Enhancing Virtual Learning Communities by Finding Quality Learning Content and Trustworthy Collaborators

Irene Y.L. Chen [1],Stephen J.H. Yang [2], Leon Y.Y Lin[3], Jeff J.S Huang [4]
Ching Yun University ,Taiwan[1] *National Central University, Taiwan*[2,3,4]
irene@mail.cyu.edu.tw[1], *jhyang@csie.ncu.edu.tw*[2], *darknesscount @gmail.com*[3],
Jeff-huang@csie.ncu.edu.tw[4]

Abstract

Virtual learning communities encourage members to learn and contribute knowledge. However, knowledge sharing requires mutual-trust collaboration between learners and contribution of quality knowledge. This task cannot be accomplished by simply storing learning content into repositories. It requires a mechanism to help learners find relevant learning content as well as knowledgeable collaborators to work with. In this paper, we present a peer-to-peer based social network to enhance the quality of e-learning regarding knowledge sharing in virtual learning communities. From the technical aspect, we will present advanced semantic search mechanism for finding quality content and trustworthy collaborators. From the social point of view, we will address how to support trustworthy social network to encourage learners to share. Results of this research demonstrate that applying such mechanism to knowledge sharing do improve the quality of e-learning in virtual learning communities.

Keywords：P2P network, Social Network, Virtual learning communities, knowledge sharing

1: Introduction

The explosion in Web based technology has led to increasing volume and complexity of knowledge which stimulates the proliferation of virtual learning communities (VLCs). VLCs are information technology based cyberspaces in which individuals and groups of geographically dispersed learners accomplish their goals of e-learning. One of VLCs' purposes is to encourage knowledge sharing so that valuable knowledge embedded in the network can be effectively explored. Most of the learners participate in VLCs with the expectations that they can acquire and share valuable knowledge to fulfill their needs.

The emergent VLCs over the past decade have stimulated research interests by academia and practitioners. Bruckman [5] found that the learning potential of the Internet technology can come from the peers and elders. Jin [10] provided a conceptual framework for the development of a prototype system of the virtual community based interactive learning environment. Wachter et al. [19] pointed out that an enhanced learning environment is possible only if one goes beyond mere on-line course delivery and creates a community of learners and other related resource groups. Wasko and Faraj [20] found that knowledge sharing has been a motivation for participation in virtual communities. In addition, many Web-based or agent-based models and software have been proposed to support interaction, discussion, and collaboration in VLCs [16] [26][11][3]. Prior studies have provided evidences demonstrating the importance of knowledge exchange in enhancing the learning performance. They also have called for the attention of providing mechanisms to support knowledge sharing in VLC environments.

However, knowledge sharing in some VLCs has not lived up to the expectation. Two barriers preventing efficient and effective knowledge sharing are: (1) the difficulty in finding quality knowledge, (2) the difficulty in finding trustworthy learning collaborators to interact with.

The objective and contribution of this research is applying peer-to-peer (P2P) based social network with trust management mechanism to overcome the aforementioned barriers. In order to help learners find quality content and trustworthy collaborators, we provide peer ranking mechanism and classify peers based on their content's quality. We have enhanced typical keyword search with keyword thesaurus search and semantic search to improve the performance of content discovery. We also have enhanced conventional on-line group discussion by finding trustworthy collaborators who are more willing to share.

2: Finding relevant and quality learning content

One of the motivation for participating VLCs is knowledge sharing. Without high-quality content, a VLC cannot achieve its intended purpose of encouraging knowledge sharing. The information areas for course materials, discussion forum, newsletters, and recommended-articles in a VLC's website constitute its knowledge/experience repository. Whether learners can effectively explore and exploit the knowledge within a VLC significantly influences the extent to which knowledge sharing can be achieved. High-quality content can attract learners participate in the knowledge activities and continually accumulate and enrich the knowledge in the repository, which in turn, facilitates knowledge sharing.

2.1: Knowledge domain and quality control

To facilitate content resource management, we classify resources based on their knowledge domains and their quality. We utilize ACM Computing Classification System 1998 (http://www.acm.org/class/1998/) as our classification base of knowledge domain. In order to organize and provide better resource management, each peer in our P2P network need to classify content and evaluate the quality of content based on their reputation, number of times having been accessed per day, and the matching degree by which the content classification conform to knowledge domain. 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: quality of a content resource

REP: reputation represents the rating of the resource, the higher it is, the better the reputation is.

