Chih-Ming Chen*, Zong-Lin You "Community Detection with Opinion Leaders' Identification for Promoting Collaborative Problem-based Learning Performance", *British Journal of Educational Technology*, 2018. (Accepted) **[SSCI 收錄] (post print version)**

Community Detection with Opinion Leaders' Identification for Promoting Collaborative Problem-based Learning Performance

Chih-Ming Chen[1*], Zong-Lin You[2]

[1]Graduate Institute of Library, Information and Archival Studies, National Chengchi University, No. 64, Section 2, ZhiNan Road, Wenshan District, Taipei City 116, Taiwan, R.O.C.

E-mail: chencm@nccu.edu.tw

Tel: +886-2-29393091 Ext. 88024

Fax: +886-2-29384704

Chih-Ming Chen is currently a distinguished professor in the Graduate Institute of Library, Information and Archival Studies at National Chengchi University, Taipei, Taiwan. His research interests include digital library, e-learning, data mining, machine learning and intelligent agents on the web.

[2]Graduate Institute of Library, Information and Archival Studies, National Chengchi University, No. 64, Section 2, ZhiNan Road, Wenshan District, Taipei City 116, Taiwan, R.O.C.

E-mail: s891510@gmail.com

Tel: +886-2-29393091 Ext. 80640

Fax: +886-2-29384704

Zong-Lin You is currently a librarian in the National Chengchi University Libraries, Taipei, Taiwan. His research interests include e-learning and data mining.

Abstract

In the 21st century when knowledge-based economy is emphasized, the cultivation of autonomous learning and problem-solving capabilities presents the importance. With web-based collaborative problem-based learning, learners could more conveniently cultivate their problem-solving abilities through autonomous learning. Nevertheless, learners are often guided to solve a target problem by the information announced by teachers during the collaborative problem-based learning (CPBL) processes. Individual learners often could not effectively absorb such standard information, thus ignoring the important information from teachers. In the information communication theory, the two-step flow of communication through opinion leaders has been proved that it can better change audience's attitudes than the one-step flow of communication through mass media. This study thus employs the modularity Q function as the fitness function of genetic algorithm (GA) to optimally detect learning communities and uses PageRank measure to accurately find out community opinion leaders according to the social network interaction data of learners in the CPBL process. Based on quasi-experimental design, this study examines whether learners in the experimental group using the two-step flow of communication through opinion leaders to convey information for solving the target CPBL

missions could more significantly enhance web-based CPBL performance, social network interaction, and group cohesion than learners in the control group using the one-step flow of communication through teachers' information. Analytical results show learners in the experimental group remarkably outperform those in the control group on learning performance and peer interaction under a CPBL environment. Particularly, female learners in the experimental group notably outperform female learners in the control group on learning performance, while there is no significant difference in male learners between both groups. More importantly, learners in the experimental group present significantly higher group cohesion than those in the control group. This study confirms that using the two-step flow of communication instead of the one-step flow of communication traditionally used in web-based learning environments could significantly promote web-based CPBL performance, social network interaction, and group cohesion.

1. Introduction

With the development of information and communication technology (ICT), the computer-supported collaborative learning in past years expects to develop a web-based collaborative learning model better than the collaborative learning with face-to-face interaction in authentic learning environment. Koschmann, Kelson, Feltovich and Barrows (1996) stated that students with web-based collaborative problem-based learning (CPBL) outperformed those with traditional face-to-face collaborative problem-based learning on cognitive performance. This is because it could benefit real-time message communications, enhance learner interaction, and facilitate more effective collaborative learning than traditional face-to-face interaction. Aiming at health education, Naidu and Oliver (1996) designed a computer-assisted course with CPBL as the teaching strategy. The results showed that the integration of computer-based learning environment with CPBL indeed could enhance learners' thinking and problem-solving ability. Liu et al. (2010) confirmed that a web-based CPBL platform can promote elementary school students' research skills, decision-making, execution, and evaluation ability. Hung et al. (2012) discovered that experienced learners could effectively help novices' learning in the web-based problem-based learning environment to achieve collaborative learning. Generally, learners often accept the guiding information announced by teachers for the CPBL. However, such standard information not only could not be effectively absorbed by individual learners, but also might result in learners ignoring important information due to formal or serious information announcement type. On the other hand, in the information communication theory, the two-step flow of communication has been widely applied to information dissemination in community media, shopping choices in daily life, and elections (Choi, 2015; Zhang, 2015; Childers & Rao, 1992; Hong, 2016) and the two-step flow of communication through opinion leader in each community has been proved that it could better change audience's attitudes than the one-step flow of communication directly through mass media. Although it is common to apply the two-step flow of communication to communications, research on the application to teaching and even e-learning is little. It therefore induces the research motivation of this study to examine whether developing a two-step flow of communication model through opinion leaders in web-based CPBL communities could better enhance CPBL performance than the one-step flow of communication model through teachers' information posting. This study defined that the two-step flow of communication in an online learning context is that an online instructor's teaching guidance is indirectly conveyed to all members of each learning community through opinion leader, not is directly conveyed by the instructor.

