# Rough Set and Trust Assessment-based Potential Paths Analysis and Mining for Multimedia Social Networks

Lijun Yang, Zhiyong Zhang*, Jiexin Pu
*Electronics Information Engineering College, Henan University of Science and Technology, Luoyang 471003, China*
*E-mail: z.zhang@ieee.org*

### *Abstract*

*The emerging multimedia social network is a convenient platform on which multimedia content, such as e-books, digital images, audios and videos can be distributed and shared. In the open network scenario, users can randomly share or spread digital contents protected by copyright; thus, Digital Rights Management (DRM) has become a growing problem. This paper's objective is an exploration of potential contents propagations path between users. A novel qualitative and quantitative approach to analysis and mining potential paths is proposed for multimedia social networks. The qualitative analysis and formal description were primarily based on the theory of rough set. And also, the quantitative computing of trust measurement was performed in order to determine a trusted potential path, which is within the range of user-defined trust threshold values. The proposed method would effectively and accurately find the potential paths of the digital contents distributing and sharing as to further resolve the burning DRM.*

**Keywords***: Multimedia Social Networks, Digital Rights Management, Rough Set, Potential Path, Trust Measurement*

## 1. Introduction

The rapid development of network socialization has led to the emergence of different services and functions provided by many multimedia social networks (MSNs). These MSNs offer network tools, services, and applications for multimedia content (e.g., words, images, or videos) that can be shared among users in the same group or between different groups within social networks. Currently, the more prevalent local and international MSNs are Youtube, GoogleVideo, Tudou and Youku, among others. Given that these networks offer direct, rapid, and flexible transmissions of audio-video digital content, they have facilitated the sharing of digital multimedia content between users. However, due to the convenience of information exchange and information interaction provided by open multimedia network environments, and the harmless reproduction and easy distribution of digital content, valuable digital content protected by copyrights have also been shared and spread randomly within these networks. Therefore, issues on digital rights management (DRM) have become more prominent in recent years.

Existing digital content security and copyright protection [1] research is centered on cryptology protection and access control [2,3] based on the goal of secure sharing of digital contents (rights). However, these technologies have failed to meet the requirements of MSN, hence paving the way toward a new scenario. The MSN is characterized by a small world, so it has large promoting action on spreading multimedia contents/rights. Moreover, user relationship is characterized by rapid transitivity, dynamics, and strengthening of weak ties; thus, it immeasurably extends the range of multimedia contents/rights in the network. With explosive growth of user scale and content sharing behavior in this network, controlling content dissemination has become more difficult. Direct or indirect relationships exist between users in any MSN. The direct relation can be observed; whereas the indirect relation cannot be observed easily because it is relatively concealed. Therefore, finding the implicit propagation path (concealed or potential) between users has become another aspect of DRM research in the context of MSNs. Based on the structural property of MSN, this paper adopts the binary relation rough method in rough set to explain the rough path or potential propagation path existing between users. This paper identifies a believable potential path by combining its trust computation.

# 2. Related Literature

## 2.1 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 network can be established by real life acquaintances (e.g., friends, classmates, colleagues, or business partners) or strangers but those who share common bonds (e.g., friend recommendation or having common interests). Based on the establishment of certain relationships in a social network, users can have real-time information exchange and resource sharing with one another. To formally study the characteristics of a social network, scholars model it as a diagram based on the relationship between people. In a social network, each user is regarded as one point in the diagram and the relationship between users is represented by the edge. The strength of the degree of relation between individuals is very important for content sharing and spreading in social networks. Therefore, the link relationship network between users in a social network is expressed as: $G=(V,E,\varpi)$, where $V$ refers to congregation of all users, $E$ refers to congregation of edges between user nodes in the network, and $\varpi$ refers to weight information between users.

