A social network-based system for supporting interactive collaboration in knowledge sharing over peer-to-peer network

Stephen J.H. Yang[a],*, Irene Y.L. Chen[b]

[a]Department of Computer Science and Information Engineering, National Central University, No. 300, Jhongda Road, Jhongli City, Taoyuan County 320, Taiwan
[b]Department of Information Management, Ching Yun University, No. 229, Chien-Hsin Road, JhongLi 320, Taiwan

Abstract

Knowledge sharing enables people in virtual communities to access relevant knowledge (explicit or tacit) from broader scope of resources. The performance in such environments is fundamentally based on how effectively the explicit and tacit knowledge can be shared across people, and how efficiently the created knowledge can be organized and disseminated to enrich digital content. This study will address how to apply social network-based system to support interactive collaboration in knowledge sharing over peer-to-peer networks. Results of this study demonstrate that applying such social network-based collaboration support to knowledge sharing helps people find relevant content and knowledgeable collaborators who are willing to share their knowledge.
© 2007 Elsevier Ltd. All rights reserved.

Keywords: Social network; Interactive collaboration; Knowledge sharing; Peer-to-peer; Instant messenger

1. Introduction

The explosion in Web-based technology has led to increasing volume and complexity of knowledge, which stimulates the proliferation of the virtual communities (e.g., peer-to-peer (P2P) network and instant messenger (IM)). The Internet provides an environment for people to contribute their knowledge and acquire others'. One of the Internet's intended purposes is to encourage knowledge sharing so that valuable knowledge embedded in the network can be effectively explored. Most Internet users expect that they can acquire and share valuable knowledge to fulfill their needs. However, knowledge sharing in some virtual communities has not lived up to the expectation. Two barriers preventing efficient and effective knowledge sharing are (Chen and Yang, 2006):

(1) The difficulty in finding relevant knowledge.
(2) The difficulty in finding relevant collaborators to interact with.

Knowledge sharing requires collaboration between the consumers and contributors of knowledge; namely the collaborators. This task cannot be accomplished simply by storing knowledge in the repository. It also requires a mechanism, which helps people find the collaborators with relevant knowledge. Collaboration over the Internet communities has characterized itself by heavily relying on interaction among the collaborators (Biström, 2005; Eikemeier and Lechner, 2003). Collaborators can be any virtual users who interact to achieve the goals of resources discovery, access, knowledge sharing, group communication and discussion. The collaboration for knowledge sharing should be enacted without spatial and temporal limitations. In addition, it should take place over medium such as the Internet and therefore beyond the geographical boundaries.

This research studies how to enhance knowledge sharing through a social network-based collaboration support. Social networks are built upon an idea that there exists a determinable networking structure of how people know each other. In such networks, people are connected through common social relationships either directly or

*Corresponding author. Tel.: +886 3 422 7151; fax: +886 3 422 2681.
   E-mail addresses: jhyang@csie.ncu.edu.tw (S.J.H. Yang),
irene@cyu.edu.tw (I.Y.L. Chen).

indirectly (Churchill and Halverson, 2005). Social networks identify two important relationship ties—knowledge relationship tie and social relationship tie. The two relationship ties are metrics used to measure the degree of people's knowledge matching with a query, and the degree of people's social relationship with other people.

Our social network-based collaboration support is implemented with P2P and IM technology to enhance content discovery and interactions among collaborators in knowledge sharing. P2P makes each peer act as both client and server (Aberer et al., 2002; Gnutella). It provides a metaphor commonly addressed in virtual communities that an individual could be a consumer as well as a producer of knowledge. This makes P2P particularly suitable for facilitating efficient knowledge sharing in Internet communities. P2P collaboration is supported by IM (Herbsleb et al., 2002; Isaacs et al., 2002a), which provides real-time interaction and communication. IM reveals not only how many people are online at a given time, but also who is online (Alexander, 2005; Grinter and Palen, 2002; Isaacs et al., 2002b; Nardi et al., 2000). P2P, IM, and social networks share many concepts in common. For example, they are both distributed networking structures; a peer in a P2P network is like a node in a social network; a link in a P2P network is like a relationship tie in a social network.

The major contribution of this paper is that it provides a mechanism, which adopts our social network-based system in the domain of a campus-wide P2P network to overcome the two aforementioned barriers. The studies and experiments presented in this paper demonstrate that applying our social network-based system can help people find relevant and quality resources over a campus-wide P2P network as well as relevant people who are willing and capable of sharing their knowledge. For finding relevant and quality resources, we utilized a rating mechanism to calculate a resource's reputation (Mitchell, 1997). For finding relevant people, we developed social net and knowledge net to calculate the degree of relationship ties between people in order to find the most related people who are willing and capable of sharing their knowledge (resources). In detail, we take a three-step approach.

(1) Firstly, we construct a P2P network of content repositories with 156 peers in a university.
(2) Secondly, upon receiving a request of collaboration from a user, we dynamically generate a network depending on the knowledge retrieved from the P2P network.
(3) Finally, we create an instance messenger (IM)-based virtual community to enable real-time collaborations between peers identified by the constructed network.

The remainder of the paper is organized as follows. In Section 2, we review related work. In Section 3, we introduce how to find relevant knowledge via our constructed P2P network. In Section 4, we present how to find relevant and knowledgeable collaborators via our

social network-based system. In Section 5, we discuss IM-enabled group collaboration and analyze experimental results. We conclude this study with future research in Section 6.

2. Related work

2.1. Knowledge sharing and virtual communities

The Internet gives rise to virtual communities that aim at facilitating collaboration by providing an environment for mutual sharing and interaction. A collaborative process in such an environment involves intensive online knowledge discovery and knowledge sharing between collaborators, such as knowledge consumers and knowledge contributors.

The emergent virtual communities over the past decade have stimulated research interests by academia and practitioners. For example, Zhang and Tanniru (2005) built an agent-based model for virtual learning communities (VLCs). The results of a series of comparison experiments indicated that every member tends to be active in the initial stage of the development of a VC; when the community reaches a reasonable size, each participant is important and function together to form a healthy and stable population. Bruckman (2002) found that the potential of Internet technology could come from the peers and elders. Jin (2002) provided a conceptual framework for the development of a prototype system of the virtual community based interactive environment. Wachter et al. (2000) pointed out that an enhanced environment is possible only if one goes beyond mere on-line course delivery and creates a community of people and other related resource groups. Wasko and Faraj (2005) found that knowledge sharing has been a motivation for participation in virtual communities. Although prior studies have provided evidence demonstrating the importance of knowledge sharing in enhancing the performance, few have attempted to provide mechanisms to support the knowledge sharing in virtual communities.

2.2. Collaboration support and social networks

One of the important supports in collaboration systems is the awareness of collaborators and their actions (Carroll et al., 2003; Diamadis and Polyzos, 2004). For example, who are the collaborators? What do they know? And what are they doing now? Fischer et al. (2005) claimed that people's creativity is the result of interaction and collaboration with other individuals in a society. In other words, creativity is a social activity that embodies in collective knowledge and essential collaborators. This indicates that finding knowledge and finding collaborators are two important supports for knowledge sharing in collaboration systems.

Besides the supporting collaboration tools, one factor critical to the success of knowledge sharing in virtual communities is volunteered knowledge contribution, which

is influenced by the social relationships in a network. Social relationships are composed of nodes and ties. Nodes are the individual actors within the network, and ties are the links between the actors. There can be many kinds of ties between the nodes. In its most simple form, a social network is a map of all of the relevant ties between the nodes being studied.

Researchers have recognized that a broader sense of social network is a self-organized structure of people, information, and communities (Kautz et al., 1997; Raghavan, 2002). A social network can be modeled by a net structure consisting of nodes and edges. Nodes represent individuals or organizations. The edges connecting nodes are called ties, which represent the relationships between the individuals and organizations. The strength of a tie indicates how strong the relationship is. Many kinds of ties may exist between nodes. In this paper, we will address the knowledge relationship tie and the social relationship tie. Social networks can be represented as matrices; therefore, the properties of the social networks can be analyzed by graph theory.

Our social network-based P2P network is designed based on Hill and Dunbar's theory (Hill and Dunbar, 2002). They suggest that the maximum size of a social network averages around 150 individuals. Thus, we confine the number of nodes in our social network-based P2P network to around 150. Another good reason to justify the number of nodes in our P2P network is that small communities are more prone to collaborate than large ones.