Community detection aims to explore the structure-functionality relationship in complex networks, which involves two issues—the quantitative function for community as well as algorithms to discover communities (Ma, Wang, & Yu, 2018). Community detection algorithms can be divided into two categories, one is disjoint community detection, in which a node can be a member of only one community at most, and the other is overlapping community detection, in which a node can be a member of more than one community (Chintalapudi & Krishna Prasad, 2017). Aiming at community detection, a lot of past research proposed different algorithms that aimed at improving the ability of detecting meaningful communities, while keeping computational complexity as low as possible. Modularity, proposed by Newman and Girvan (2004) based on the graph theory, is one of the most representative quality criteria for measuring the quality of disjoint community detection. Thus, measuring community quality with modularity (Newman & Girvan, 2004) was used to detect CPBL learning communities in the study.

As the need for Web-based collaboration accelerates, it becomes increasingly important to understand small-group leadership in online environments (Gressick & Derry, 2010) because the leaders of small-groups have considerable influence on group process as well as outcomes (Bass, 1990). Facilitating collaborative learning performance by providing learners with collaboration scripts is regarded as a promising approach (Weinberger, Kollar, Dimitriadis, Mäkitalo-Siegl, & Fischer, 2009). Collaboration scripts, which consist of at least five components, including learning objectives, type of activities, sequencing, role distribution, and type of representation, can scaffold specific collaborative learning processes to facilitate social and cognitive processes of collaborative learning by shaping the way learners interact with each other (Kobbeet et al., 2007; Kollar, Fischer, & Hesse, 2006). Particularly, role distribution is considered in collaboration scripts because it is an important component affecting collaborative learning performance. Specially, opinion leaders in two-step flow of communication model are the brokers between mass media and the public as well as are the key person in information transfer. Opinion leader proposed by Lazarsfeld et al. (1994) was considered that the influence of mass media was not as good as opinion leaders; delivering messages to opinion leaders and having opinion leaders pass the messages to the followers could better affect the attitudes of ones in the same community. Katz and Lazarsfeld (1995) further analyzed the effect of information generated from opinion leaders on market consumption, fashion, public affairs, and movies. The results showed the higher effects of opinion leaders than mass communications on life issues of market consumption, movies, and fashion. Apparently, opinions proposed by an opinion leader were more easily promoted and accepted in a community to result in critical effects on information communication. This observation inspires us to explore how opinion leaders in collaborative learning communities affect collaborative learning performance in a CPBL environment.

To determine the learning communities in a CPBL environment, the social network interaction data in the CPBL processes of learners were utilized for finding out the community with which learners appear closer interactions, based on genetic algorithm based community detection scheme with modularity Q function (Newman & Girvan, 2004). Furthermore, PageRank measure (Page, Brin, Motwani & Winograd, 1998) was used for searching the most influential opinion leader in each learning community. This study aimed to examine whether the two-step flow of communication through opinion leader in each community could better enhance CPBL performance, social network interaction, and group cohesion than the one-step flow of communication through teachers in the CPBL processes.

2. Community Detection with Opinion Leader Identification for Collaborative Problem-based Learning

2.1 Collaborative problem-based learning (CPBL) system

The presented CPBL procedure involves four major learning stages for solving a target problem: 1) identifying the problem and situation; 2) designing the problem-solving method; 3) solving the problem; 4) reflecting on the process and its results. The four problem-solving learning stages were summarized as corresponding to "cognition", "action 1", "action 2", and "reflection" mental processes. The CPBL system used in this study provides a friendly user interface so that the instructor can conveniently design the problem-solving learning scaffolds based on the four-stage problem-based learning procedures to assist learners in completing problem solving procedures. Based on the designed learning scaffolds, the CPBL system asks learners to solve a semi-structured problem through higher-order thinking. A report concerning the solving of the target problem is completed by the writing of a report in each stage. Figure 1 shows an example of the user interface that the course instructor can use to plan the learning scaffolds in the first learning stage of a task related to the ‘‘global warming problem’’ in order to assist students’ learning of the experimental and control groups. Figure 2 shows an example of the user interface that the learner can use to write up a task report in the first learning stage of a task related to the ‘‘global warming problem’’ according to the learning scaffolds designed by the course instructor. The learning scaffolds provide students with the well-organized basic knowledge, designed learning guideline, gathered reference websites, gathered reference videos, or predesigned forms that students can easily follow or fill in. The learning scaffolds aim at guiding the learning directions of students and assisting them to learn in solving complex problems that would otherwise be beyond their current abilities.