Apart from its other properties, such as higher aggregation and being scale-free and a "small world," a social network has another important property, namely, community structure [4]. Community structure is an inherent basic structure of a social network, which reflects the distribution of users, node source, and other inherent properties. Furthermore, a social network consists of several locally dense "communities." Each community represents an actual social organization formed on the basis of social relationship or interest. In a community, members communicate more closely. Similarly, the topology of a social network shows that members also connect closely within the group. Hence, a more obvious partition exists between them and other nodes outside the community. Fig. 1 shows a simple social network structure map formed by three communities. Accordingly, one community is composed of a series of nodes. The connection between nodes in each community is relatively close, but that between communities is very loose. Hence, the connection between nodes in the same community is very strong, whereas that between communities, specifically those between two nodes in different communities are relatively weak. Granovetter [5] proposes the concept of "strength of weak link," and uncovered the different roles played by strong and weak relationships in the information spread process. He believes that information spread by strong relation might be limited to a smaller range, and does not have large value for users in the community. Meanwhile, relative to strong relation, a weak relation tends to transfer non-recurring information between different communities. Therefore, more alternation and information spread are performed between communities through weak relations, hence making it an "information bridge."

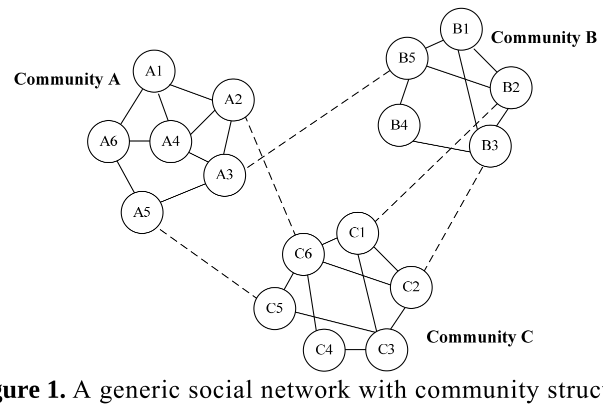


**Figure 1.** A generic social network with community structure

### 2.2 DRM study of a social network

In recent years, scholars here and abroad have carried out extensive studies on DRM issues in the social network platform. Wang et al. [6] report that traditional access control is conducted in data management system; however, it is unsuitable for collaborative open social network scenarios. They propose a kind of trust-based access control (TBAC) framework, and compared with other traditional methods, the TBAC framework can provide secret protection to support higher complex relations. As for content access control in MSNs, Sachan et al. [7] propose a fine-grained multimedia access control model based on bit-vector transform, which verifies the safety and storage and enforcement efficiency through mathematics and simulation analysis. Barbara et al. [8] emphasize that the reinforced social network access control system is the key factor in solving safety and privacy issues in online social networks. To resolve current limitations, the generated data have been adopted to assume a test social network. This test network is used to verify the effectiveness of the proposed previously semantic reasoning method. And introduce an expandable fine-grained access control model based on semantic web online social networks. This model consists of authorization, management and filter, and uses OWL (Web Ontology Language) and SWRL (Semantic Web Rule Language). Lian et al. [9] introduce content distribution and copyright verification system based on media index and watermarking technology in connection with copyright protection of media content in MSNs and the illegal distribution of media content over the Internet. This system has been proven to be robust. Chung et al. [10] also propose a new video matching algorithm in connection with copyright protection of video content in YouTube and other MSNs; they also designed an intelligent copyright protection system based on this algorithm. Accordingly, their proposed algorithm can effectively perform video matching. Moreover, their system can be used in copyright protection within video sharing networks.

Therefore, the most common research methods in the copyright issue of digital content in social network scenarios are based on access control and digital watermarking. However, the current paper explores this issue from a new perspective, that is, to seek the potential spread path in social network.

## 3. Discovery of Potential Paths in MSNs based on the Rough Set

The topological structure of real social networks shows community structural characteristics, in which a network consists of several communities. The nodes within a community are very closely connected. In comparison, connections between communities are very loose. Therefore, the potential relationships and paths between different users can be identified through the weaker connections between communities using the binary relation rough method in the rough set.

### 3.1 Rough set

The rough set (RS) theory was first proposed in 1982 by Zdzisław Pawlak, a Polish scientist. Based on set theory, RS is a mathematical tool for processing incomplete information. Using the RS theory, we can find implicit knowledge and uncover the potential law. Partial definitions of the rough set related to this paper are given below.