TOA: the total number of times a resource has been accessed per day. TOA represents the degree of popularity, the higher it is, the more popular it is.

MD: the matching degree of how a content classification conforms to knowledge domain, the higher it is, the better is 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: 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:

$$QoP = \frac{\sum_{j=1}^{NoD} QoP_j}{NoD} =$$

$$QoP = \frac{\sum_{j=1}^{NoD} \left(\frac{\sum_{i=1}^{NoR} (REC_{(i,j)} \times TOA_i \times MD_{(i,j)})}{NoR} \right)}{NoD}$$

where

NoD: number of knowledge domains, which represents the scope of this peer.

2.2: P2P-Based Content Search

Based on the content classifications and their quality control, we now present our P2P environment and illustrate how to use it to find more relevant quality content. We have constructed a P2P (peer-to-peer) environment as shown in Figure 1, each peer in our P2P environment 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 user's search query and processing search requests received from other peers.

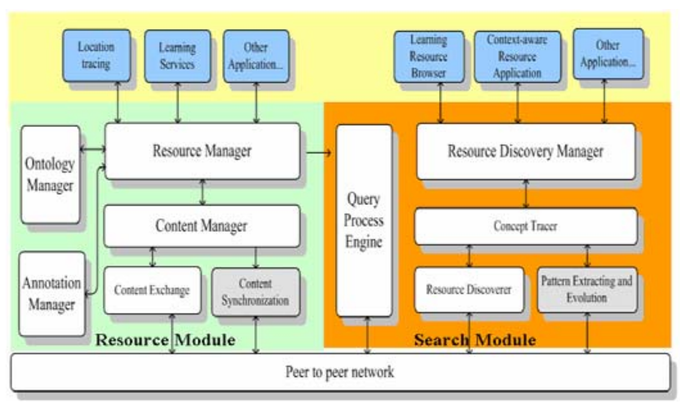

Figure 1. System architecture of peer-to-peer network

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 learning content, learning 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 users cannot specify search request clearly, the Query Process Engine automatically generates one for users based on users' surrounding context. The Resource Discovery Manager is designed to process search requests received from other peers by providing a concept map to guide the searching process. The concept map is derived from the keyword and keyword thesaurus analyzed from users' requests; the concept map is extended or redrawn to match users' search requests.

We have enhanced and implemented P2P in our previous researches [22][23]. For content discovery, our P2P 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.

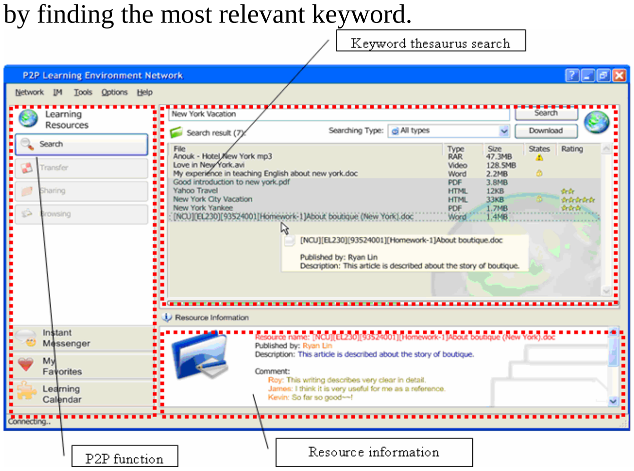

Figure 2. P2P network with keyword thesaurus search

As shown in Figure 2, the basic keyword search is enhanced by utilizing keyword thesaurus. Our P2P matches not only a single keyword but also a set of related keyword previously classified and saved by our content repository. The search results are shown in the main window along with resource's file name, type, size, state, and rating. 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 concept map approach to construct the relationship of a keyword concept and its related concept [7]. For example, if a user input the concept "Web services", the system will prompt a concept map with three nodes and two edges. One edge connects from Web services to Semantic Web, and the other connects from Web service to DAML-S. If the user continues to press the node of "Semantic Web", the concept map will grow further to the one shown in Figure 3. If the user then double clicks the node of "XML", the system will proceed to do the search and come out with the results. In the upper left hand side window of Figure 3 is the description of the concept, the lower left hand side window shows the types of resources and their abstracts related to the concept, and the lower right hand side shows the detailed information regarding the resource selected from the lower left hand side window.

