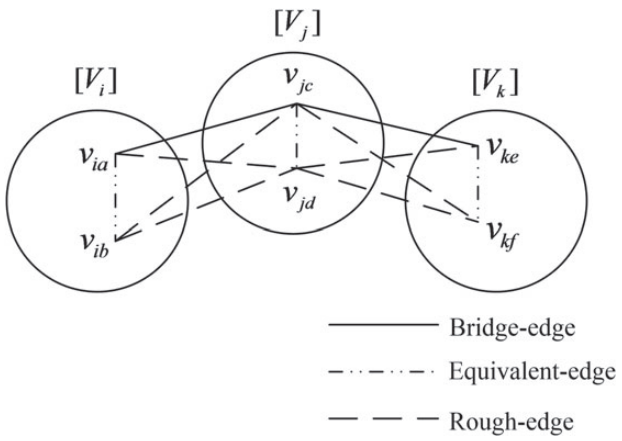

FIGURE 1. Potential paths between non-overlapped communities.

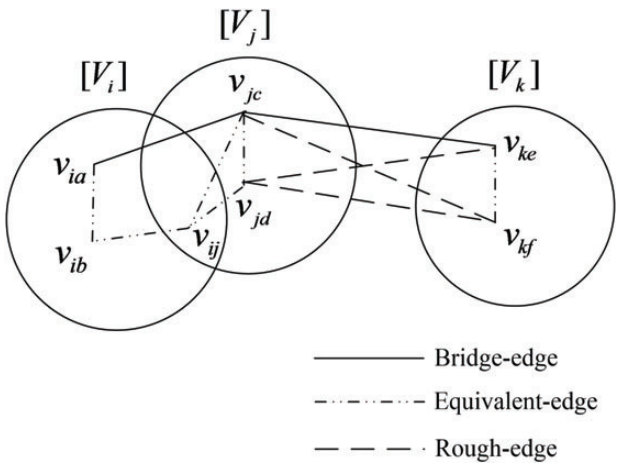

FIGURE 2. Potential paths between overlapped communities.

gained. Based on Definitions 6 and 7, we present all potential paths, namely, $PP_{v_{ia} \to v_{ke}} = <v_{ia}, v_{jd}>, <v_{jd}, v_{ke}>$, $PP_{v_{ia} \to v_{kf}} = <v_{ia}, v_{jd}>, <v_{jd}, v_{kf}>$, $PP_{v_{ib} \to v_{ke}} = <v_{ib}, v_{jd}>, <v_{jd}, v_{ke}>$ and $PP_{v_{ib} \to v_{kf}} = <v_{ib}, v_{jd}>, <v_{jd}, v_{kf}> / <v_{ib}, v_{jc}>, <v_{jc}, v_{kf}>$ in Fig. 1.

However, there also exist overlapped communities in social networks based on a division of only non-exclusive factors as hobbies and interests. Overlapping means that some nodes may belong to more than one community. Figure 2 shows potential paths in overlapping communities. Node v_{ij} belongs to two communities and has an equivalence relation with other nodes from two communities, $[V_i]$ and $[V_j$. Thus, a rough relation between $[V_i]$ and $[V_j]$ does not exist. According to Definition 4, three rough edges exist in Fig. 2, namely, $<v_{jc}, v_{kf}>$, $<v_{jd}, v_{ke}>$ and $<v_{jd}, v_{kf}>$.

4. TRUST MEASUREMENTS OF POTENTIAL PATHS AND CREDIBLE POTENTIAL PATHS

Potential paths in the MSN are described by the rough set method in the preceding section. However, many potential paths can be mined in MSNs. Higher time complexity is needed to search all potential paths, and the probability of potential paths can be reduced gradually by increasing path length. Thus, a measurement standard should be established for potential paths. An MSN is an individual relational network based on interpersonal relationships among users. In networks, users can share and spread digital content based on a certain trust relationship, which is an important factor in the spreading and sharing of digital content. Therefore, we can mine credible potential paths by calculating the trust value of potential paths. One direct approach is to set a potential path trust threshold, which involves seeking potential paths from a starting node to an ending node and calculating the trust value of the potential paths. The self-defined trust threshold is beneficial for users to set and control the baseline of digital content sharing among interested and trustworthy users. Indeed, the trust computation and dynamic adjustment depend on the users' past activity, share history and current relationships [19]. The obtained trust value is then compared with the path trust threshold. If the trust value of a potential path is more than or equal to the threshold, this path is considered a credible potential path.