**Definition 1** (Approximation Space) [11] Here, $U$ indicates a non-empty finite set composed of objects of study, called discourse domains. $R \subseteq U \times U$ is the aggregation of equivalence relations to $U$, and $M = (U, R)$ is the approximation space. If $S \subseteq U \times U$ is a binary relation to $U$ different from $R$, where $S$ is common binary relation, fuzzy relation, compatible relation, partial order relation, and so on, then $S$ is a binary relation in approximation space $M = (U, R)$.

**Definition 2** (Indiscernibility relation) [11] Equivalence relation $R$ divides $U$ into a series of discontinuous subsets represented by $U / R = \{G_1, G_2, \cdots, G_n\}$, and $G_n$ is an equivalence class of domain $U$ on $R$. If two objects $u, v \subseteq U / R$ belong to the same equivalence class, $u$ and $v$ are said to have an indiscernibility relation.

### 3.2 Formal description of potential paths in MSNs

The MSN is a dynamic network. With numerous changes over time, new entities and links have continuously joined these MSNs, while the old ones disappear. This paper provides a snapshot of a certain period in the growth of MSNs. First, we divide MSNs based on community structure, namely, the different communities where the users are located (equivalence relation), and provide a formal description of the potential paths in MSNs using the binary relation rough method in the rough set.

**Definition 3** (Approximation social network space) In MSNs, domain $V$ is set to users in all nodes within the network, and $V$ is divided based on different communities (equivalence relation $R$ ) to form an approximate MSN space $G=(V,R)$. Then, $V/R=\{[V_1],[V_2],\cdots,[V_n]\}$ is the division of $V$ on $R$, where $[V_i](1\le i\le n)$ is called community equivalence class in approximate MSNs. Nodes in the same community equivalence class have an equivalence relation. The connected edge of any nodes in the same community is called equivalent edge.

**Definition 4** (Rough Relation) In an approximate MSN space, $G=(V,R)$ , $V/R=\{[V_1],[V_2],\cdots,[V_n]\}$ is the division of $V$ on $R$, and $S$ is the direct binary relation different from $R$. As for any two community equivalence classes, $[V_i],[V_j]$ $(1\le i,j\le n 且 i\ne j)$, $\forall a\in[V_i]$ and $\forall b\in[V_j]$ if $<a,b>\in S$, and $S^*=\{<[V_i],[V_j]>\}$ is named as rough relation of $S$ in an approximate MSN space $G=(V,R)$. The weak connection edge formed by relation $S$ for connecting different communities is called the bridge edge, and the two endpoints forming a bridge edge are called bridge nodes.

**Definition 5 (** $S^*$ path) In an approximate MSN space, $G=(V,R)$**,** $V/R=\{[V_1],[V_2],\cdots,[V_n]\}$, $S$ is the direct binary relation different from $R$ **.** $S^*$ is the rough relation of $S$ **.** If $S^*=<[V_1],[V_2]>,<[V_2],[V_3]>\cdots,<[V_{n-1}],[V_n]>$**,** then $<[V_1],[V_2]>,<[V_2],[V_3]>\cdots,<[V_{n-1}],[V_n]>$ is considered the $S^*$ path from $[V_1]$ to $[V_n]$.

**Definition 6** (Potential path) In an approximate MSN space $G=(V,R)$, $V/R=\{[V_1],[V_2],\cdots,[V_n]\}$. $S^*$ is the rough relation of $S$. If $<[c],[V_1]>,<[V_1],[V_2]>,\cdots,<[V_{n-1}],[V_n]>,<[V_n],[d]>$ is the $S^*$ path from $[c]$ to $[d]$. For any $a\in[c]$, $b\in[d]$ and $v_{i1},v_{i2},\cdots,v_{ix},\cdots v_{im}\in[V_i]$ $(i=1,2,\cdots,n)$, then $<a,v_{1x}>,<v_{1x},v_{2x}>,<v_{2x},v_{3x}>,\cdots,<v_{(n-1)x},v_{nx}>,<v_{nx},b>$ is called the potential path on $S$ from $a$ to $b$. The ordered pair in the potential path is called the rough edge (notes that each ordered pair in potential path $<x_{(i-1)x},x_{ix}>(i=1,2,\cdots,n)$ cannot be described correctly by relation $S$, that is, $<x_{(i-1)x},x_{ix}>\notin S$ ). Thus, the edges in the potential path are composed of rough edges.