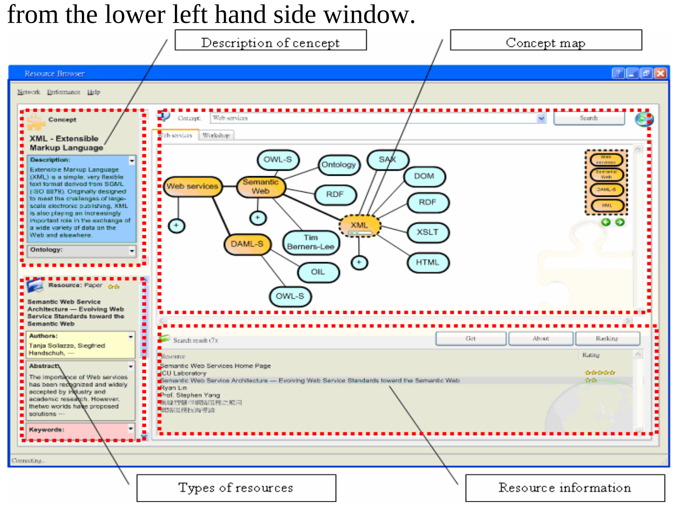

Figure 3. P2P with concept map

3: Find Trustworthy and Social Related Learning Collaborators

Social interaction ties are the structural links created through the social interactions between individuals in a network [25]. Prior studies suggested that an individual's centrality in an electronic network of practice can be measured using the number of social ties an individual has with others in the network [1]. Some academics addressed the importance of social interaction ties in knowledge exchange. For example, Tsai and Ghoshal [17] found that social interaction tie has positive impacts on the extent of inter-unit resource exchange. Wasko and Faraj [20] found that the centrality built up by the social interaction ties that any individual creates in a network significantly and positively impacts the helpfulness and volume of knowledge contribution.

A VLC's knowledge has both explicit and tacit components. The explicit knowledge can be easily browsed over the Internet. Yet, its implicit knowledge resides in the heads of the community members themselves and is shared with others through social interaction. Posting and responding to messages creates a social interaction tie between learners. The more social interaction ties a learner has with others, the easier he/she may acquire or share relevant knowledge to others. Therefore, social interaction ties are positively associated with the knowledge quality in a virtual learning community.

3.1: P2P-based Social Network Support

Social network [18] is used to describe a learner's social relationship with other learners in a VLC. We implement a hierarchical P2P-based social network support for knowledge sharing. As shown in Figure 4, a P2P knowledge network (K-network) is established to connect active learners into a pool of active peers, i.e., the learners (peers) that are online and available from the Internet. The pool can be an entire P2P network or a specific virtual community. Each peer (e.g., a ~ f) appeared in Figure 4 represents knowledge repository or a knowledgeable individual.

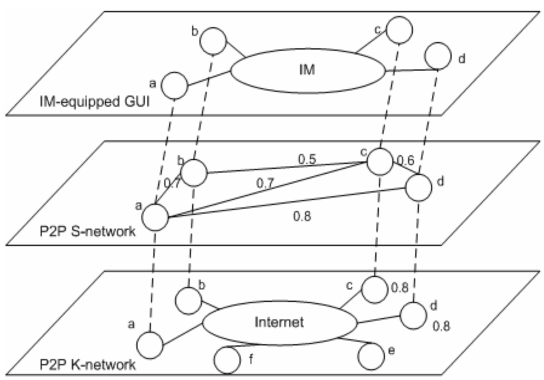

Figure 4. P2P-based social network support for knowledge sharing

If a peer in a P2P K-network, e.g., peer "*a*", requests a specific piece of knowledge, the social network support will dynamically generates a P2P S-network based upon the requester's social relationships with other peers who own the requested knowledge. As shown in Figure 4, peers that do not know about the relevant knowledge are filtered out and will not appear on the P2P S-network (e.g., peers *e* and *f*). Weighted edges in the generated S-network are called trust association (TA) to represent the levels that the peers can help the requestor (peer *a*). Using the example shown in Figure 4, peer *d* is more trustworthy than peer *c* because the TA between peers *a* and *d* is 0.8, which is greater than the TA between peers *a* and *c*, which is 0.7. Based upon the generated S-network, an IM-equipped (instant messenger) Graphical User Interface (GUI) is created to help the requestor discuss with the other peers in real time.