4.1. Rough edge trust

This paper aims to find the potential path between communities in MSNs. Definition 4 indicates that a rough relation is obtained by using an equivalence relation and weak relation on the bridge edge between communities. Therefore, we must determine the trust value of the equivalent edge in any community as well as that of the bridge edge between communities to obtain the trust value of the rough edge.

4.1.1. Equivalent edge trust

DEFINITION 8 (Equivalent edge trust (EET)). *In an MSN, all users in the same community have equivalent relations. They have a history of direct sharing and distributing digital content among themselves, and the trust relationship among users in the same community is a direct trust relationship.*

The EET in a community is the direct trust between two nodes connected on equivalent edges in that community. In this case, X and Y are two users in the same community, and EET_X^Y demonstrates the direct trust value from X to Y, as shown in Fig. 3.

EET calculation can adopt a trust model for MSNs [19]. In an empirical study, we made on user behaviors and analysis of related factors for media social network, we found that 82.5% of respondents share with relatives, friends and colleagues based on mutual direct/inferred trust relationships and media contents experience feedback [20]. For the nature of trust and the properties of digital content sharing, the paper focuses on the following factors: credible feedback (Rc), which is a user trust evaluation on shared content, feedback weight factor (Fb), time decay function [$\omega(t)$] and user sharing similarity (Sm) of the shared content. These shared contents are used to calculate EET (EET_X^Y) from user X to Y.

FIGURE 3. Equivalent edge trust.

DEFINITION 9 (Digital content credible feedback (Rc)). *After each sharing session, the user can send feedback about the session based on the degree of security and credibility of the shared digital content. This feedback is expressed as* $Rc \in \{0, 1\}$, *where 1 indicates that the shared digital content is safe and consistent with the statement of the sharer about the content, and 0 indicates that the shared digital content is unsafe and not credible.*

DEFINITION 10 (Feedback weight factor (Fb)). *Malicious sharing users also exist in MSNs. These users often provide exaggerated or false feedback about shared content and lie about their sharing sessions to increase their trust values or slander other sharing users. Therefore, feedback weight factor* Fb *is introduced to balance the credibility of the digital content feedback. For* Fb, *the trust value of feedback of the user at a previous sharing is used to represent his or her feedback credibility, which is expressed as* $Fb \in [0, 1]$.

DEFINITION 11 (Time decay function ($\omega(t)$)). *In accordance with the dynamics of trust and time decay, trust relation strength changes over time; thus, a user's recent sharing behaviors can better reflect his or her credibility. More distant sharing sessions have less influence on the current trust evaluation, which indicates that credible feedback of sharing history information is moderately significant in trust evaluation. Therefore, time decay function* $\omega(t)$ *is defined in the following equation as*

$$\omega(t) = \left\lceil (t_{\text{present}} - t_{\text{share}})/\delta \right\rceil, \tag{2}$$

Where t_{present} *represents the current time,* t_{share} *represents the location of the sharing period and* δ *refers to the trust decay period. Trust relation decays once every* δ *sharing period, that is,* $\delta \geq 1$. *Users can define the size of* δ *based on a detailed sharing scenario. Specifically, if* δ *is larger, trust relation decays more slowly through the sharing period decay, whereas the trust value decays faster if* δ *is smaller.*

DEFINITION 12 (User sharing similarity (Sm)). *In digital content-sharing communities, users prefer to share digital*

content with other users who have the same sharing activities and interests, which makes establishing a sharing relation easier. Trust value rapidly increases with the sharing of preferred digital content. User sharing preferences are mostly represented by multi-group $\vec{p}$, as shown in the following equation:

$$\vec{p} = (k_1, k_2, \ldots, k_n), \tag{3}$$

where n refers to the class of shared digital content in sharing scenarios and k_i represents the proportion of sharing times of i class of the digital content to the total sharing times, $k_i \in [0, 1]$. Sharing similarity between users X and Y $Sm(X, Y)$ can be represented by the cosine similarity, as shown in the following equation:

$$Sm(X, Y) = (\vec{p}_X \cdot \vec{p}_Y)/(|\vec{p}_X||\vec{p}_Y|), \quad Sm \in [0, 1]. \tag{4}$$