The other two reasons to confine the number of nodes in our P2P network to around 150 on-campus nodes are concerns of copyright and message flooding. Firstly, to reduce the possibility of misusing copyrighted digital content that do not belong to the university (or is not subscribed to the university), we confine the users of our system to the nodes with on-campus IP address. The idea is similar to book publishers' policy that only the end users who search for information via on-campus IPs are allowed to access subscribed digital content. The other reason to set an upper bound of around 150 nodes is to avoid the so-called flooding phenomenon (Ravishankar, 2002). Flooding can be defined by the equation

$$\sum_{i=1}^{TTL} n^i = \frac{n(n^t - 1)}{n - 1}.$$

For an n-nary node in a P2P network, its centrality is n. TTL is the abbreviation of time to live, which is the period of time during which a request is valid. A request will become invalid when the TTL is expired. If the TTL is too short, a request may turn invalid before it can be satisfied. If the TTL is too long, a flooding phenomenon is prone to occur when the request cannot be satisfied in a short time. The values of n and TTL depend on the total number of nodes in a P2P network, thus the more nodes, the larger n and longer TTL, which will lead to potential flooding phenomenon. As a result, we confine the number of nodes

to 150 to avoid the flooding phenomenon in current experiment. In our research, the scope of a social network is limited to 156 peers in a university.

2.3. Peer-to-peer and instant messenger

Two of the very good examples of virtual communities are P2P and IM. P2P is a connected network of content repositories, which is referred to as one of the resources of accessible explicit knowledge. IM enables interactive group collaboration through which people can interactively communicate, discuss, and exchange ideas to derive tacit knowledge. The discussion and interaction details recorded in IM are considered as one of the resources of tacit knowledge.

Besides the well-known Gnutella protocol for file search and file sharing, many researchers have applied P2P to collaboration. An academic P2P framework- Edutella, was developed for learning communities (Brase and Painter, 2004; Nejdl et al., 2002). This framework uses the exchange of Resource Description Framework (RDF) metadata to enhance the description and discovery of learning resource. In contrast to conventional P2P keyword matching, Edutella proposed a semantic matching approach. Each peer on the Edutella P2P network maintains a data repository, which serves as a service registry containing metadata to facilitate the description and discovery of learning resource. Biström (2005) suggested using the P2P network in collaborative learning environment. In his study, Bistrom addressed certain essential requirements for successful collaborative learning, and how P2P networks should be developed with additional communication functionalities to better support collaborative learning. Eikemeier and Lechner (2003) proposed a P2P collaboration tool—iK[now], for domain specific ad-hoc collaboration. iK[now] enhanced a typical P2P platform with IM communication support in order to foster communication among P2P communities. iK[now] extended P2P search functionality to locate experts who are currently on the P2P network, and then establish an IM network among the experts. Groove (2006) is a commercial P2P collaboration support tool with which users can join various groups based on interest and expertise. The communication among users can be encrypted, but this results in computation overhead. Unlike most academic P2P platforms, users of Groove are not anonymous; they need to register and identify themselves, which also results in administration overhead.

P2P and social networks share many concepts in common. For example, they are both distributed networking structures; a peer in a P2P network can be viewed as an analog of a node in a social network; a link in a P2P can be viewed as an analog of a relationship tie in a social network. In contrast to most P2P searches that emphasize on search queries and protocols, our social network-based P2P search aims at reducing search time and decreasing message traffic by minimizing the number of messages circulating in the network.

The goal of this research is to overcome the aforementioned two barriers that hinder efficient and effective knowledge sharing. In the following two sections, we will address how to overcome the two barriers by classifying knowledge and defining its quality control mechanism, and classifying collaborators and defining their relationship ties. Based on the classifications and definitions, we present the implementation of our P2P and IM collaboration support environment and demonstrate how to utilize such environment to find relevant knowledge and right collaborators.

3. Finding relevant knowledge via P2P network

Bieber et al. (2002) pointed out that a virtual community's knowledge has both explicit and tacit components. Explicit knowledge is mostly in the forms of documents and tangible artifacts that can be expressed in words, languages, diagrams, and formulas (Polanyi, 1997). In contrast, tacit knowledge is mostly in the forms of concepts and intangible personal experiences. Explicit knowledge can be codified, represented and shared asynchronously via reading and studying documents, while tacit knowledge cannot be explicitly described and mostly can only be perceived in human minds and shared synchronously via discussion and mentoring (Gall, 1987). Examples of explicit knowledge include books, research papers, and video recordings. Examples of tacit knowledge include know how, cognitive skills, and beliefs. In this paper, we refer to content found by P2P as one of the resources of explicit knowledge, while the discussion and interaction details recorded in IM is one of the resources of tacit knowledge.

3.1. Architecture of P2P network

To facilitate knowledge discovery, we have constructed a P2P network. As shown in Fig. 1, each peer in the P2P network consists of two modules: Resource Module and Search Module. The Resource Module is designed to formally describe resources contained in a peer. The Search Module is responsible for generating users' search queries and processing their search requests.

The Resource Module contains several managers to organize and manage the resources kept in the peer. The Resource Manager is the coordinator, which handles all kinds of resources from various managers. These resources can be content, services, or other applications provided by the peer. The managers include the Content Manager that handles the content repository, the Ontology Manager that provides semantic metadata of contents, and the Annotation Manager that processes annotation imposed to the content.

The Search Module contains a Query Process Engine and a Resource Discovery Manager. The Query Process Engine is an interface designed to generate search request. If a user cannot specify a search request clearly, the Query Process Engine automatically generates one for him/her based on his/her surrounding context. The Resource Discovery Manager is designed to process search requests received from peers by providing a concept map to guide the search process. The concept map is derived from the
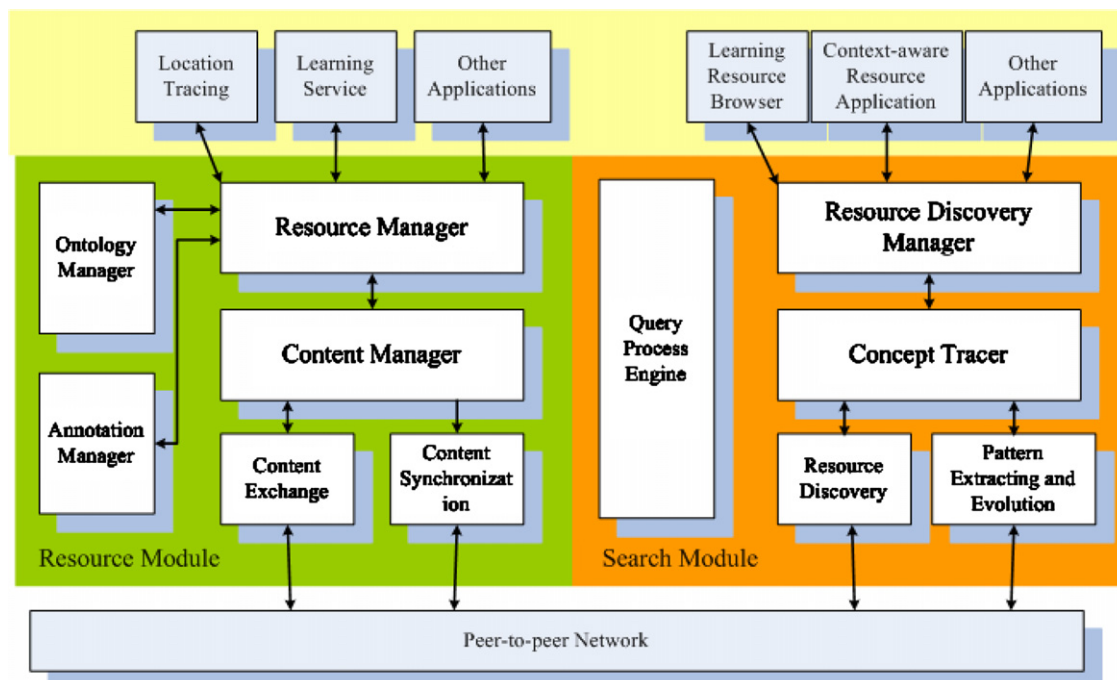

Fig. 1. System architecture of peer-to-peer network.

keyword and keyword thesaurus analyzed based on users' requests; the concept map is then extended or redrawn to match users' search requests.

3.2. P2P-based content discovery

For content discovery, our P2P network provides the functions of basic keyword search, keyword thesaurus and concept map-based search. Based on the content classifications and their quality control, the keyword thesaurus is used to extend search scope by finding more relevant keywords. In contrast, concept map based search is used to derive a more precise search scope by finding the most relevant keyword.

As shown in Fig. 2, the basic keyword search is enhanced by utilizing the keyword thesaurus. Our P2P network matches not only a single keyword but also a set of related keywords previously classified and saved in our content repository. The search results are shown in the main window along with the file name, type, size, state, and rating of the resource. For example, a keyword search of "New York Vacation" will derive a keyword thesaurus such as "New York City Life," "New York Travel", and even "New York Yankee".