The problem-solving learning report is completed according to learners’ answers. The next learning stage can be preceded after the learner completes the learning procedure at the previous learning stage and passes the instructor’s evaluation. When a learner accesses to the platform for problem-based mission learning, the system would automatically record the learner’s learning progress, discussion messages and interaction relationship, and assignment records. Opinion leaders used the instant message function on the CPBL system to send the messages related to the problem-based learning missions from the instructor to the members of the same learning community based on the two-step flow of communication, whereas the instructor directly used the bulletin board to announce the messages related to the problem-based learning missions to all the learners based on the one-step flow of communication.

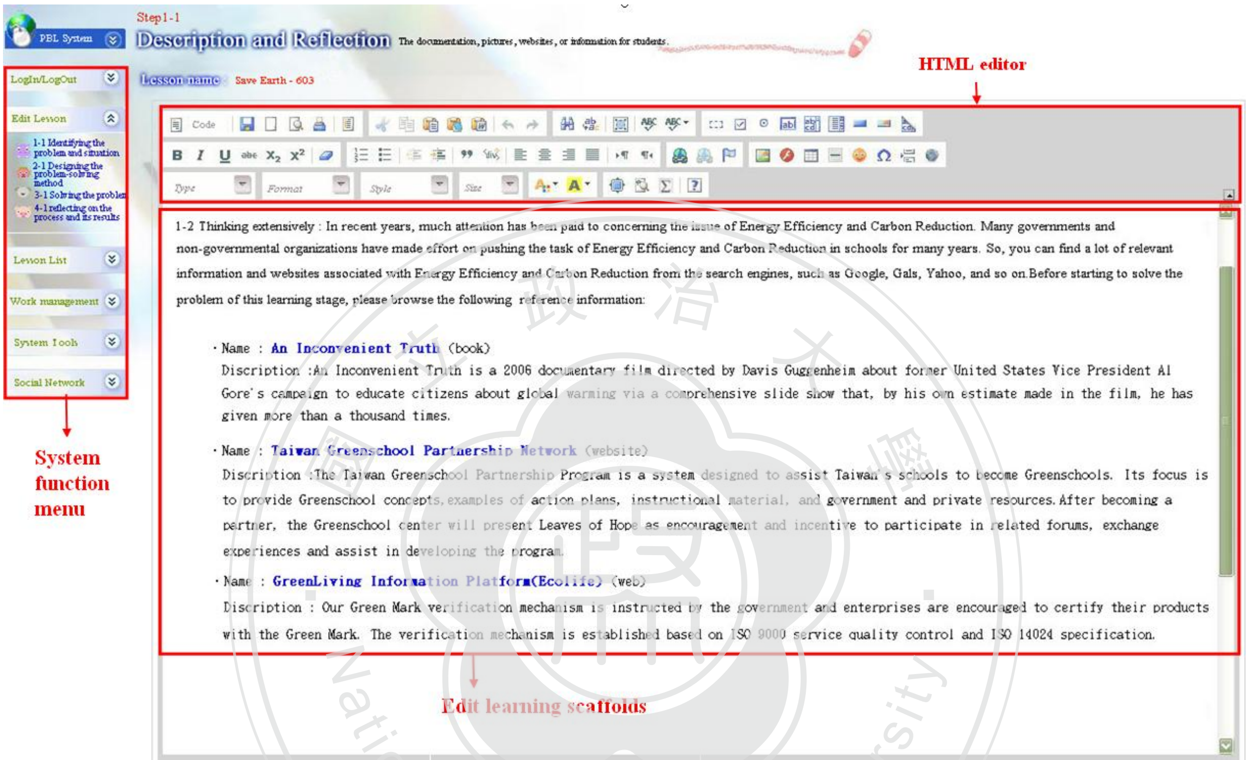


Figure 1. Teacher scaffolding design interface in the CPBL system

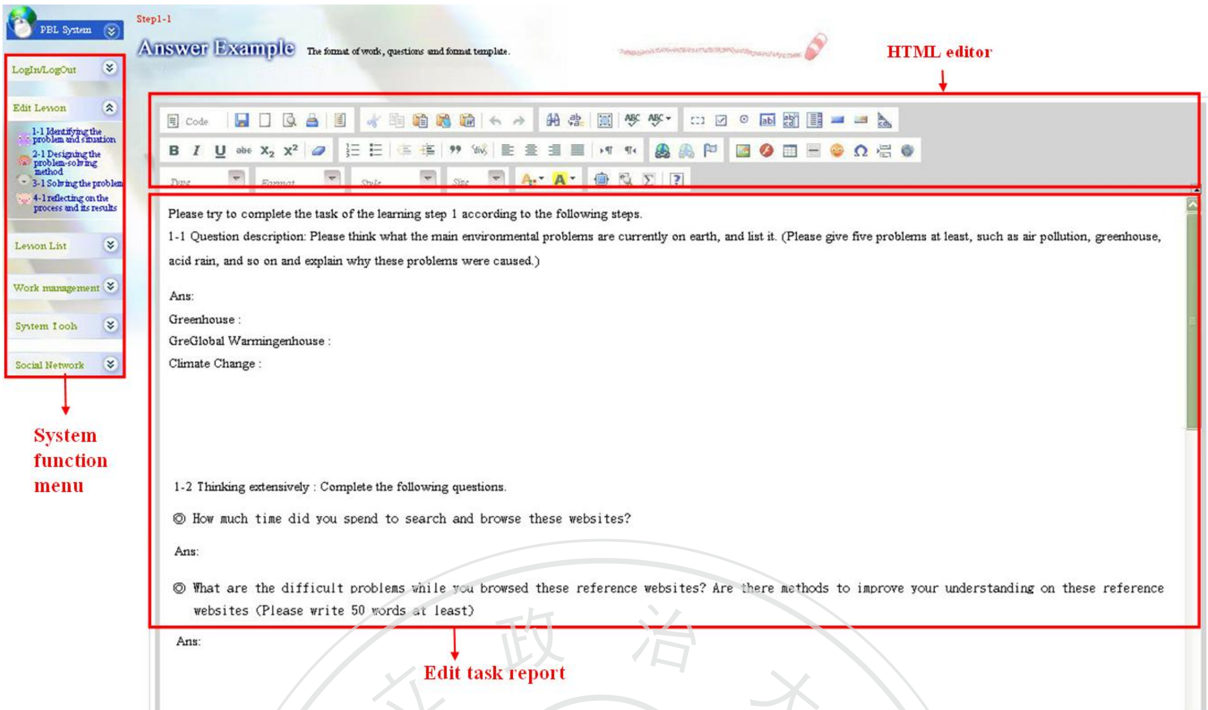


Figure 2. Learner answer scaffolding interface in the CPBL system

2.2 Community detection for collaborative problem-based learning

2.2.1 Modularity Q function

To confirm the quality of identified communities and solve the problem of proper number of community groups, Newman and Girvan (2004) proposed the evaluation function Q, called “modularity”. Modularity Q function was a numerical indicator for evaluating the appropriateness of specific community network division that it could be used for testing the community grouping. The larger modularity Q revealed the stronger community structure in the network. Therefore, modularity Q function could be used for searching the optimal community grouping in social networks. Newman & Girvan (2004) indicated that the concept of a community formed by a group of people was the connection in a group larger than the connection between groups, as the conceptual diagram of community detection in Fig. 3.

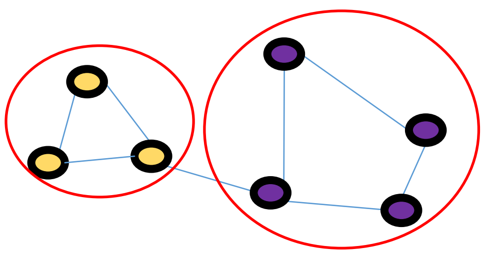


Figure 3. Community detection diagram