For each sharing session, if a user saves the related sharing information and the total sharing period of user X and Y is set to N_t, the obtained EET value of any two users, X and Y, in the community is expressed in the following equations:

$$\mathrm{EET}_X^Y = \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{Sm(X, Y) \times Fb(X, Y) \times Rc_i(t)}{\omega_i(t)}. \tag{5}$$

4.1.2. Bridge edge trust

DEFINITION 13 (Bridge edge trust (BET)). *In an MSN, a weak edge that connects the different communities is called a bridge edge. Bridge edge trust is the direct trust between two users connected to a bridge edge.*

As shown in Fig. 4, X and Y are two users in different communities that have a direct trust relationship. In this study, BET_X^Y represents the BET from X to Y.

These bridge nodes have a direct digital content-sharing history, and the relation between bridge nodes refers to the direct trust relation. The trust relation computation is the same with the direct trust evaluation in a community. Thus, the BET calculation is similar to the EET calculation, as shown in the following equation:

$$\mathrm{BET}_X^Y = \frac{1}{N_t} \sum_{i=1}^{N_t} \frac{Sm(X, Y) \times Fb(X, Y) \times Rc_i(t)}{\omega_i(t)}. \tag{6}$$

4.1.3. Rough edge trust

DEFINITION 14 (Rough edge trust (RET)). *Rough edge trust is the trust between two users with a rough edge.*

Definition 4 indicates that rough edge is obtained by assuming a relationship between the equivalent and the bridge edges. Thus, RET can be obtained by integrating equivalent edge trust and bridge edge trust. In addition, RET calculation includes the following two situations.

(a) One-degree rough relation: A node has an indirect and uncertain relation with another node in a different community based on an equivalence and weak relation on a bridge edge. This kind of rough relation is called one-degree rough relation. Figure 5 shows that W has a one-degree rough relation with Y, and the trust value calculation is shown in the following equation:

$$RET_W^Y = EET_W^X * BET_X^Y. \tag{7}$$

(b) Two-degree rough relation: A node has an indirect and uncertain relation with another node in a different community based on a two-degree rough and equivalence relation. This kind of rough relation is called a two-degree rough relation. Figure 5 shows that W has a two-degree rough relation with Z.

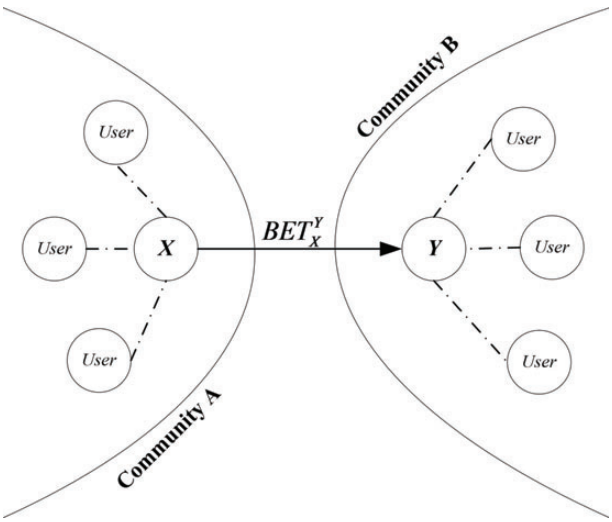

FIGURE 4. Bridge edge trust.

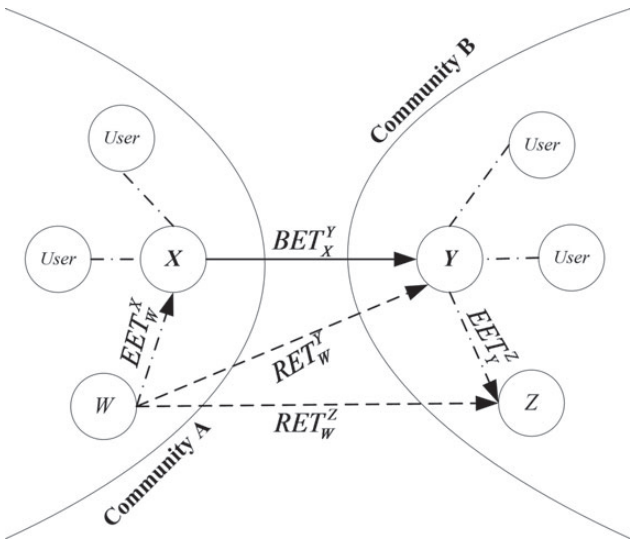

FIGURE 5. Rough edge trust.