For semantic search, we utilize the concept map approach to construct the relationship of a keyword concept and its related concept (Chau and Yeh, 2004). For example, if a user inputs the concept "Web services",

the system will produce a concept map with three nodes and two edges. One edge connects Web services and Semantic Web, and the other connects Web service and DAML-S. If the user continues to press the node "Semantic Web", the concept map will grow further to the one shown in Fig. 3. If the user then double clicks the node "XML", the system will proceed to search and come out with the results. On the upper left side of Fig. 3, there is the description of the concept. The lower left side of Fig. 3 shows the types of resources and their abstracts related to the concept, while the lower right side shows the details of the resource selected from the lower left side window.

3.3. Enhancing knowledge discovery with quality control

The motivation for participating in virtual communities is knowledge sharing (Wasko and Faraj, 2005). Without high-quality knowledge as its content, a virtual community cannot achieve its intended purpose of encouraging knowledge sharing. Content resource management can be used as a mean to achieve this goal.

To facilitate content resource management, we classify resources based on their knowledge domains and their quality. A peer can have a variety of resources and a single resource can be replicated in more than one peer. We utilize ACM Computing Classification System (1998) as our classification base of knowledge domain. In order to organize and provide better resource management, each
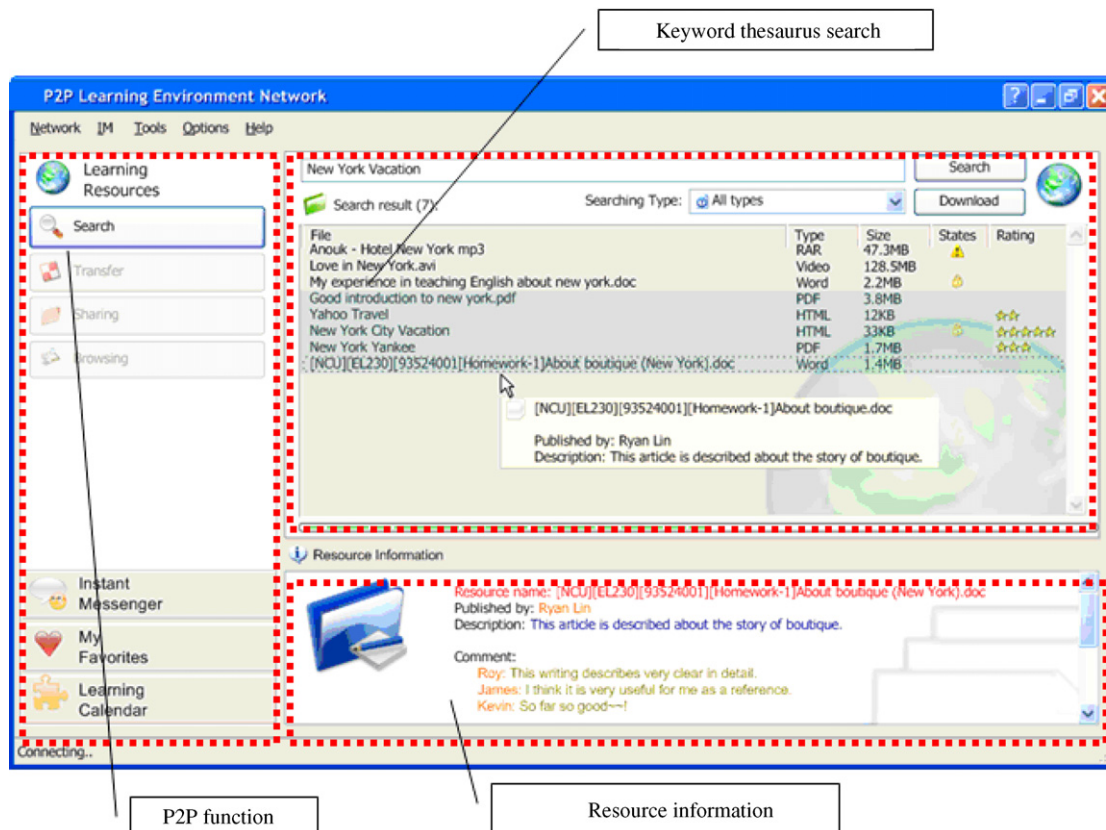

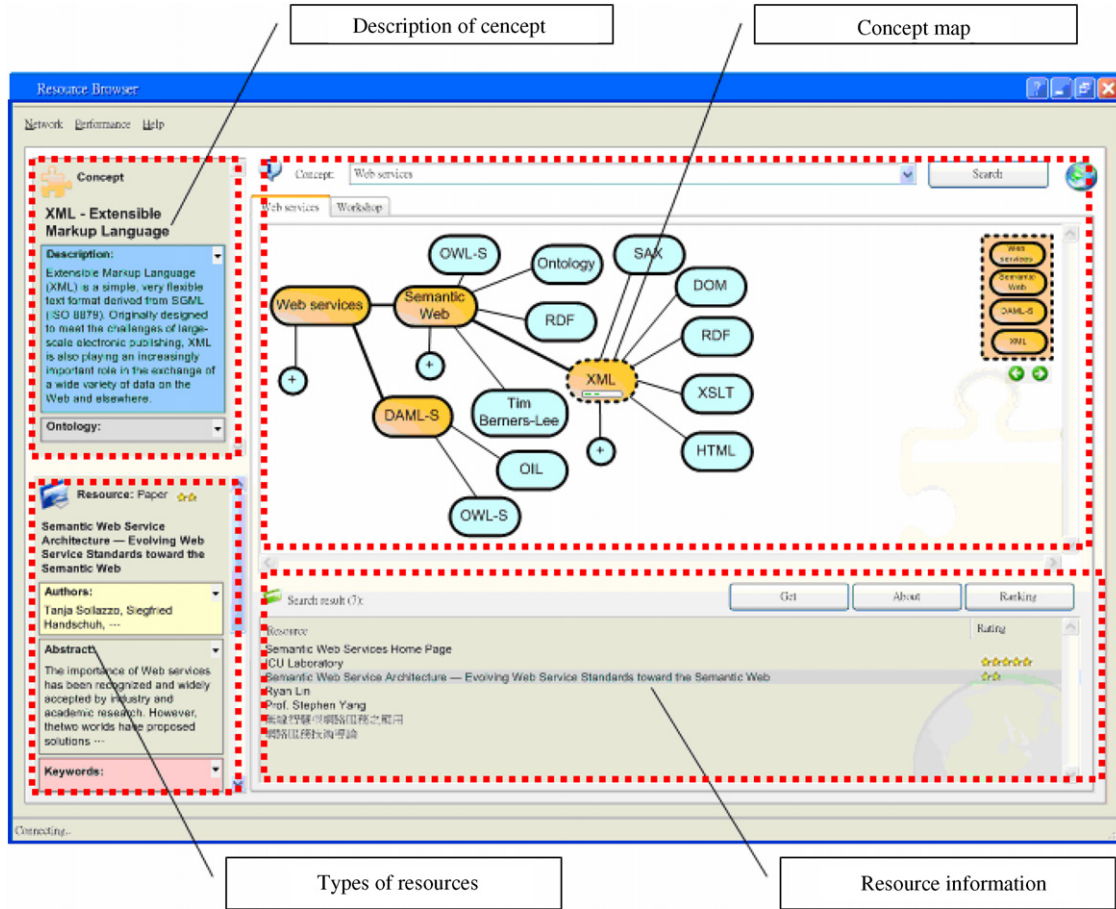Fig. 2. P2P network with keyword thesaurus search.

Fig. 3. P2P network with concept map.

peer in our P2P network needs to classify the content and evaluate its quality of content based on their reputation, accessed frequency (per day), and the matching degree to which the content classification conforms to knowledge domain. In our previous study, we have utilized a rating mechanism to evaluate Web services' reputation (Yang et al., 2006). In this study, we utilize the same rating mechanism to calculate a resource's reputation based on a 95% confidence interval in terms of probability (Mitchell, 1997), we will explain this in more details when we address a peer's reputation in later section.

The quality of resource i in knowledge domain j is given as

$$QoR_{(i,j)} = REP_{(i,j)} \times TOA_i \times MD_{(i,j)},$$

where QoR is the quality of a content resource, REP the reputation represents the rating of the resource. The higher the reputation, the better the rating of the resource, TOA the total number of times a resource is accessed per day. TOA represents the degree of popularity. The higher the TOA, the more popular the resource, MD the matching degree to which a content classification conforms to knowledge domain. The higher the MD, the better the matching.

The quality of a peer with respect to a certain knowledge domain, j, is the summation of quality of resource i over

the number of content resources, as given below

$$QoP_j = \frac{\sum_{i=1}^{NoR} QoR_{(i,j)}}{NoR},$$

where QoP the quality of a peer, NoR the number of content resources, which represents the volume of content in a peer.