Based on above concept, Newman and Girvan (2004) proposed a modularity Q function to measure the quality of community and listed as Equation 1.

$$Q = \sum_{s=1}^{c}\left[\frac{l_s}{m} - \left(\frac{k_s}{2m}\right)^2\right] \quad (1)$$

where c stands for the number of community detected, m is the total number of edges in the entire network, l_s is the number of edges in the sth group, and k_s is the number of dimensions in the sth group.

From Equation 1, the modularity Q function appears between -1 and 1; the closer to 1 presents the better grouping quality of community. However, when the number is below 0, the grouping quality does not show reference value. The modularity Q function used as the fitness function of genetic algorithm was utilized in this study to detect communities with closer collaboration and interaction in the CPBL system. PageRank was further used to find out the opinion leaders of such communities for the reference of enhancing CPBL performance with two-step flow of communication.

2.3 Opinion leader detection based on Pagerank

There are four approaches to identifying opinion leaders, including sociometric methods, key-informant methods, self-designating methods, and observation (Rogers, 1995). Sociometric methods involve extensive analyses of leadership nominations within members of a peer group and were employed in the study. Pagerank measure, proposed by the founder of Google in 1998 (Page, Brin, Motwani & Winograd, 1998), is a kind of sociometric methods aiming to sequence pages for measuring the importance of a website based on link-based analysis. Google optimizes the search engine through Pagerank to promote the ranking of pages with higher importance and further enhance the correlation and quality of search results in order to meet users' needs. The higher Pagerank value shows the more popular of the page. Pagerank measure was applied in this study to search for opinion leaders in the community network, where learners are regarded as nodes, the interaction among learners is the link, and opinion leaders in communities with higher influence are found out according to interactions for two-step flow of communication in order to promote the CPBL performance.