The trust value calculation is shown in the following equation:

$$RET_W^Z = RET_W^Y * EET_Y^Z. \tag{8}$$

4.2. Trust value of potential and credible potential paths

Definition 7 indicates that all edges on the potential path are rough edges. The trust value of the potential paths (expressed as T_{PP}) is a product of all RET values in the path, as shown in the following equation:

$$T_{\mathrm{PP}_{v_1 \to v_n}} = T_{\mathrm{PP}}(v_1, v_2, \ldots, v_n) = \prod_{i=1}^{n} \mathrm{RET}_{v_i}^{v_{i+1}} (i = 1, 2, \ldots, n-1), \tag{9}$$

where the relation between v_1 and v_2, v_2 and v_3, $\ldots$, v_{n-1}, and v_n is the rough relation.

Figure 6 shows that a potential path $<W, Y><Y, Q>$ from W to Q exists. Moreover, based on Equation (9), the trust value of this potential path is $T_{\mathrm{PP}_{W \to Q}} = T_{\mathrm{PP}}(W, Y, Q) = \mathrm{RET}_w^Y * \mathrm{RET}_Y^Q$; the calculation procedures of RET_W^Y and RET_Y^Q are described in Section 4.1.

Based on the trust value of the potential paths, we attempt to find the credible potential paths by using the following steps. First, the trust value T_{PP} of the potential path is calculated. Secondly, the user defines a trust threshold, denoted by $T_{\mathrm{threshold}}$. Finally, T_{PP} and $T_{\mathrm{threshold}}$ are compared.

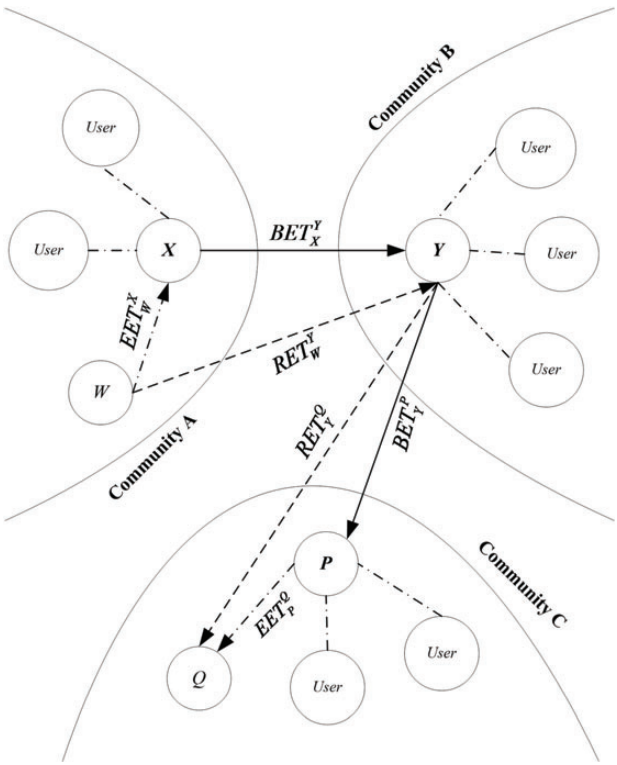

FIGURE 6. Potential path trust.

DEFINITION 15 (Credible potential path (CPP)). *When* $T_{\mathrm{PP}}(v_1, v_2, \ldots, v_{i-1}) \geq T_{\mathrm{threshold}}$ *and* $T_{\mathrm{PP}}(v_1, v_2, \ldots, v_{i-1}, v_i) < T_{\mathrm{threshold}}$, *a potential path* $PP(v_1, v_2, \ldots, v_{i-1})$ *is deemed a credible potential path from* v_1 *to* v_{i-1}.