The essential technique in such social network support is how to construct a social network with trustworthy collaborators who can work with. The construction of such a social network is mainly based upon calculations of social network association between peers in the P2P environment. Each pair of peers is associated with two kinds of association- trust association and knowledge association, which will be addressed in the following subsections, respectively.

3.2: Trustworthy Social Network

The concern of trustworthiness in a social network can be classified into three levels- infrastructure, understanding and policy. Infrastructure is the first level, which focuses on keeping a trusted infrastructure. For example, the underlying software and hardware of a Web-based VLC must be trustworthy. The network should guarantee that network transmission is reliable and secure. Understanding is the second level. Huhns and Buell [9] pointed out that we are more likely to trust something if we understand it. One needs to confirm with confidence to the things he/she requested. An approach is to analyze experiences and estimate degree of trust based on one's past experiences [15], such as rating service, reputation mechanism, and referral network for exchanging experiences and reputation based on a third party certification group [8] or a peer-to-peer sharing mechanism [24]. Policy is the third level, which is used to describe requirements of trust, security, privacy and societal conventions to reach high-level trustworthy objectives [9][15]. In general, the policy provides many specific description-methods for requesting party to define what states and situations could accept. In other words, policy works like a rule set used to decide what behaviors and states could acquire authorizations. In this paper, we present an experience-based evaluation of learners' trustworthiness based on understanding and policy levels.

3.3: Calculation of Trust Association

Trust association is a confidence of how a pair of peers (learners) on the social network treats each other. It also indicates how a learner is associated with other learners directly connected to her on the social network. For a pair of learners who are socially related, as denoted by the requesting learner *i* and the requested learner *j*, the trust association between the two learners is denoted by TA(i,j). TA(i,j) indicates the confidence of trustworthiness of the requesting learner *i* to the requested learner *j*. TA(i,j) is used to determine whether the requested learner conforms to the requesting learner's requirements of trustworthy. The value of TA(i,j) is denoted by percentage, the higher the confidence is, the higher the trust association is. For example, if the value of TA(Chris, Albert) is 78%, which means the requesting learner Chris has 78% confidence that the requested learner Albert is trustworthy.

We utilize sampling of binomial probability to calculate the value of TA(i,j), based on a 95% confidence interval in terms of probability [13]. We first define the following terms

- S is a set of interaction instances representing samples of the requested learner's past interactions, $S = \{s_1, s_2, \ldots s_n\}$.
- Tr is a set of trust evaluation values containing past experience instance, and is denoted by $Tr = \{tr_1, tr_2, \ldots tr_n\}$.
- $Rating : S \rightarrow Tr \;\; Rating(s)$: The Rating function maps the interaction instance s to past experience instance, tr. In other words, the function associates past service instance with past experience instance, the experiences are collected by learners other than the requesting learner.
- $Accpet : Tr \rightarrow \{0,1\}$ A requirement hypothesis can be denoted as $Accpet$ function. The output of Accept function is 1 when past experience instance is accepted by the requesting learner, otherwise is 0.

$$\text{Accpet}(\text{tr}) \equiv \begin{cases} 1 & \text{Accept} \\ 0 & \text{otherwise} \end{cases}$$

Based on the usage of Large-Sample of Hypothesis for a Binomial Proportion to evaluate the simple error and true error of a hypothesis addressed in [12][13], the result of the hypothesis assesses the sample is a Boolean value (true or false). Thus we can see that the hypothesis assesses the sample as a Bernoulli trial and the distribution of Bernoulli trial is a binomial distribution. The binomial distribution approximates the normal distribution when the number of sample is enough. Simple error is correct rate in samples and true error is correct rate in population. We will get a confidence interval according to the simple error and the area of confidence interval represents a probability which true error fall in the interval. In the normal distribution, the true error is 95% probabilities falling within the range of $mean \pm 1.96 \times SD$ (Standard Deviation) in compliance with the experience rule. In other words, we can utilize the confidence interval to evaluate lowest true error of the evaluating hypotheses.