3. Research Methodology

3.1 Experimental design

Two classes of Grade 4 students from an elementary school in Taoyuan City, Taiwan were recruited as the research subjects to precede a target problem-based learning mission on the CPBL system through problem discussion and interactions. Based on the quasi-experimental design, the two classes were randomly divided into the experimental group using two-step flow of communication and the control group applying one-step flow of communication to conduct the problem-based learning mission. The problem-based learning mission in the CPBL system were divided into four learning stages of “cognition”, “action 1”, “action 2”, and “reflection”. The instructor designed the problem-based learning mission with scaffolding support for the four stages in the back-end of CPBL and learners asynchronously preceded collaborative learning to solve the four-stage learning mission set by the teacher. The entire learning process lasted for 4 weeks. Namely, each learning stage lasted for a week.

The teacher’s evaluation at the first learning stage of “cognition” was used to test the difference in the prior knowledge of solving problem-based learning mission between two groups. The teacher then explained the problem-based learning mission and guided learners complete the four-stage problem-based learning mission. At the first learning stage of “cognition”, learners would interact and discuss with real-time messages on the CPBL system, which would collect the social interaction information among learners. After that, genetic algorithm matched with the modularity Q function as the fitness function was further utilized to determine the optimal community group, and PageRank measure was used to decide the opinion leader of each community. From the learning stages 2 to 4, the research subjects in two groups accepted the problem-based learning mission at each corresponding learning stage and completed such mission with different information communication models. The teacher’s evaluation scores on solving the problems in learning mission at the learning stages 2 to 4 were referred to the learning performance. The experimental design is further explained as follows.

(1) Pretest stage

The research participants in two groups were explained the experimental procedures and demonstrated the CPBL platform before the instruction experiment to ensure the smooth operation of the experiment. The research subjects would fill in the group cohesion scale before the experiment in order to understand the initial states of group cohesion in both groups.

(2) Learning stage using different information communication strategies for both groups

The research subjects in two groups would precede the four-stage problem-based learning mission related to the ‘‘global warming problem’’ on the CPBL platform for 4 weeks. Both the groups preceded free discussion on the CPBL platform, which would record the interactions, in the first week. From the second week, the modularity Q function matched with genetic algorithm was utilized for community detection of learners in the experimental group, according to the interactions on the platform. Besides, the community members with the highest PageRank values in all the communities detected from the experimental group would be regarded as opinion leaders, while the control group was not proceeded community detection for opinion leaders and all information was directly announced by the teacher to the entire class.

Different information communication strategies would be adopted for different groups from week 2 to week 4. The experimental group applied two-step flow of communication, where information is transferred to opinion leaders who then transfer to the other community members. One-step flow of communication model was utilized in the control group, where information is directly transferred to all learners by the teacher. Learners in two groups receive identical conditions, except for the communication model. In the learning stage, the opinion leaders in the experimental group got the teaching guidance including the hints and the reference materials of solving the target CPBL missions from the instructor via the instant message function, followed by searching more information related to the teaching guidance from the Internet, and then transferred the absorbed and integrated information to their peers in the same learning community by using the instant message function. If the opinion leaders could not fully understand the teaching guidance from the instructor, they could ask the instructor via the instant message function. Additionally, some students actively shared useful information getting from the Internet to their opinion leaders, and opinion leaders relayed the information to other group members to create a positive learning circulation. In other words, the opinion leaders in the experimental group not only play the role of information transformation, but also play the role of information filter and integrator. In contrast, the instructor used the bulletin board to announce the teaching guidance related to the problem-based learning missions to all the learners in the control group. Namely, the instructor did not have direct interaction with all the learners in the control group.

(3) Posttest stage

After the experiment, the research subjects of both groups were requested to fill in the group cohesion scale and some research subjects in the experimental group were invited to perform semi-structured interviews. The data were analyzed to deduce the difference in group cohesion between both groups and to understand how the opinion leaders in the experimental group facilitated knowledge sharing.

3.2 Research participants

Table 1 shows the number of research subjects including 49 students, 27 males and 22 females from two classes in an elementary school. One class was randomly assigned to the experimental group applying two-step flow of communication to perform the CPBL, and the remaining class was assigned to the control group using one-step flow of communication for the CPBL. The experiment was preceded on the CPBL platform, and opinion leaders in the experimental group were included at the learning stages 2, 3, and 4 for the information transfer with two-step flow of communication, while the control group utilized one-step flow of communication with the teacher posting messages on the CPBL platform.