The quality of a peer with respect to all knowledge domain contained in this peer is the average of QoP_j, which is given as follows:

$$QoP = \frac{\sum_{j=1}^{NoD} QoP_j}{NoD} = QoP$$
$$= \frac{\sum_{j=1}^{NoD} (\sum_{i=1}^{NoR} (REP_{(i,j)} \times TOA_i \times MD_{(i,j)})/NoR)}{NoD},$$

where NoD is the number of knowledge domains, which represents the scope of this peer's knowledge.

Based on the content classifications and the quality control, our P2P environment can be enabled to find more relevant and quality contents. In order to encourage peers to continually contribute quality content, each day our system will rank the peers according to their scores of reputation. The top 10% peers will be called "honorary peers". The underlying concept of "honorary peers" is that peers making more contribution deserve higher priority to

download content from the network. In other words, we value the honorary peers by allocating them higher download priority.

4. Finding relevant collaborators via social network

The key idea of our three-layer social network is illustrated in Fig. 4. For a given query requesting for collaborators with certain knowledge, a social network containing relevant collaborators with the requested knowledge will be dynamically constructed within the scope of a P2P network.

As shown in Fig. 4, the first layer is the P2P knowledge net (K-net), which is established to incorporate peers who own the requested knowledge into a pool of active peers. In this research, we confine the scope of the pool to a small-scale P2P network with about 50–150 peers within a university. A peer in the pool can be either a knowledge repository or a knowledgeable individual. A weighted edge between two peers is a knowledge relationship tie, which is used to measure the degree to which a peer's knowledge matches the other's query. For example, if a peer's query in such a P2P K-net (e.g., peer *Steve*) is requesting for peers with *Software Engineering* (SE) knowledge, a P2P K-net will be dynamically generated based on the query. As shown in Fig. 4, there are four edges dynamically generated on P2P K-net based on *Steve*'s query. Among them, the edge weighted (0.8) connecting *Chris* and *Steve* indicates that *Chris*'s knowledge matches best with *Steve*'s query. The edge weighted (0) connecting *Mary* and *Steve* indicates that *Mary*'s knowledge does not match *Steve*'s query.

P2P social net (S-net) is the second layer. A weighted edge connecting two peers is a social relationship tie, which is used to measure the degree of social familiarity between the two peers. Peers on K-net without the requested knowledge will be removed from S-net (e.g., *Mary*). Using the example shown in Fig. 4, *Steve* is more familiar with *Albert* than *Chris* because the social relationship tie between *Steve* and *Albert* is (0.9), which is greater than the social relationship tie between *Steve* and *Chris* (0.8). Based upon the generated S-net, an IM enabled group collaboration, shown on the third layer in Fig. 4, is invoked to help *Steve* communicate with the peers found in Layer-2 (*Chris* and *Albert*). Peers that appear on S-net and have negative relationships with the requester will be removed from the group collaboration (e.g., *Bob*).

This example shows that the essential challenge of constructing this three-layer social network is how to calculate knowledge relationship tie and social relationship tie.

4.1. Calculation of knowledge relationship tie

In our research, we consider a peer's knowledge domain, proficiency, and reputation of contribution as key indicators determining its capability to participate in collaborations. Therefore, as shown in K-net, we calculate a peer's knowledge relationship tie based on the three indicators.

As addressed by Barak and Rafaeli (2004), one of the most common hierarchies for ranking knowledge is the Bloom taxonomy (Bloom et al., 1956). In this study, we use the revised Bloom taxonomy in the form of matrix (Anderson et al., 2001) to classify a peer's domain knowledge and its proficiency in such a domain. As shown in Fig. 5(a) and (b), Bloom taxonomy is a matrix consisting of two dimensions: Knowledge dimension and Cognitive Process dimension. Knowledge dimension indicates the types of knowledge; Cognitive Process dimension indicates cognitive processing of knowledge. Each cell in a Bloom matrix is associated with a value ranging from 0 to 1,
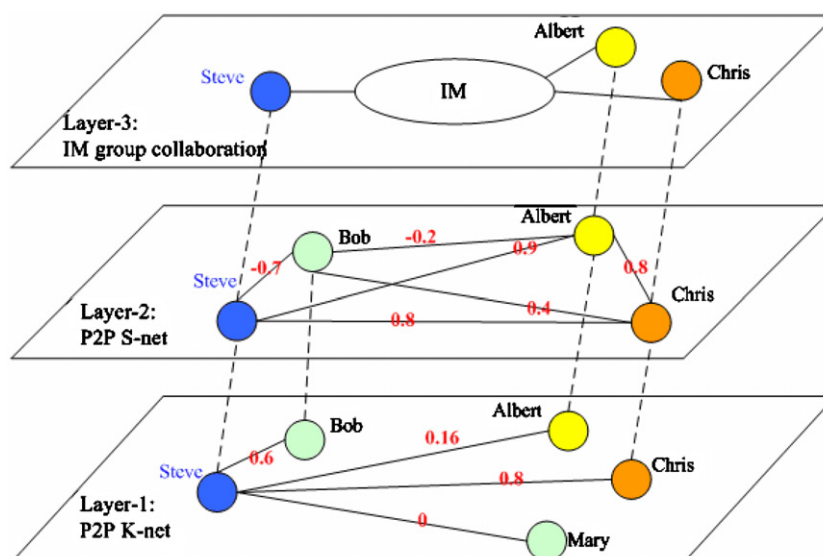

Fig. 4. A three-layer social network.

a

Knowledge dimension	Cognitive Process Dimensions					
	Remember	Understand	Apply	Analyze	Evaluate	Create
Factual knowledge	**0.9**	0.8	0.4	0.4	0	0
Conceptual knowledge	0.3	0.3	0.3	0.1	0	0
Procedural knowledge	0.6	0.5	0.3	0.2	0	0
Metacognitive knowledge	0	0	0	0	0	0

b

Knowledge dimension	Cognitive Process Dimensions					
	Remember	Understand	**Apply**	Analyze	Evaluate	Create
Factual knowledge	0.2	0.7	0.7	0.7	0.8	0.8
Conceptual knowledge	0.8	0.7	**1.0**	0.8	0.7	0.7
Procedural knowledge	0.2	0.2	0.3	0.2	0	0
Metacognitive knowledge	0	0	0	0	0	0

Fig. 5. (a) Example of *Albert*'s Bloom taxonomy matrix. (b) Example of *Chris*'s Bloom taxonomy matrix.

representing the level of proficiency. The values are manually given by Bloom experts after conducting evaluation tests for classifying individuals' proficiency toward Knowledge dimension and Cognitive Process dimension based on the Bloom taxonomy (Anderson et al., 2001; Bloom et al., 1956). For example, let Figs. 5(a) and (b) indicate peer Albert and peer Chris's knowledge proficiency regarding the knowledge domain of "Software Engineering", respectively. In Fig. 5(a), the cell [Factual knowledge, Remember] has a value (0.9), which indicates that Albert is good at memorizing factual knowledge about "Software Engineering". In Fig. 5(b), the cell [Conceptual knowledge, Apply] has a value (1.0), which indicates that Chris is excellent at applying conceptual knowledge of "Software Engineering".

4.1.1. Generation of Bloom matrix

Generally speaking, the domain experts and the Bloom experts are different individuals. In this paper, the domain expert is the course instructor who happened to be the first author of this paper, and the Bloom expert is the second author of this paper. The evaluation questions (in the format of multiple choice) used for the generation of a Bloom matrix are given by the domain experts while they prepare examination papers for quiz, midterm or final examinations. The collection of evaluation questions can be further organized into a test bank. At this question generation stage, the evaluation questions have not been classified into categories based on Bloom matrix (i.e., cells in a two-dimensional Bloom matrix, e.g., [Factual knowledge, Remember]). In this approach, all of the examination papers (containing the Bloom evaluation questions) can be generated from the test bank, thus there is no extra efforts for the instructor (domain expert). Students take the same copies of examination papers at the same time for both Bloom evaluation test and other regular examinations (the quiz, midterm and final examinations), thus there is no additional work for students.

The only overhead comes from the efforts of Bloom experts, who need to manually classify each evaluation question into an appropriate category. After this stage, each evaluation question is associated with a category. For example, a Bloom matrix (Fig. 5(a)) associated with the domain knowledge of software engineering. There are two examples of evaluation questions Q1 and Q2 as shown in the following. Both Q1 and Q2 are in the domain of software engineering. Q1 will be classified by a Bloom expert into the category of [Factual knowledge, Remember] because this question is used to evaluate whether students can memorize the software process model. Q2 will be classified into the category of [Conceptual knowledge, Apply] because this question is used to evaluate whether students can apply the concept of waterfall approach to certain real-world applications. The number of questions in each category may be different depending on the number of questions given by the domain expert. Thus, the final scores of the evaluation test need to be normalized and the results will range from 0 to 1. The calculation and the normalization of the values in a Bloom matrix are performed automatically.