$[V_i]$, $[V_j]$, and $[V_k]$ are the three community equivalences after the equivalent relation division of MSNs, $S=\{<v_{ia},v_{jc}>,<v_{jc},v_{ke}>\}$ (Fig. 2). Definition 4 shows that $S^*=\{<[v_i],[v_j]>,<[v_j],[v_k]>\}$. Definitions 5 and 6 are based on finding potential paths existing in the figure, that is, $<v_{ia},v_{jd}>,<v_{jd},v_{ke}>$ is the potential path in $S$ from $v_{ia}$ to $v_{ke}$; $<v_{ia},v_{jd}>,<v_{jd},v_{kf}>$ is the potential path in $S$ from $v_{ia}$ to $v_{kf}$; $<v_{ib},v_{jd}>,<v_{jd},v_{ke}>$ is the potential path in $S$ from $v_{ib}$ to $v_{ke}$; and $<v_{ib},v_{jd}>,<v_{jd},v_{kf}>$ is the potential path in $S$ from $v_{ib}$ to $v_{kf}$.

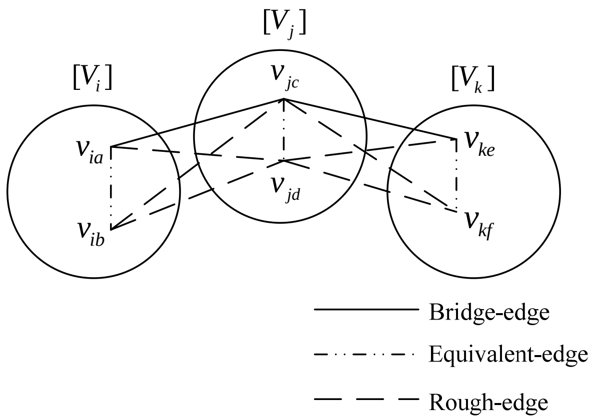


**Figure 2.** The discovery of potential paths between communities

## 4. Trust Measurement of the Potential Paths

The rough set method has been used to describe the potential path in MSNs in the preceding section. However, many potential paths can be found in large-scale and complex MSNs. Higher time complexity is needed for searching all potential paths, and the probability of potential path can be reduced gradually with the increase of path length. Hence, a measurement standard should be established for the potential paths. The MSN, an individual relational network, is established based on interpersonal relationships among users. Network users share and spread digital content based on a certain trusting relationship. This trusting relationship between users is an important factor that influences the spreading and sharing of digital content. Therefore, we can find credible potential paths by calculating the trust value of these potential paths. One direct idea is to set a path trust threshold. Here, we seek a potential path starting from the initial node and then calculate the trust value in this path. The obtained trust value is then compared with the set path trust threshold. If the trust value in the potential path is more than or equal to this threshold, it is considered a credible, potential path.

### 4.1 Trust of Rough-edge

This paper aims to find the potential path between communities in an MSN. Definitions 4 indicate that, rough relation is obtained by reasoning equivalence relation and weak relation (bridge relation) between communities. Therefore, to obtain the rough-edge trust, we first determined the equivalent-edge and bridge-edge trust value between communities.

(1) Equivalent-edge Direct Trust (EDT)

In an MSN, all users in the same community have equivalent relations. They have direct digital content sharing and spreading history, and the trust relationship among users in the community is a direct trust relationship. The EDT in a community is the direct trust between two nodes of connected equivalent edges in that community. In this case, V and U are two users in the same community, and $EDT_v^u$ demonstrates the direct trust value from V to U (Fig. 3).


**Figure 3.** Equivalent-edge Direct Trust

(2) Bridge-edge Direct Trust (BDT)

In MSNs, a weak connected edge within the connecting community is called a bridge-edge. Bridge-edge direct trust (BDT) is the direct trust between two users with a weak connected edge. V and U are two users with a direct, weak relation in different communities (Fig. 4). Here, $BDT_v^u$ represents BDT from V to U.