5. ALGORITHM FOR MINING POTENTIAL PATHS AND CREDIBLE POTENTIAL PATHS

As stated in Section 3.2, the potential paths between two overlapped communities do not exist. Thus, we propose algorithms for mining potential paths and credible potential paths for non-overlapping communities in MSNs; this algorithm is a qualitative approach (Algorithm 1). In addition, a quantitative approach (Algorithm 2) for calculating the trust value in these potential paths was performed to determine the credible potential paths, which is within the range of user-defined trust threshold values. The following algorithms show how to mine the potential paths and the credible potential paths.

In the following algorithms, the MSN consists of three equivalent community classes, expressed as $[V_1]$, $[V_2]$, $[V_3]$. According to Definition 5, a rough relation matrix between three equivalent community classes is then established, which is $3 * 3$ matrix $M_{S*} = (w_{ij})$. The detailed algorithms are described as follows:

Algorithm 1 describes the potential path mining between i and j, which includes judgment and cycle statement. Its computational overhead is $O(n)$, where n denotes the number of nodes in $[V_3]$. The efficient method results in a little and limited overhead for the management of social network.

Further, the credible potential path mining algorithm is provided as follows:

6. EXPERIMENT AND ANALYSIS

In Section 5, the (credible) potential path mining algorithms are proposed. In order to verify their effectiveness, we made the following simulation experiments. First, we need to find a MSN including several non-overlapped communities as in Fig. 1. Secondly, based on these found communities, the goal of the experiments is to find the general and credible potential paths between any two user nodes of multimedia content dissemination, as is helpful to further positively solve DRM application of MSNs. Some of the algorithms for finding a community were proposed related to social network and social computing [21]. Here, we adopt the following method to accomplish the first step.

By using a representative real-world MSN YouTube dataset, including 1 138 499 linked users and 2 990 443 edges (http://socialnetworks.mpi-sws.org/data-imc2007.html), we found a random MSN with non-overlapped communities, as shown in Fig. 7. Three sharing communities are involved with the random MSN, and they are written as Community1, Community2 and Community3, which are connected by some

Algorithm 1 Potential path mining between any two nodes in MSNs.

Input:
A_M // Adjacency matrix of MSN
$M_{S^*} = (w_{ij})$ //Rough relation matrix between equivalence community classes
BE[][] //Bridge-edges array in MSN
i // Starting node and $i \in [V_1]$
j // Ending node and $j \in [V_2]$
x // The node in $[V_3]$

Output:
$PP_{i\to j}$ // The potential paths from i to j

```
BEGIN
  if (w12 = 0&&  w13 = 1&&  w23 = 1)        //<[V1], [V3]> ∈ S* and <[V2], [V3]> ∈ S*
      for (x = 0; x<n; x++)                 // Makes an ergodic process of all nodes in [V3]
                                            //n denotes the number of nodes in [V3]
        {
          if (<i, x>! == BE[][]&&<x, j>! == BE[][])     //<i, x> ∉ S and <x, j> ∉ S
          PP_{i→j} ← <i, x>, <x, j>
          return PP_{i→j}
        }
  else if (w12 = 1&&w13 ∧ w23 = 0&&<i, j>! == BE[][])
      for (x = 0; x<n; x++)
        {
          PP_{i→j} ← <i, j>
          return PP_{i→j}
        }
  else if (w12 = 1&&w13 = 1&&w23 = 1)
      for (x = 0; x<n; x^++)
        {
          if (<i, j>! == BE[][]&&<i, x>! == BE[][]&&<x, j>! == BE[][])
              PP_{i→j} ← <i, j>
              PP_{i→j} ← <i, x>, <x, j>
          return PP_{i→j}
          else if (<i, j> == BE[][]&&<i, x>! == BE[][]&&<x, j>! == BE[][])
              PP_{i→j} ← <i, x>, <x, j>
          return PP_{i→j}
        }
  else return 0
END
```

Algorithm 2 Credible potential path mining between any two nodes in MSNs.

Input: All potential paths $PP_{i\to j}$ from i to j; number of share cycles *ShareNum*; trust calculation window size *WindowSize*; feedback weight factor *Rate*; and trust threshold $T_{threshold}$.

Output: The trust values $T_{PP_{i\to j}}$ of $PP_{i\to j}$ and all the credible paths $CPP_{i\to j}$.