In our approach, each student not only gets a test score, but also gets a Bloom matrix classification when he/she finishes each test. The values in some cells of the Bloom matrix will be zero if a student scores 0 point in related categories or if there is no evaluation questions in those categories.

Q1. What follows is NOT a software process model?
  (1) the waterfall approach
  (2) evolutionary development
  (3) verification and validation
  (4) formal transformation.
Q2. Given the following systems, which one is the most appropriate to be developed by applying the waterfall approach?
  (1) an anti-lock braking control system
  (2) a virtual reality system
  (3) an accounting system
  (4) an interactive system.

4.1.2. Knowledge relationship tie

Huhns and Buell (2002) pointed out that people are more likely to trust something proved. Similarly, people are more likely to collaborate with someone with good reputation. Many approaches have been proposed to calculate degree of trust (reputation) based on experiences, such as rating mechanism and referral networks. In our previous research, we developed a rating mechanism to evaluate Web services' reputations. The mechanism applied sampling of binomial probability to calculate a service's reputation based on a 95% confidence interval in terms of probability (Mitchell, 1997). The binomial distribution approximates the normal distribution when the number of samples is large enough. For the normal distribution, the confidence interval is 95% probabilities falling in the range of $\text{mean} \pm 1.96 \times \text{SD}$ (Standard Deviation) in compliance with the experience rule. We applied the same rating mechanism in this research to evaluate peers' reputation. Due to length limitation, readers please refer to (Yang et al., 2006) for detailed discussions about the calculations of trust and reputation.

Consider that peer i's query is requesting for peer j whose knowledge proficiency conforms to a requested knowledge domain k. Peer i's query can be calculated by

$$KQ_{(k)}(i,j) = K_{(k)}^{\text{proficiency}}(j)K_{(k)}^{\text{conformance}}(i)^{\text{T}},$$

$$KQ_{(k)}^{\text{serialized}}(i,j) = \sum_{m=1}^{4}\left(\sum_{n=1}^{6} KQ_{(k)}(m,n)\right),$$

$$K_{(k)}^{\text{tie}}(i,j) = KQ_{(k)}^{\text{serialized}}(i,j)K_{(k)}^{\text{reputation}}(j),$$

where $KQ_{(k)}(i,j)$ is a Bloom taxonomy matrix representing a query by peer i requesting for peer j whose knowledge proficiency conforms to a requested knowledge domain k.

$K_{(k)}^{\text{proficiency}}(j)$ is a Bloom taxonomy matrix representing peer j's knowledge proficiency with respect to a requested knowledge domain k.

$K_{(k)}^{\text{conformance}}(i)$ is a Bloom taxonomy matrix representing a conformance requirement requested by peer i to peers whose knowledge proficiency conforms to a requested knowledge domain k.

$KQ_{(k)}^{\text{serialized}}(i,j)$ is the serialization of $KQ_{(k)}(i,j)$, which is a 6 by 4 Bloom taxonomy matrix.

$K_{(k)}^{\text{tie}}(i,j)$ is a real number between 0 and 1 representing the knowledge relationship tie between peer i and peer j with respect to knowledge domain k.

$K_{(k)}^{\text{reputation}}(j)$ is a real number between 0 and 1 representing peer j's reputation regarding contribution to the requested knowledge domain k.

The value of $K_{(k)}^{\text{tie}}(i,j)$ indicates the degree of knowledge match up between peer j's knowledge and peer i's query; the higher the value, the stronger the tie. For example, consider a peer, *Steve*, who requests for peers who have knowledge "*Software Engineering*" and the proficiency of applying conceptual knowledge of "*Software Engineering.*" *Steve*'s request can be denoted as $K_{(\text{SE})}^{\text{conformance}}(Steve)$. Based

on the aforementioned equations and the example shown in Figs. 4 and 5, we found two peers, *Albert* and *Chris*, whose knowledge relationship ties, $K_{(\text{SE})}^{\text{tie}}(Steve, Albert)$ and $K_{(\text{SE})}^{\text{tie}}(Steve, Chris)$ are greater than zero, which means both *Albert* and *Chris* conform to *Steve*'s request in terms of knowledge relationships. Moreover, since $K_{(\text{SE})}^{\text{tie}}(Steve, Chris)$ is greater than $K_{(\text{SE})}^{\text{tie}}(Steve, Albert)$, *Chris* is more knowledgeable than *Albert* in terms of meeting *Steve*'s request of "Applying conceptual knowledge of *Software Engineering.*" The detail calculation of knowledge relationship tie is illustrated as follows:

Let *Steve*'s request be $K_{(\text{SE})}^{\text{conformance}}(Steve)^{\text{T}}$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^{\text{T}}.$$

Let *Albert*'s knowledge proficiency, $K_{(\text{SE})}^{\text{proficiency}}(Albert)$ be the matrix shown in Fig. 5(a), then

$KQ_{(\text{SE})}(Steve, Albert)$

$= K_{(\text{SE})}^{\text{proficiency}}(Albert)K_{(\text{SE})}^{\text{conformance}}(Steve)^{\text{T}}$

$$= \begin{bmatrix} 0.9 & 0.8 & 0.4 & 0.4 & 0 & 0 \\ 0.3 & 0.3 & 0.2 & 0.1 & 0 & 0 \\ 0.6 & 0.5 & 0.3 & 0.2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^{\text{T}}$$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

After serialization

$KQ_{(k)}^{\text{serialized}}(Steve, Albert)$

$$= \sum_{m=1}^{4}\left(\sum_{n=1}^{6} KQ_{(k)}(m,n)\right) = 0.2.$$

Let *Albert*'s knowledge reputation be $K_{(\text{SE})}^{\text{reputation}}(Albert) = 0.8$

$K_{(k)}^{\text{tie}}(Steve, Albert)$

$= KQ_{(k)}^{\text{serialized}}(Steve, Albert)K_{(SE)}^{\text{reputation}}(Albert)$

$= 0.2 \times 0.8 = 0.16.$

As a result, the knowledge relationship tie *between Steve* and *Albert* is 0.16.

Let *Chris*'s knowledge proficiency, $K_{(\text{SE})}^{\text{proficiency}}(Chris)$ be the matrix shown in Fig. 5(b), then the calculation of the degree to which *Albert*'s proficiency conforms to *Steve*'s

request is denoted as

$KQ_{(SE)}(Steve, Chris)$

$= K_{(SE)}^{proficiency}(Chris)K_{(SE)}^{conformance}(Steve)^{T}$

$$= \begin{bmatrix} 0.2 & 0.7 & 0.7 & 0.7 & 0.8 & 0.8 \\ 0.8 & 0.7 & 1 & 0.8 & 0.7 & 0.7 \\ 0.2 & 0.2 & 0.3 & 0.2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^{T}$$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

After serialization

$$KQ_{(k)}^{serialized}(Steve, Chris) = \sum_{m=1}^{4}\left(\sum_{n=1}^{6} KQ_{(k)}(m,n)\right) = 1.0.$$

Let *Chris*'s knowledge reputation be $K_{(SE)}^{reputation}(Chris) = 0.6$, then

$K_{(k)}^{tie}(Steve, Chris)$

$= KQ_{(k)}^{serialized}(Steve, Chris)K_{(SE)}^{reputation}(Chris)$

$= 1.0 \times 0.6 = 0.6.$

As a result, the knowledge relationship tie *between Steve and Chris* is 0.6.

In summary, this example shows that *Chris* is more knowledgeable than *Albert* in terms of helping *Steve* apply conceptual knowledge of "*Software Engineering*".

4.2. Calculation of social relationship tie

The social relationship tie indicates the degree of social familiarity between pairs of peers on the S-net. For a pair of peers, as denoted by peer i and peer j, the social relationship tie between them is the product of their social familiarity and social reputation.

$S^{tie}(i,j) = S^{familiarity}(i,j) \times S^{reputation}(j),$

where $S^{tie}(i,j)$ is the social relationship tie between peer i and peer j, $S^{familiarity}(i,j)$ is the social familiarity between peer i and peer j, and $S^{reputation}(j)$ is peer j's social reputation.