Table 1. Statistics of research subjects

Group	Number	Male	Female
The control group with one-step flow of communication	25	13	12
The experimental group with two-step flow of communication	24	14	10
sum	49	27	22

3.3 Research tools

3.3.1 Social networks analysis tool

UCINET (University of California at Irvine Network) was used in this study for analyzing social networks measurement of degree centrality, betweenness centrality, and closeness centrality to judge whether two-step flow of communication could facilitate web-based collaborative learning interaction or not. In UCINET, the data in the social networks are stored, displayed, and described with matrix, and the social networks analysis results are visualized.

3.3.2 Group cohesion scale

The group cohesion scale used in this study contains social cohesion and task cohesion, which are revised by referring to the questions compiled by Zaccaro (1991) and Seibold & Kelly (1988). A total of 13 questions were developed for the scale, presenting favorable reliability (Crobach's alpha=.939, N=96) (see appendix).

4. Experimental Results

4.1 Analysis of learning performance between both groups

To avoid some other variables influencing the learning performance assessment, this study designed the same CPBL missions and conducted the same experimental period with four weeks for the two different learning groups. The results at the learning stage 1 were preceded the independent-samples *t* test to understand the difference in the prior knowledge of learners in both groups on solving the target CPBL missions between both groups. Table 2 shows the results. The results show that the initial problem-solving abilities between learners using two different communication models on the target CPBL missions do not appear significant difference ($t=-1.579$, $p=.121<.05$), i.e. the prior knowledge of learners in both groups is the same at the learning stage 1.

Table 2. Independent-samples *t* test of initial prior knowledge of learners between the experimental and control groups using different information communication strategies

Test item	Group	Number	Mean	Standard deviation	Mean-equivalent t test	
					t	Significance (two-tailed)
The performance of the learning stage 1	The experimental group	24	77.08	7.790	-1.579	.121
	The control group	25	74.00	5.577		

The number of learners in both groups passing the four stages and the pass rate are shown in Table 3. The pass rate of both groups at the learning stage 1 and stage 2 is 100%; the pass rate of the experimental and control groups at the learning stage 3 appears 100% and 76%, respectively; and, the pass rate of the experimental and control groups at the learning stage 4 shows 33.33% and 16%, respectively.

Table 3. The number of passed learners and the pass rate in the experimental and control groups in the four-stage problem-based learning

Group	Learning stage 1		Learning stage 2		Learning stage 3		Learning stage 4	
	No. of passed learners	Pass rate	No. of passed learners	Pass rate	No. of passed learners	Pass rate	No. of passed learners	Pass rate
The experimental group	24	100%	24	100%	24	100%	8	33.33%
The control group	25	100%	25	100%	19	76%	4	16%

Furthermore, the average performance of both groups at the learning stages 2, 3, and 4 were proceeded the independent-samples *t* test to evaluate the learning performance. Table 4 shows the results, where there is significant difference in the learning stages 2 and 3 between two groups, and the experimental group is higher than the control group. It was therefore inferred the higher learning performance of the experimental group than the control group. The average performance of the experimental group from the learning stages 2 to 4 is better than that of the control group, showing that the experimental group, with the influence of opinion leaders, outperforms the control group on learning performance.

Table 4. Independent-samples *t* test result of the learning performance for the experimental and control groups at various learning stages

Learning stage and group		Number	Mean	SD	*t*	Significance (two-tailed)
Learning stage 2	The experimental group	24	77.92	5.299	-3.535	.001
	The control group	25	72.60	5.228		
Learning stage 3	The experimental group	24	83.96	8.338	-4.353	.000
	The control group	25	57.16	29.580		
Learning stage 4	The experimental group	24	42.08	25.148	-.486	.629
	The control group	25	38.60	25.021		
Average performance of learning stages 2 to 4	The experimental group	24	67.99	9.272	-2.932	.006
	The control group	25	56.12	17.888		

Independent-samples *t* test was further applied to analyze the difference in learning performance between learners with different gender in both groups. Table 5 shows the results. The results reveal that female learners using two-step communication model through opinion leaders present better learning performance than female learners applying one-step flow of communication model through teacher's website announcement, but no significant difference appears between male learners in both groups. It is possibly because female learners are more easily affected by opinion leaders or peers than male learners.

Table 5. Independent-samples *t* test results of learning performance between learners with different gender in both groups

Learning stage and group		Number	Mean	SD	*t*	Significance (two-tailed)
Average performance of male learners in both groups from the learning stages 2 to 4	The experimental group	14	67.95	6.479	-1.419	.174
	The control group	13	62.12	13.438		
Average performance of female learners in both groups from the learning stages 2 to 4	the experimental group	10	73.5	6.609	-3.186	.006
	the control group	12	58.94	14.083		

After passing the learning stage 1, learners in the experimental group was detected 4 communities by using GA with the modularity Q function as the fitness function. Table 6 shows the results. One-way analysis of variance (ANOVA) was utilized in this study to analyze the learning performance among four communities. No significantly difference in learning performance appears among four groups (F=66.490, p=1.386>.05), indicating that information transfer through opinion leaders could let learners receive good information, without information gap.