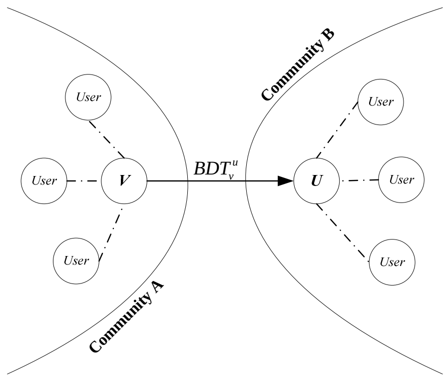


**Figure 4.** Bridge-edge Direct Trust

The BDT relation refers to the direct trust relation in a weak relation edge between communities. These bridge-edge users also have the same direct digital content sharing history. The BDT calculation between communities is same to EDT calculation in any community. The BDT calculation and EDT calculation can accord to a trust model for MSN in the literature [12].

The definitions of rough relationship indicate that the rough-edge is obtained from reasoning equivalent-edge and bridge-edge. That is, rough-edge trust (RT) can be obtained by integrating equivalent- and bridge-edge trust. Further, RT calculation can be divided into two situations detailed below.

(1) Degree I rough relation: A certain node has an indirect and uncertain relation with a certain node in another community by equivalence and bridge relation. This kind of rough relation is called Degree I rough relation. In Fig. 5, W has Degree I rough relation with U, and the calculation of trust value is shown in Equation (7) as follows:

$$RT_w^u = EDT_w^v * BDT_v^u \tag{7}$$

(2) Degree II rough relation: A certain node has an indirect and uncertain relation with a certain node in another community by Degree I rough relation and equivalent relation. This kind of rough relation is called Degree II rough relation. In Fig. 5, W has Degree II rough relation with Z, and the calculation of trust value is shown in Equation (8) as follows:

$$RT_w^z = RT_w^u * EDT_u^z \tag{8}$$

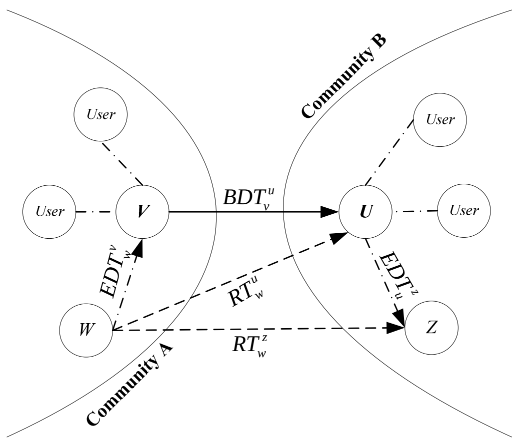


**Figure 5.** Rough-edge Trust

### 4.2 Trust value of potential path and credible potential path

Definition 6 indicates that edges of potential path are all rough edges. The trust value of the potential path (expressed as $T_{PP}$) is a product of all RT values in the path, as shown in Equation (9):

$$T_{PP}(v_1, v_2, \cdots, v_n) = \prod_{i=1}^{n} RT_{v_i}^{v_{i+1}} (i = 1, 2, \cdots, n) \qquad (9)$$

Fig. 6 also shows the definition of potential path, in which a potential path exists from W to T $< W, U >< U, T >$. Moreover, Eq. (9) indicates that the trust value of this potential path is $T_{PP}(W, U, T) = RT_w^u * RT_u^t$, where $RT_w^u$ and $RT_u^t$ are rough-edge trust values. The calculation procedure is described in Section 4.1.