Social familiarity indicates the level of familiarity ranging from casual to close. Each peer existing in a social network needs to specify its social familiarity with a new peer first connecting to the social network, by filling forms and answering questionnaires. If an existing peer does not specify social familiarity with a new peer, the default value is zero meaning that there is no relationship between them. A new peer would presumably initially have a social familiarity score of 0 with all other existing peers in the social network. As it interacts with others the score will

change depending on how successful the interaction is. The scores of social familiarity can be updated manually as peers interact with others. Peers in the social network can explicitly update their social familiarity by changing the scores of social familiarity. Social familiarity can exhibit different levels of familiarity relationships, such as friends, teammates, organization colleagues, or virtual community members. Meanwhile, social familiarity can be positive or negative values ranging between −1 and 1, indicating bad or good relationships. In our experiment, we found that it is very difficult to obtain negative reports from testers. This phenomenon is more obvious when testers realize that the information will be used for the calculation of social relationships. Besides the fear of retaliation, people are prone to build good reputation for popularity, thus most of them would rather give a neutral response (the default value of familiarity is 0) than disclose a negative relationship. To perform quantitative analysis, we define social familiarity between peers i and j as follows:

$S^{familiarity}(i,j) = 0$, if there is no relationship between peer i and peer j.
$S^{familiarity}(i,j) = 0.8-1.0$, if peer i considers peer j a friend with positive relationship.
$S^{familiarity}(i,j) = 0.5-0.7$, if peer i considers peer j a teammate with positive relationship.
$S^{familiarity}(i,j) = 0.3-0.4$, if peer i considers peer j an organization colleague with positive relationship.
$S^{familiarity}(i,j) = 0-0.2$, if peer i considers peer j a virtual community member with positive relationship.
$S^{familiarity}(i,j) = 0$ to -0.2, if peer i considers peer j a virtual community member with negative relationship.
$S^{familiarity}(i,j) = -0.3$ to -0.4, if peer i considers peer j an organization colleague with negative relationship.
$S^{familiarity}(i,j) = -0.5$ to -0.7, if peer i considers peer j a team-mate with negative relationship.
$S^{familiarity}(i,j) = -0.8$ to -1.0, if peer i considers peer j a friend with negative relationship.

Each peer has a social reputation which is the product of the peer's social rating (Yang et al., 2006) and the average of the peer's social familiarities. Social rating indicates the degree of popularity of a target peer. It is an average rating given by other peers who know the target peer. Social reputation represents the degree of confidence other peers have in a target peer. The social reputation of peer j is computed as follows:

$S^{reputation}(j) = AVG[S^{familiarity}(j, NoP(j))] \times S^{rating}(j)$

$$= \frac{\sum_{m=1}^{M} m \in NoP(j)[S^{familiarity}(j,m)]}{M} \times S^{rating}(j),$$

where $AVG[S^{familiarity}(j, NoP(j))]$ is an average value of peer j's social familiarities, $S^{rating}(j)$ is peer j's social rating, $NoP(j)$ is a set of peers connected to j. The number of peers in the $NoP(j)$ is M.

Please refer to Fig. 4 where you can find all the values of the $S^{\text{familiarity}}$ that are needed to follow the following examples. The detail calculation of social relationship tie is illustrated as follows:

(1) Let *Albert* be a good friend of *Steve*, $S^{\text{familiarity}}(Steve, Albert) = 0.9$.
(2) Let *Albert*'s social rating be 0.6.
(3) Let *Albert* be directly connected to three peers (*Steve*, *Chris*, and *Bob*).

Then *Albert*'s social reputation is

$S^{\text{reputation}}(Albert)$

$$= \frac{\sum_{m1}^{3} m \in NoP(Albert)[S^{\text{familiarity}}(Albert, NoP(Albert))]}{3} \times S^{\text{rating}}(Albert)$$

$$= \frac{[S^{\text{familiarity}}(Albert, Steve) \times S^{\text{familiarity}}(Albert, Chris) \times S^{\text{familiarity}}(Albert, Bob)]}{3} \times S^{\text{rating}}(Albert)$$

$$= \frac{0.9 + 0.8 + (-0.2)}{3} \times 0.6 = 0.3.$$

So the social relationship tie between *Steve* and *Albert* is

$S^{\text{tie}}(Steve, Albert)$

$\quad = S^{\text{familiarity}}(Steve, Albert) \times S^{\text{reputation}}(Albert),$

$S^{\text{tie}}(Steve, Albert) = 0.9 \times 0.3 = 0.27.$

(1) Let *Chris* be a good friend of *Steve*, $S^{\text{familiarity}}(Steve, Chris) = 0.8$.
(2) Let *Chris*'s social rating be 0.8.
(3) Let *Chris* be directly connected to three peers (*Steve*, *Albert*, and *Bob*).

Then *Chris*'s social reputation is

$S^{\text{reputation}}(Chris)$

$$= \frac{\sum_{m=1}^{3} m \in NoP(Chris)[S^{\text{familiarity}}(Chris, NoP(Chris))]}{3} \times S^{\text{rating}}(Chris)$$

$$= \frac{[S^{\text{familiarity}}(Chris, Steve) \times S^{\text{familiarity}}(Chris, Albert) \times S^{\text{familiarity}}(Chris, Bob)]}{3} \times S^{\text{rating}}(Chris)$$

$$= \frac{0.8 + 0.8 + 0.4}{3} \times 0.8 = 0.52.$$

So the social relationship tie between *Steve*'s and *Chris* is

$S^{\text{tie}}(Steve, Chris)$

$\quad = S^{\text{familiarity}}(Steve, Chris) \times S^{\text{reputation}}(Chris)$

$S^{\text{tie}}(Steve, Chris) = 0.8 \times 0.52 = 0.42.$

This example concludes that the social relationship tie between *Steve* and *Chris* is stronger than that between *Steve* and *Albert*. We can also infer that *Chris* is more likely to share his knowledge with *Steve*.

5. Experiments and discussions of interactive collaboration

We have conducted quantitative and qualitative experiments to evaluate the performance of our social network, P2P network, and IM group collaboration. A total of 156 undergraduate students (computer science major) participated in this experiment. They were required to install the social network-based P2P network and IM on their own computers. In order to establish the three-layer social networks, each student (peer) was asked to fill out forms and answer questions to help the system identify and calculate peers' knowledge and social relationship ties. Adopting the calculation presented in this paper, a social network-based P2P network and IM is established.

5.1. Experiment

To facilitate the experiment, we have developed a P2P network, called SOtella as shown in Fig. 6. SOtella is implemented based on open-source software, Edutella. In order to control the experiment scale and monitor SOtella's performance, we confine the search scope of SOtella to a small-scale network including about 150 peers within a university. Each peer in SOtella is associated with knowledge relationship tie and social relationship tie as we presented in the three-layer social networks.

We also constructed an IM-enabled group collaboration tool as shown in Fig. 7. Peers can communicate with collaborators that are able to fulfill their needs and improve collaboration. From the peers found on K-net and S-net as presented in Layer-1 and Layer-2, groups can
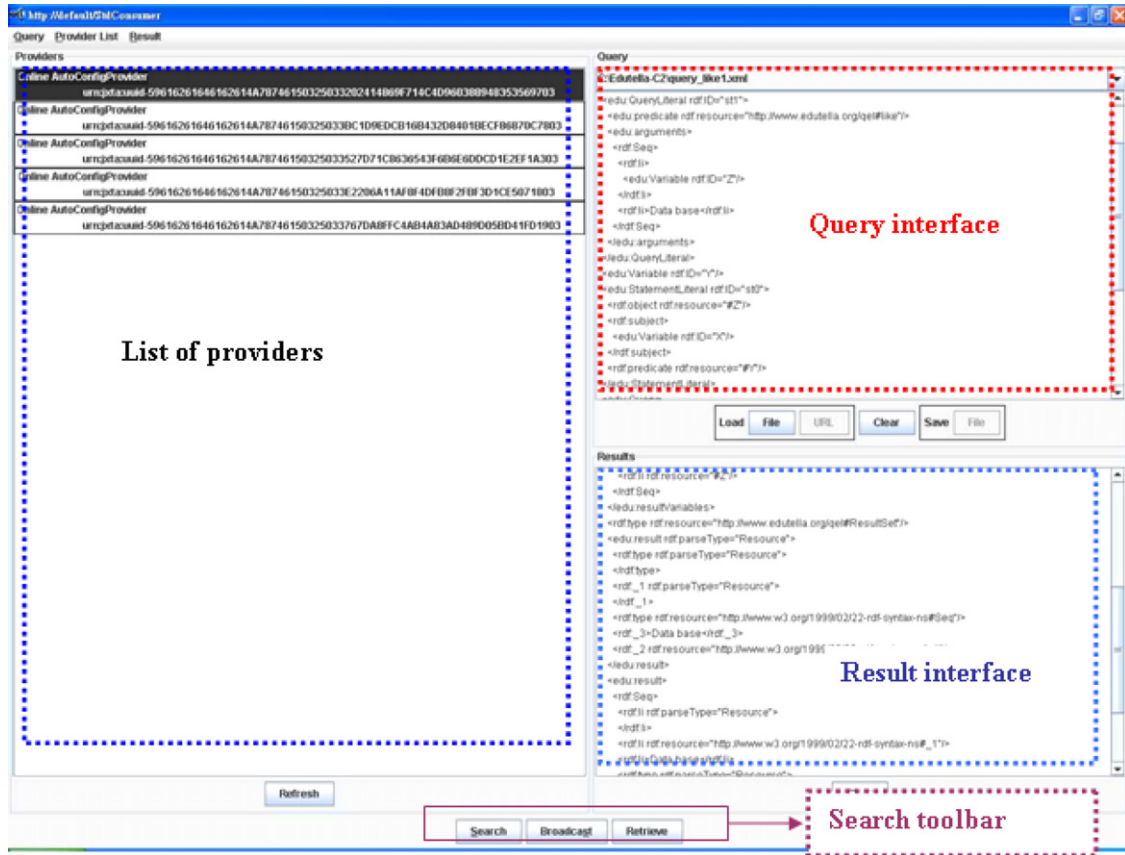
Fig. 6. Screen shots of Sotella.

be formed based on peers' knowledge relationship tie and social relationship tie. For example, the peer *Steve* can request for a group formation with four members who have the strongest knowledge relationship ties to the subject "Applying conceptual knowledge of *Software Engineering.*" For adaptive group formation, peers can provide feedback to the social network via the rating mechanism and improve the calculation of the two types of tie on our three-layer social networks.