Table 6. Descriptive statistics of four communities in the experimental group

Community	Number	Mean	SD	*F*	Significance (two-tailed)
Community 1	4	65.63	5.543	66.490	1.386
Community 2	8	69.84	7.148		
Community 3	6	70.42	8.862		
Community 4	5	75	4.146		

4.2 Analysis of social network interaction between both groups

In the social networks analysis, when a learner not interacting with other learners cannot be analyzed by using social networks analysis. Therefore, all learners in both groups not interacting with other learners were removed from the social networks analysis. The interaction of the experimental and control groups at the learning stages 1-4 were preceded the overall social networks interaction analysis. Table 7 shows the social networks analysis results between both groups with different information communication models.

Table 7. Overall social networks analysis results between both groups with different information communication models

Communication model	Overall network centrality		
	Degree centrality	Closeness centrality	Betweenness centrality
The experimental group (n=23)	0.5368	0.5756	0.3875
The control group (n=24)	0.4308	0.4533	0.1452

Degree centrality refers to the linkage degree to others; the larger degree centrality stands for the more active network interaction. The degree centrality of the experimental group, 0.5368, is higher than it of the control group, 0.4308, revealing more active information transfer among learners in the experimental group. Closeness centrality refers to the closeness among learners; the shorter distance among learners would show higher closeness centrality that the information could be faster acquired. The results show the higher closeness centrality, 0.5756, of the experimental group than it, 0.4533, of the control group, showing that the information transfer through opinion leaders could shorten the distance among learners. Betweenness centrality refers to the degree of an individual controlling resource communications in the network. When a learner is in the shortcut between other learners, the betweenness value is higher. The result reveals that the betweenness centrality, 0.3875, of the experimental group is far higher than it, 0.1452, of the control group. The transfer through opinion leaders could have better linkages among learners and faster information communication to enhance the information transfer efficiency of collaborative learning teams.

The interactions of both groups with different communication models were observed in this study. Figure 4 shows the social networks interactions among learners in the control group with one-step flow of communication model through teacher's website announcement. In the figure, the linked line reveals the interactions among learners, and the direction of arrows shows the respondents of learners. From Fig. 4, the control group is about a close group, but the linkage among nodes is chaotic. It shows that learners in the control group do not seriously discuss, and the situation of collaborative learning is not obvious. Figure 5 shows the social network interactions of learners in the experimental group with two-step flow of communication through opinion leaders. Apparently, there is the appearance of sub-groups, and the linkages among nodes are easier than those in the control group. Apparently, learners in the experimental group present good communication intention and collaborative learning in each community detected by the community detection scheme.

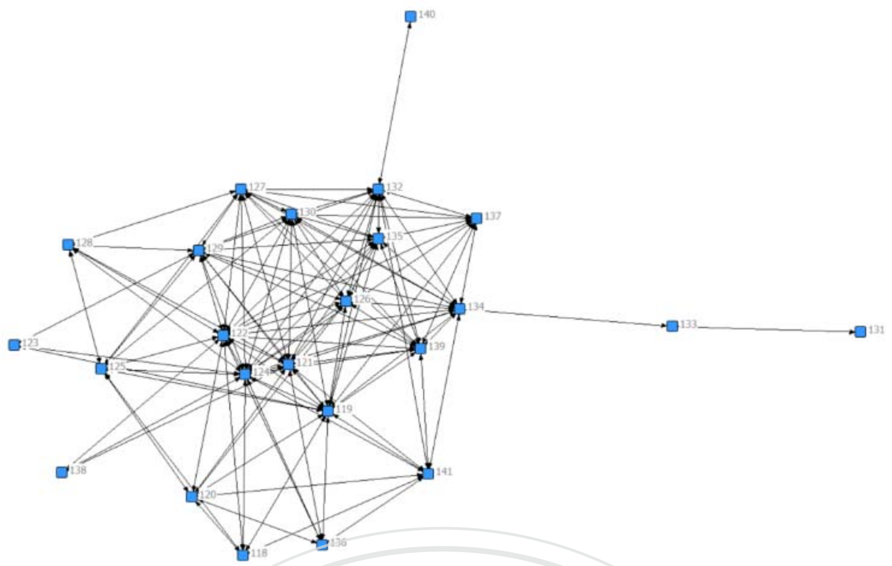


Figure 4. The control group with one-step flow of communication model through website announcement (n=24)