Let $Accpet$ function be the hypothesis and then we can evaluate the possible true error of the hypothesis based on the past instances S according to the Evaluating Hypotheses theory [13]. Whether the tr ($tr \in E$) is accepted by $Accpet$ is a binomial distribution which approximates the normal distribution when the number of samples is large enough. Thus we can utilize the normal distribution to calculate that the sample error closes with the true error. The true error is of 95% probabilities falling within a confidence interval, which will be approved as a trustworthy learner in the general application.

We define the confidence symbol as the lowest bound of the true error. The trust of service conforms to the request's requirement when the confidence is higher.

$$\hat{p} = \frac{1}{n} \sum_{s \in S} Accpet(Rating(s)), \; SD = \sqrt{\frac{\hat{p} \times (1 - \hat{p})}{n}},$$

$$z_{95\%} = 1.96$$

$$Confidence \equiv \max\{\hat{p} - z_{95\%} \times SD, 0\}$$

As the number of samples increases, the standard deviation decreases relatively and the confidence will be closer to the true error. For example, the past instances of a requested learner is denoted as S, and let $|S| = 256$. The requesting leaner proposes a Requirement Hypothesis $Accpet$. If the result of calculation is $\hat{p} = 0.6$, the confidence can be calculated from the following equation.

$$\hat{p} = \frac{1}{256} \sum_{s \in S} Accpet(Rating(s)) = 0.6, \; z_{95\%} = 1.96$$

$$Confidence = \hat{p} - z_{95\%} \times \sqrt{\frac{\hat{p} \times (1 - \hat{p})}{256}} \cong 0.6 - 0.060012 \cong 0.53998$$

The calculated confidence, i.e. TA(i,j) is 53.99%, which means the requesting learner has 53.99% confidence that the requested learner can meet the trustworthy requirement based on 95% confidence interval. Hence we can assert that the trustworthiness of the requested learner is 56.83% (53.99% over 95%) conforming to the requesting learner's requirements.

3.4: Calculation of knowledge association

Learners' knowledge association can be described by learners' domain of knowledge along with their proficiency pertaining to the corresponding domain. We use ACM Computing Classification System to classify domain of knowledge, and use Bloom taxonomy matrix [4][2] to classify learners' proficiency in that domain. As shown in Figure 5, a Bloom taxonomy matrix consists of two dimensions, Knowledge dimension and Cognitive Process dimension. Each cell in the matrix is associated with a value ranging between 0 and 1, indicating the level of proficiency. For example, given a learner with a Bloom taxonomy matrix, as shown in Figure 5, indicates the learner is good at memorizing and understanding factual and procedural knowledge pertaining to the corresponding domain.

4: Concluding Remarks and Future Research

The objective and contribution of this paper is to apply social network to enhance the quality of e-learning regarding knowledge sharing in virtual learning community by overcoming two barriers- the difficulty in finding quality knowledge and the difficulty in finding trustworthy learning collaborators. The experiment results of this research demonstrate that applying such mechanism to knowledge sharing do improve the quality of e-learning in virtual learning communities. We provide several avenues for further research. It is a general problem in social network to support the discovery, access, and sharing of knowledge. Learners and other collaborators have their own needs when they learn subjects and discuss with others. Further study is needed to investigate the special requirements from different learning context in virtual learning communities, such as for a given time, where the learners are? Who are the learners with? What are the learners doing? And what resources are available for learners? We will consider such context-aware learning in our future research.

References

[1] M. Ahuja, D. Galletta, and K. Carley, Individual centrality and performance in virtual R&D groups: an empirical study. Management Science, 2003, 49(1), pp. 21-38.

[2] L. W. Anderson, D. R. Krathwohl, P. W. Airasian, K. A. Cruikshank, R. E. Mayer, P. R. Pintrich, J. Raths, and M. C. Wittrock, “A taxonomy for learning, teaching, and assessing: a revision of Bloom’s taxonomy of educational objectives,”New York: Longman, 2001.

[3] N. Avouris, V. Komis, M. Margaritis, and G. Fiotakis, “An environment for studying collaborative learning activities,” Educational Technology & Society, 7 (2), 34-41.