5.2. Quantitative analysis

For quantitative performance evaluation, we measure two indexes: *Precision* and *Recall*. *Precision* is the fraction of the found peers that the requesting peers consider relevant; *Recall* is the fraction of the relevant peers that has been found. *Precision* and *Recall* are formally defined as follows:

$$Precision = \frac{|Ra|}{|A|}, \quad Recall = \frac{|Ra|}{|R|},$$

where A contains a set of discovered peers, $|A|$ is the number of peers in A, R contains a set of discovered peers who are considered relevant, $|R|$ is the number of peers in R, Ra contains the intersection of sets R and A. $|Ra|$ is the number of peers in Ra.

In this experiment, we used four kinds of knowledge domain as the search domains: Internet computing, Web computing, Mobile Internet, and Wireless Web. As indicated in Fig. 8, for the four given search domains, the *Precision* of social relationship tie (S-Tie) search outperforms the knowledge relationship tie (K-Tie). This indicates that the discovered peers are more relevant and they are more likely to be in the same social group because they have higher social relationship tie. In contrast, the *Recall* of K-Tie search outperforms S-Tie. This indicates that peers with the same knowledge domain will most likely be found at the same time because they have higher knowledge relationship tie.

5.3. Qualitative analysis

Three months after the system started to run, we conducted a survey with the same 156 students who had participated in this experiment. Each student was given a questionnaire containing 14 questions to verify their satisfaction rates regarding our system. For this measure, a five-point Likert scale with anchors ranging from *strongly disagree* (1) to *strongly agree* (5) was adopted. We then calculated the mean value and standard deviation of each question item. The result of the survey is summarized in Table 1.
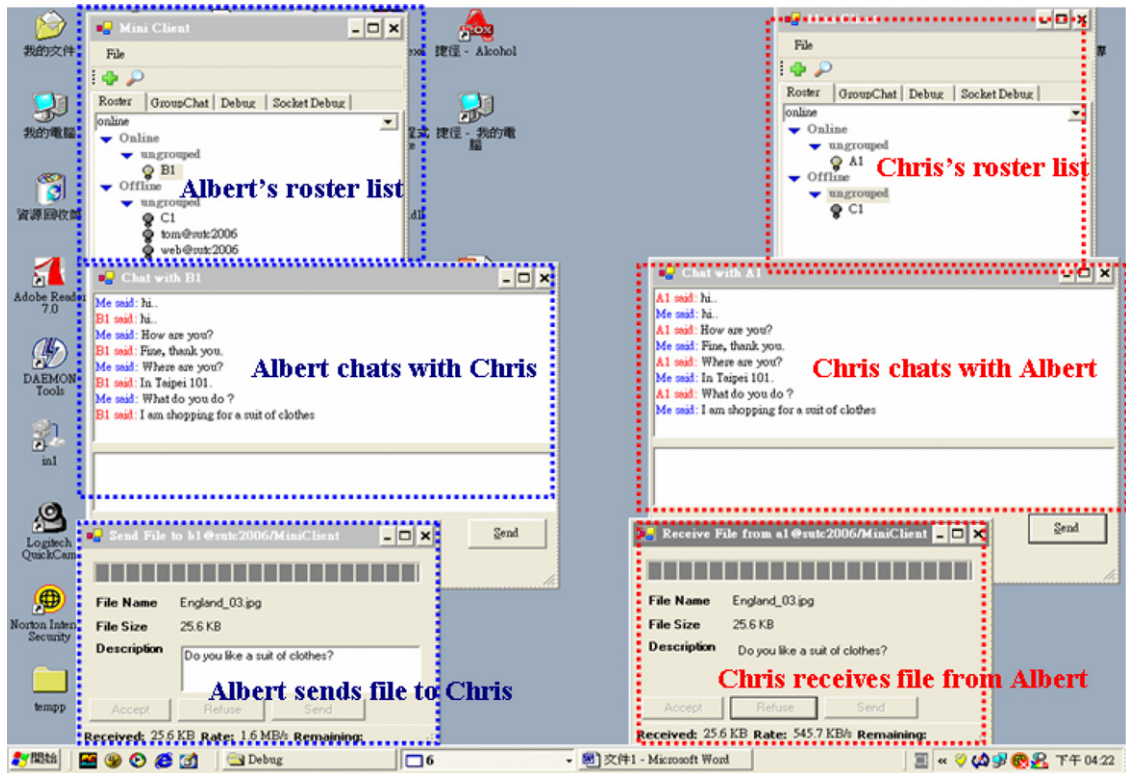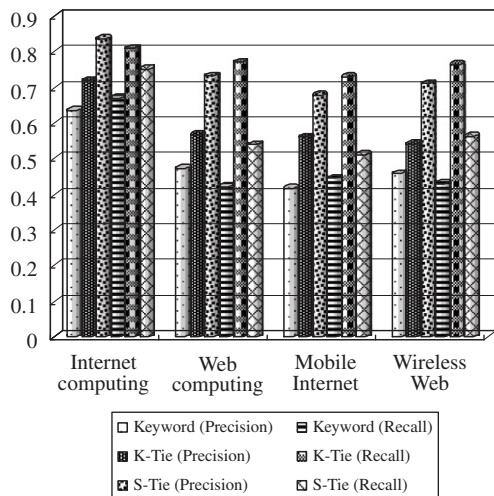
Fig. 7. IM-enabled group collaboration tool.

Fig. 8. Precision and Recall as performed by knowledge relationship tie (K-Tie) and social relationship tie (S-Tie).

Results of the survey revealed students' doubts about the stability of the P2P search results. They could not obtain the same search results from each trial even though they used the same search option. This is due to the decentralized nature of P2P network; it only searches for the peers currently on line. Another reason of non-deterministic search results is the precaution against opportunism. Some students are not willing to share the resources they obtained from the network because free-

riders are able to acquire the same resources when they used the same search option. From the survey results, we found that most of the students were not satisfied with the automatic group formation. Rather, they preferred to find their own collaborators even though they admitted that the group members picked by our system were knowledgeable. This observation suggests that in addition to knowledge competence, we should also consider users' social relationships when we form groups. However, the main purpose of such a system is to generate social networks for users who do not already know other users. If the users know each other only as peers on the network and their friendship judgments are based on previous interactions rather than personal relationships then the system is more useful and this problem is not an issue. In addition, most of the students stressed the importance of user interface design of group collaboration, and they expected better controls of dynamic group formation. We also perceived that students desired powerful collaboration instruments, such as voice communication and e-whiteboard for synchronous discussion and file sharing.

Our experiment confirms the effectiveness of our social network-based collaboration support to enhance knowledge sharing. Most students expressed their willingness to utilize the system for their daily studies. We also intend to use other test audiences (e.g., commercial users) to further verify the generalizability of the current findings. In fact, many part-time graduate students in our university have showed interests in adopting our system in their daytime

Table 1
Average scales of the survey of social network-based collaboration and group formation

No.	Questionnaire	Mean	SD
1	Are you interested in finding collaborators by using social network system?	3.45	1.34
2	Are you satisfied with the system performance in terms of connection time and searching time?	3.67	0.80
3	Can you find the same results every time you use the P2P network?	3.97	1.29
4	Do you keep your P2P network on line for most of the time?	4.47	0.75
5	Are you always willing to share the resources obtained from the P2P network?	3.65	1.75
6	Are you satisfied with the group formation provided by social network system?	3.24	1.13
7	Do you think your group member chosen is knowledgeable?	3.97	0.78
8	Do you think you can form better group by yourself?	4.56	1.10
9	Are you satisfied with the user interface design of group collaboration?	3.95	0.94
10	Is it easy to form a group?	3.67	1.28
11	Are you satisfied with the group discussion performance in terms of communication and synchronization?	3.56	0.84
12	Do you think it is important to connect to other IMs?	4.23	0.62
13	Do you think it is important to have voice-enabled discussion?	4.37	0.68
14	Do you think it is important to have e-whiteboard for synchronous file sharing?	4.18	0.61

Number of students = 156, Strongly disagree = 1, Strongly agree = 5.

jobs. Some even plan to conduct small-scale experiments in their departments, which could extend the scope of the experiment to broader communities of practice with real-world cases.