(1) To calculate direct trust value between the nodes from the same equivalence community class based on *ShareNum*, *WindowSize*, and *Rate*;

(2) Based on the potential path trust calculation method in Sections 4, the trust value $T_{PP_{i\to j}}$ of all potential paths between i and j is computed;

(3) The obtained trust value $T_{PP_{i\to j}}$ from (2) is compared with the trust threshold $T_{threshold}$. According to Definition 15, $CPP_{i\to j}$ will then be obtained;

(4) Return trust value $T_{PP_{i\to j}}$ and the credible paths $CPP_{i\to j}$.

Algorithm 3 Finding a community by using a seed node in MSNs.

Input: Graph G; initial node degree d; Size of community s
Output: interior links of a found community

(1) For nodes of G, whose degrees are greater than d, randomly select an initial one n_1 from the nodes, and append n_1 to null list L;
(2) Randomly select the second node n_2 from nodes linked by n_1, and append n_2 to list L;
(3) **if** the intersection set of linked between n_1 and n_1 is not null

then select the third n_3 from the intersection set randomly, **else** select the third n_3 from nodes linked by n_1, and append n_3 to list L;

(4) Select one node n from G which has the strongest link with all nodes of L;
(5) Repeat (4), **until** the number of L is equal to s;
(6) Return all links of L nodes from G.

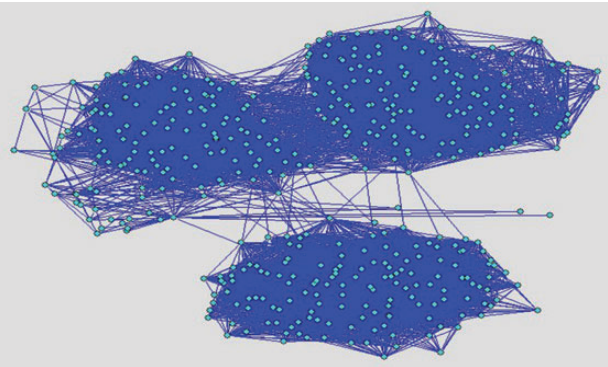

FIGURE 7. Random non-overlapped MSN found by using YouTube.

extra-community bridge edges called as weak ties that are different from the inner-community edges. We defined the size of the community followed by the 'Rule of 150', which indicates that the node number of each sharing community is 150 or so for any user. The related main parameters of the random MSN are also shown in Table 1.

Based on the MSN, C++ language programming is used to mine potential paths and credible potential paths. The experiment can achieve some functions, including finding potential paths between any two nodes, calculating the potential path trust and finding the credible potential path in the Fig. 7 non-overlapped network. The detailed experimental steps are as follows: First, the corresponding trust impact factor, the number of share cycles, trust calculation window size and feedback weight factor are simulated. Subsequently, direct trust value between all node pairs in the network can be calculated. Secondly, all potential paths between any two nodes as well as credible potential paths are found (the trust threshold value is set as 0.5 in the experiment).

We employed a computer system with an Intel Core i3-2130 Processor with 3.40 GHz, 4 GB RAM and Microsoft Windows 7 Ultimate in the experiment. The overhead of the experiment includes three parts: 2055.7167 s for establishing a social network by the YouTube dataset, 38.1953 s used for computing feedback factors relative to trust evaluation, as well as 0.5665 s for (credible) potential path mining.

Table 2 shows the obtained results after calculating the trust value of all the potential paths and mining the credible potential paths from starting node 29 to ending node 159, when the number of share cycles, the trust calculation window size and feedback weight factor are 100, 60 and 0.88, respectively. When the trust threshold is 0.5, 150 potential paths and 90 credible potential paths are found from node 29 to node 159 (in the following table, '–' denotes that the path is not a credible potential path).

Table 3 shows the results obtained by calculating the trust value of all the potential paths and seeking credible potential paths from starting node 130 to ending node 349 when the number of share cycles, trust calculation window size and feedback weight factor are 50, 20 and 0.9, respectively. The results show that 151 potential paths and 52 credible potential paths exist from node 130 to node 349 within the range of a trust threshold value of 0.5.