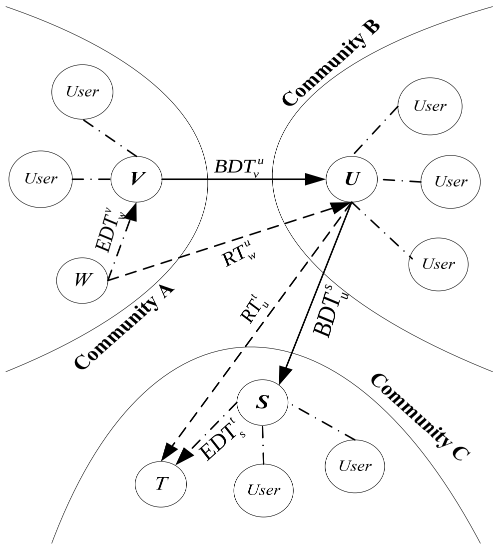


**Figure 6.** Potential Path Trust

The potential path trust calculation in Fig. 6 is based on finding a credible potential path. The method is defined as follows: first, the user custom-defines a trust threshold, which is denoted by $\alpha$. Then, trust measurement is made to identify the potential path; it is then compared with threshold $\alpha$. When $T_{PP}(v_1, v_2, \cdots, v_{i-1}) \geq \alpha$ and $T_{PP}(v_1, v_2, \cdots, v_{i-1}, v_i) < \alpha$, path $PP(v_1, v_2, \cdots, v_{i-1})$ is deemed a credible potential path.

## 5. Conclusion and Future Prospects

Using the basic properties of MSNs as a division criterion, this paper adopts the binary relation rough method based on the rough set. We also find the rough relation and potential path between MSN communities as well as present a formal description. These potential paths are implicit. Although weaker social relation is found between users in paths, indirect relation exists to a certain degree between them, and as such, the digital content is likely to spread in the future. In order to find the credible and potential path, we also combine the trust calculation method in the threshold range by considering custom-defined trust value as the threshold value.

**Acknowledgments** The work was sponsored by the National Natural Science Foundation of China (Grant No.61003234), Program for Science & Technology Innovation Talents in Universities of Henan Province (Grant No.2011HASTIT015), China Postdoctoral Science Foundation (Grant No.20100471611), and Henan University of Science & Technology Doctors Research Fund (Grant No.09001470).

## References

[1] Zhang ZY, "Security, Trust and Risk in Digital Rights Management Ecosystem", Science Press of China, Beijing, China, 2012.
[2] Zhang LL, Zhang ZY, Niu DM, Huang T, "A Novel DRM Security Scheme and its Prototype System Implementation", International Journal of Digital Content Technology and its Applications, vol.5, no.11, pp.334-342, 2011.
[3] Zhang ZY, "Digital Rights Management Ecosystem and its Usage Controls: A Survey", International Journal of Digital Content Technology and its Applications, vol.5, no.3, pp.255-272, 2011.
[4] Girvan M, Newman MEJ, "Community structure in social and biological networks", In Proceedings of the National Academy of Sciences, pp.7821-7826, 2002.
[5] Granovetter M, "The Strength of Weak Ties", American journal of Sociology, vol.78, no.6, pp.1377-1378, 1973.
[6] Wang H, Sun LL, "Trust-involved access control in collaborative open social networks", In Proceedings of the 4th International Conference on Network and System Security, pp.239-246, 2010.
[7] Sachan A, Emmanuel S, Kankanhalli M, "An efficient access control method for multimedia social networks", In: Proceedings of the 2nd ACM SIGMM Workshop on Social Media, pp.33-38, 2010.
[8] Barbara C, Elena F, Raymond H, Murat K, "Semantic web-based social network access control", Computers and Security, vol.30, no.2, pp.108-115, 2011.
[9] Lian SQ, Chen X, Wang JW, "Content distribution and copyright authentication based on combined indexing and watermarking", Multimedia Tools and Applications (Springer), vol.57, no.1, pp.49-66, 2012.
[10] Chung MB, Ko IJ, "Intelligent copyright protection system using a matching video retrieval algorithm", Multimedia Tools and Applications (Springer), vol.59, no.1, pp. 383-401, 2012.
[11] Pawlak Z, "Rough Sets", Journal of Information and Computer Sciences, vol.38, no.11, pp.341-356, 1982.
[12] Zhang ZY, Wang KL, "A Trust Model for Multimedia Social Networks", Social Networks Analysis and Mining (Springer), DOI: 10.1007/s13278-012-0078-4, 2012.