6. Conclusion and future research

The major contribution of this study is applying social network-based collaboration to enhance knowledge sharing in a P2P network by overcoming the difficulty in finding relevant content and collaborators to interact with. So far, all the graphical representation, equations, and examples are based on direct relationship for the ease of illustration. In our future research, we will continue to expand such mathematical model of social network for more advanced reasoning of social network analysis.

In the real-world society, mostly we consider individuals who have direct relationships such as "Steve is a friend of Irene". Nevertheless, we also need to take into account some cases of isolated individuals and individuals with indirect (or transitive) relationships. Some individuals could be isolated and are thus disconnected from the society. On the other hand, an individual may know the friends of his/her friends. For example, if Steve is a friend of Irene, and Irene is a friend of Albert, then we might derive a transitive relationship between Steve and Albert. To fully mimic the real-world society, a social network comprising both direct and transitive relationships should be modeled as a non-connected graph (network). So far, all graphical representation, equations, and examples presented in this paper are based on connected graph and direct relationships. In our future research, we will continue to expand the mathematical model from connected graph to non-connected graph, and continue to develop new computing mechanism for deriving transitive relationship.

We plan to continue our research work in the following two directions. First, it is a general problem for a social network to support the discovery, access, and sharing of knowledge. Users and other collaborators may have their own needs when they access subjects and discuss with others. We plan to conduct further study to investigate the special requirements from different social networks or social relationships in virtual communities. Second, to take into account the context of collaboration, we plan to explore applied social networks to combine collaborative domain and collaborators' ontology.

Acknowledgments

References

Aberer, K., Punceva, M., Hauswirth, M., Schmidt, R., 2002. Improving data access in P2P systems. IEEE Internet Computing 6 (1), 58–67.

ACM Computing Classification System, 1998. Available from 〈http://www.acm.org/class/1998/〉.

Alexander, P., 2005. Should Your Business Use Instant Messaging? 〈www.msnbc.com〉.

Anderson, L.W., Krathwohl, D.R., Airasian, P.W., Cruikshank, K.A., Mayer, R.E., Pintrich, P.R., Raths, J., Wittrock, M.C., 2001. A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. Longman, New York.

Barak, M., Rafaeli, S., 2004. On-line question-posing and peer-assessment as means for web-based knowledge sharing in learning. International Journal of Human–Computer Studies 61 (1), 84–103.

Bieber, M., Engelbart, D., Furuta, R., Heltz, S.R., Noll, J., Preece, J., Strhr, E.A., Turoff, M., van de Valle, B., 2002. Toward virtual community knowledge evolution. Journal of Management Information Systems 18 (4), 11–35.

Biström, J., 2005. Peer-to-peer networks as collaborative learning environments. In: Paper Presented at HUT T-110.551 Seminar on Internetworking.

Bloom, B.S., Engelhart, M.B., Furst, E.J., Hill, W.H., Krathwohl, D.R., 1956. Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook 1: The Cognitive Domain. Longmans Green, New York.

Brase, J., Painter, M., 2004. Inferring metadata for a semantic Web peer-to-peer environment. Educational Technology & Society 7 (2), 61–67.

Bruckman, A., 2002. The future of e-learning communities. Communications of the ACM 45 (4), 60–63.

Carroll, J.M., Neale, D.C., Isenhour, P.L., Rosson, M.B., McCrickard, D.S., 2003. Notification and awareness: synchronizing task-oriented collaborative activity. International Journal of Human–Computer Studies 58 (5), 605–632.

Chau, R., Yeh, C.H., 2004. Fuzzy Conceptual indexing for concept-based cross-lingual text retrieval. IEEE Internet Computing 8 (5), 14–21.

Chen, I.Y.L., Yang, S.J.H., 2006. Peer-to-peer knowledge sharing in collaboration supported virtual learning communities. In: Proceedings of the Sixth IEEE International Conference on Advanced Learning Technologies (ICALT), Netherlands, July 2006, pp. 807–809.

Churchill, E.F., Halverson, C.A., 2005. Social networks and social networking. IEEE Intelligent Systems 20 (5), 14–19.

Diamadis, E.T., Polyzos, G.C., 2004. Efficient cooperative searching on the Web: system design and evaluation. International Journal of Human–Computer Studies 61 (5), 699–724.

Edutella. Available from ⟨http://edutella.jxta.org/⟩.

Eikemeier, C., Lechner, U., 2003. Introducing domain specific ad-hoc collaboration: the peer-to-peer tool iKnow. In: Proceedings of the 12th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE'03).

Fischer, G., Giaccardi, E., Eden, H., Sugimoto, M., Ye, Y., 2005. Beyond binary choices: integrating individual and social creativity. International Journal of Human–Computer Studies 63 (4–5), 482–512.

Gall, M.D., 1987. Discussion method. In: Dunkin, M.J. (Ed.), The International Encyclopedia of Teaching and Teacher Education. Pergamon, Oxford.

Gnutella. Available from ⟨http://www.gnutella.com/⟩.

Grinter, R., Palen, L., 2002. Instant messaging in teen life. In: Proceedings of ACM 2002 Conference on Computer-Supported Collaborative Work (CSCW'02), New Orleans, LA.

Groove, 2006. Available from ⟨http://www.groove.net/⟩.

Herbsleb, J.D., Atkins, D.L., Boyer, D.G., Handel, M., Finholt, T.A., 2002. Introducing instant messaging and chat in the workplace. In: Proceedings of ACM Conference on Human Factors in Computing Systems CHI '02, Minneapolis, MN.

Hill, R., Dunbar, R., 2002. Social network size in humans. Human Nature 14 (1), 53–72.

Huhns, M.N., Buell, D.A., 2002. Trusted autonomy. IEEE Internet Computing, 92–95.

Isaacs, E., Walendoski, A., Whittaker, S., Schiano, J.D., Kamm, C., 2002a. The character, functions, and styles of instant messaging in the workplace. In: Proceedings of ACM 2002 Conference on Computer-Supported Collaborative Work (CSCW'02), New Orleans, LA.

Isaacs, E., Walendowski, A., Ranganathan, D., 2002b. Hubbub: a sound-enhanced mobile instant messenger that supports awareness and opportunistic interactions. In: Proceedings of ACM Conference on Human Factors in Computing Systems CHI '02, Minneapolis, MN, pp. 179–186.

Jin, Q., 2002. Design of a virtual community based interactive learning environment. Information Sciences 140 (1–2), 171–191.

Kautz, H., Selman, B., Shah, M., 1997. Referral Web: combining social networks and collaborative filtering. Communications of ACM 40 (3), 27–36.

Mitchell, T., 1997. Machine Learning. WCB McGraw-Hill, New York, pp. 128–141.

Nardi, B., Whittaker, S., Bradner, E., 2000. Interaction and outeraction: instant messaging in action. In: Proceedings of ACM 2000 Conference on Computer-Supported Collaborative Work (CSCW'00), Philadelphia, PA.

Nejdl, W., Wolf, B., Qu, C., Decker, S., Sintek, M., 2002. Edutella: a P2P networking infrastructure based on RDF. In: Proceedings of the 11th International World Wide Web Conference, May 2002, Hawaii, USA.

Polanyi, M., 1997. The tacit dimension. In: Prusak, L. (Ed.), Knowledge in Organizations. Butterworth-Heinemann, Woburn, MA.

Raghavan, P., 2002. Social networks: from the Web to the enterprise. IEEE Internet Computing 6 (1), 91–94.

Ravishankar, C.V., 2002. The Gnutella Protocol Specification Version 0.4. Available from ⟨http://www.clip2.com⟩.

RDF, Resource Description Framework. Available from ⟨http://www.w3.org/RDF/⟩.

Wachter, R.M., Gupta, J.N.D., Quaddus, M.A., 2000. It takes a village: virtual communities in supporting of education. International Journal of Information Management 20 (6), 473–489.

Wasko, M.M., Faraj, S., 2005. Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. MIS Quarterly 29 (1), 35–57.

Yang, S.J.H., Hsieh, J.S.F., Lan, B.C.W., Chung, J.Y., 2006. Composition and evaluation of trustworthy Web services. International Journal of Web and Grid Services 2 (1), 5–24.

Zhang, Y., Tanniru, M., 2005. An agent-based approach to study virtual learning communities. In: Proceedings of the 38th Hawaii International Conference on System Sciences.