[4] S. Benjamin, “ Bloom Taxonomy of Educational Objectives,” Handbook 1: Cognitive Domain. New York. Longmans Green, 1956.

[5] A. Bruckman, “The future of e-learning communities,” *Communications of the ACM*, 2002, 45(4), pp. 60-63.

[6] R. S. Burt, “*Structural Holes: The Social Structure of Competition*,” Harvard University Press, Cambridge, MA., 1992.

[7] R. Yeh Chau, and C.H, “Fuzzy Conceptual Indexing for Concept-based Cross-lingual Text Retrieval,” IEEE Internet Computing, 2004, 8(5), pp.14–21.

[8] T. Grandison, and M. Sloman, “A Survey of Trust in Internet Applications,” *IEEE Communications Surveys*, 2000, pp.2-16.

[9] M.N. Huhns, and D.A. Buell, “Trusted autonomy,” *IEEE Internet Computing*, 2002, pp. 92-95.

[10] Q. Jin, “Design of a virtual community based interactive learning environment,” *Information Sciences*, 2002, 140(1-2) , pp. 171-191.

[11] E. Matusov, R. Hayes, and M. J. Pluta, “Using discussion webs to Develop an Academic Community of Learners,” *Educational Technology & Society*, 2005, 8 (2), pp. 16-39.

[12] W. Mendenhall, and R.J. Beaver, “Introduction to Probability and Statistics,” *Duxbury Press*, 1999,pp. 442-446.

[13] T. Mitchell, “Machine Learning,” *WCB McGraw-Hill*, 1997, pp.128-141.

[14] R. Putnam, “Tuning in, Tuning out: The Strange Disappearance of Social Capital in America,” *Political Science and Politics*, 1995, pp.664-683.

[15] M.P. Singh, “Trustworthy Service Composition: Challenges and Research Questions,” *Proceedings of the Autonomous Agents and Multi-Agent Systems Workshop on Deception, Fraud and Trust in Agent Societies*, 2002,pp. 39-52.

[16] N. Taurisson, and P. Tchounikine “Supporting a Learner Community with Software Agents,” *Educational Technology & Society*, 2004, 7 (2), pp.82-91.

[17] W. Tsai, and S. Ghoshal, “ Social Capital and Value Creation: The role of intrafirm networks,” *Academy of Management Journal*, 1998, 41(4), pp. 464-476.

[18] Y. Upadrashta, J. Vassileva, and W. Grassmann, “Social Networks in Peer-to-Peer Systems,” In *Proceedings of the 38th Hawaii International Conference on System Sciences*, 2005

[19] R. M. Wachter, J. N. D. Gupta, and M. A. Quaddus, “ It takes a village: Virtual communities in supporting of education,” *International Journal of Information Management*, 2000, 20(6), pp. 473-489.

[20] M. M. Wasko, and S. Faraj, “Why should I share? Examining social capital and knowledge contribution in electronic networks of practice,” *MIS Quarterly*, 2005, 29(1), pp. 35-57.

[21] Wellman, B. “An Electronic Group is Virtually a Social Network,” In Sara Kieslered., Culture of the Internet, Lawrence Erlbaum, Hillsdale, NJ, 1997, pp. 179-205.

[22] S.J.H. Yang, “Context Aware Ubiquitous Learning Environments for Peer-to-Peer Collaborative Learning,” *Journal of Educational Technology and Society*, 2006. 9(1), pp.188-201.

[23] S.J.H. Yang, I. Chen, and N. Shao, “Ontological Enabled Annotations and Knowledge Management for Collaborative Learning in Virtual Learning Community,” *Journal of Educational Technology and Society*, 2004, 7(4), pp.70-81

[24] P. Yolum, and M.P. Singh, “An Agent-Based Approach for Trustworthy Service Location,” *Proceedings of 1st International Workshop on Agents and Peer-to-Peer Computing (AP2PC)*, 2002, pp. 45-56.

[25] G. Zhang, Q. Jin, and M. Lin, “A Framework of Social Interaction Support for Ubiquitous Learning,” In *Proceeding of the 19th International Conference on Advanced Information Networking and Applications (AINA’05)*, 2005.

[26] Y. Zhang, and M. Tanniru, “An agent-based approach to study virtual learning communities,” *Proceedings of the*

38th Hawaii International Conference on System Sciences, 2005.