The above experiment results show that the proposed rough set-based method can effectively mine potential paths between communities for an MSN, and the potential path trust is calculated by using the trust assessment method to find the credible potential paths according to user-defined trust threshold values.

7. CONCLUSIONS AND FUTURE WORK

This paper proposed a method for finding (credible) potential paths as well as mining algorithms for an MSN. Using the basic properties of MSNs as a division criterion, the proposed method adopts the binary relation rough method based on rough set theory to find the rough relation and potential path between MSN communities. We also combined the trust calculation method and compared it with the user-defined trust threshold value to find the credible potential paths. Finally, two algorithms for mining potential paths and credible paths between any two nodes were proposed, and then simulation experiments were carried out to verify the

TABLE 1. Main parameters of the Fig. 7 network.

Network	Number of nodes	Number of edges	Average degree	Average clustering coefficient	Average shortest path length	Maximum path length
Community1	150	3782	50.42	0.4849	1.668	3
Community2	150	3510	46.80	0.3957	1.6866	3
Community3	150	2913	38.84	0.4470	1.9736	4
Random MSN	450	10514	46.72	0.426047	2.63869	5

TABLE 2. Two kinds of potential paths from starting node 29 to ending node 159.

Paths	General potential paths	Trust value	Credible potential paths
1	<29, 300>, <300, 159>	0.541338	<29, 300>, <300, 159>
2	<29, 301>, <301, 159>	0.482845	–
3	<29, 302>, <302, 159>	0.5213	<29, 302>, <302, 159>
4	<29, 303>, <303, 159>	0.492093	–
5	<29, 304>, <304, 159>	0.54123	<29, 304>, <304, 159>
		.....	
146	<29, 445>, <445, 159>	0.560105	<29, 445>, <445, 159>
147	<29, 446>, <446, 159>	0.438473	–
148	<29, 447>, <447, 159>	0.550362	<29, 447>, <447, 159>
149	<29, 448>, <448, 159>	0.521657	<29, 448>, <448, 159>
150	<29, 449>, <449, 159>	0.48376	–

TABLE 3. Two kinds of potential paths from starting node 130 to ending node 349.

Paths	General potential paths	Trust value	Credible potential paths
1	<130, 349>	0.82615	<130, 349>
2	<130, 150>, <150, 349>	0.475374	–
3	<130, 151>, <151, 349>	0.500712	<130, 151>, <151, 349>
4	<130, 152>, <152, 349>	0.475362	–
5	<130, 153>, <153, 349>	0.5062	<130, 153>, <153, 349>
		.....	
147	<130, 295>, <295, 349>	0.454442	–
148	<130, 296>, <296, 349>	0.480451	–
149	<130, 297>, <297, 349>	0.485857	–
150	<130, 298>, <298, 349>	0.461094	–
151	<130, 229>, <229, 349>	0.536041	<130, 229>, <229, 349>

proposed approaches. At present, two kinds of potential paths have been discovered. Our future work will evaluate security risks of content dissemination and sharing on the potential paths for MSNs.

ACKNOWLEDGEMENTS

We thank Dr Zhao Changwei for his technical assistance on experiments, and also would like to thank the reviewers and the associate editor for their valuable comments, questions and suggestions.

FUNDING

The work was sponsored by National Natural Science Foundation of China (Grant No. 61370220, 71331007), Plan for Scientific Innovation Talent of Henan Province (Grant No. 134100510006), Key Program for Basic Research of The Education Department of Henan Province (Grant No. 13A520240,

No. 14A520048). Program for Science & Technology Innovation Talents in Universities of Henan Province (Grant No. 2011HASTIT015).

REFERENCES

[1] Zhang, Z. (2012) *Security, Trust and Risk in Digital Rights Management Ecosystem*. Science Press, Beijing, China.

[2] Zhang, L., Zhang, Z., Niu, D. and Huang, T. (2011) A Novel DRM security scheme and its prototype system implementation. *Internat. J. Digital Content Technol. Appl.*, **5**, 334–342.

[3] Zhang, Z. (2011) Digital rights management ecosystem and its usage controls: a survey. *Int. J. Digital Content Technol. Appl.*, **5**, 255–272.

[4] Raad, E., Chbeir, R. and Dipanda, A. (2011) Discovering relationship types between users using profiles and shared photos in a social network. *Multimedia Tools Appl.*, **64**, 141–170.

[5] Girvan, M. and Newman, M. (2002) Community structure in social and biological networks. *Proc. Natl Acad. Sci.*, **99**, 7821–7826.

[6] Granovetter, M. (1973) The strength of weak ties. *Am. J. Sociol.*, **78**, 1377–1378.

[7] Braghin, S., Ferrari, E. and Trombetta, A. (2010) Combining Access Control and Trust Negotiations in an On-Line Social Network. *Proc. 6th Int. Conf. on Collaborative Computing*, Chicago, USA, October 9–12, pp. 1–10. IEEE Computer Society, Piscataway, USA.

[8] Wang, H. and Sun, L. (2010) Trust-Involved Access Control in Collaborative Open Social Networks. *Proc. 4th Int. Conf. on Network and System Security*, Melbourne, Australia, September 1–3, pp. 239–246. IEEE Computer Society, Piscataway, USA.

[9] Sachan, A., Emmanuel, S. and Kankanhalli, M. (2010) An Efficient Access Control Method for Multimedia Social Networks. *Proc. 2nd ACM SIGMM Workshop on Social Media*, Firenze, Italy, October 25, pp. 33–38. ACM, New York, USA.

[10] Park, J., Sandhu, R. and Cheng, Y. (2011) A user-activity-centric framework for access control in online social networks. *IEEE Internet Comput.*, **15**, 62–65.

[11] Lian, S., Chen, X. and Wang, J. (2012) Content distribution and copyright authentication based on combined indexing and watermarking. *Multimedia Tools Appl.*, **57**, 49–66.

[12] Chung, M. and Ko, I. (2012) Intelligent copyright protection system using a matching video retrieval algorithm. *Multimedia Tools Appl.*, **59**, 383–401.

[13] Carminati, B., Ferrari, E., Morasca, S. and Taibi, D. (2011) A Probability-Based Approach to Modeling the Risk of Unauthorized Propagation of Information in On-Line Social Networks. *Proc. 1st ACM Conf. on Data and Application Security and Privacy*, San Antonio, USA, February 21–23, pp. 51–61. ACM, New York, USA.

[14] Wang, T., Srivatsa, M., Agrawal, D. and Liu, L. (2011) Modeling Data Flow in Socio-information Networks: A Risk Estimation Approach. *Proc. 16th ACM Symp. on Access Control Models and Technologies*, Innsbruck, Austria, June 15–17, pp. 113–122. ACM, New York, USA.

[15] Mezzour, G., Perrig, A., Gligor, V. and Papadimitratos P. (2009) *Privacy-Preserving Relationship Path Discovery in Social Networks*. Lecture Notes in Computer Science 5888, pp. 189–208. Springer, Heidelberg, Germany.

[16] Xue, M., Carminati, B. and Ferrari, E. (2011) P^3D—Privacy-Preserving Path Discovery in Decentralized Online Social Networks. *Proc. 35th IEEE Annual Computer Software and Applications Conf.*, Munich, Germany, July 18–21, pp. 48–57. IEEE Computer Society, Piscataway, USA.

[17] Bródka, P., Stawiak, P. and Kazienko, P. (2011) Shortest Path Discovery in the Multi-layered Social Network. *Proc. 2011 Int. Conf. on Advances in Social Networks Analysis and Mining*, Kaohsiung, Taiwan, July 25–27, pp. 497–501. IEEE Computer Society, Piscataway, USA.

[18] Pawlak, Z. (1982) Rough sets. *J. Inf. Comput. Sci.*, **38**, 341–356.

[19] Zhang, Z. and Wang, K. (2012) A trust model for multimedia social networks. *Soc. Netw. Anal. Mining*, **4**, 969–979.

[20] Zhang, Z. (2012) Study on Digital Rights Management in Multimedia Social Networks. Post-Doctoral Fellowship Technical Report, Xi'an Jiaotong University, Xi'an, Shannxi, China.

[21] Yang, J. and Leskovec, J. (2012) Defining and Evaluating Network Communities Based on Ground-Truth. *Proc. IEEE Int. Conf. on Data Mining*, Brussels, Belgium, December 10–13, pp. 745–754. IEEE Computer Society, Piscataway, USA.