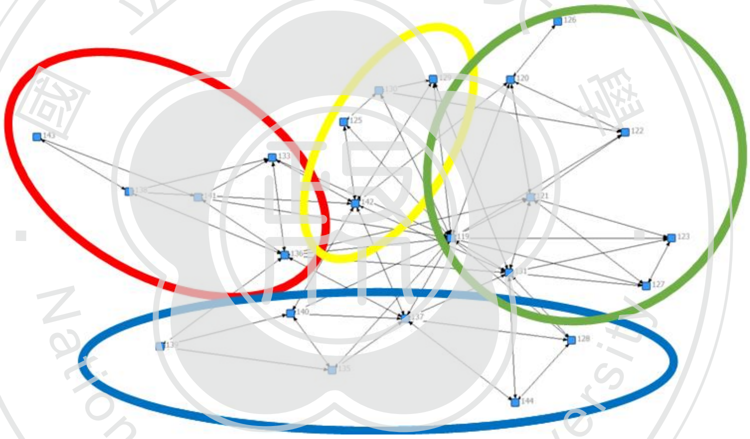
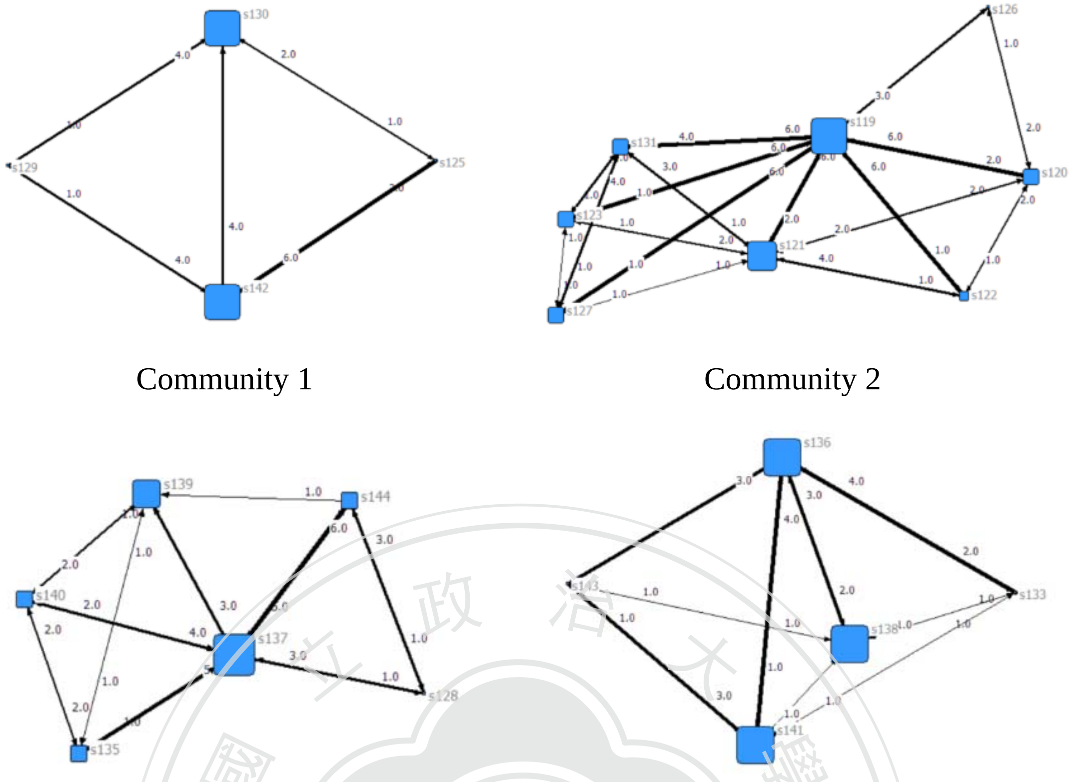


Figure 5. The experimental group with two-step flow of communication through opinion leaders (n=23)

Community interactions with the information transfer through opinion leaders were also observed in this study. The results show that, with the modularity Q function for community detection, four communities were detected in the experimental group as shown in Fig. 6. In the learning community network interaction map, the linked line represents the interactions between learners; the thicker line shows the more interactions between learners. The larger node pattern reveals the higher degree of learners in the community. According to PageRank measure, the opinion leaders of the groups are s142, s119, s137, and s136. In Fig. 6, an opinion leader appears the largest pattern and the thickest and the most node lines. It presents that opinion leaders are the ones often receive and transfer information in groups to help information transfer and enhance learners' learning performance.



Community 1 Community 2

Community 3 Community 4

Figure 6. Social network structure relationship of communities in the experimental group

4.3 Analysis of group cohesion between both groups

Table 8 shows the results of independent-samples *t* test of pretest and posttest of group cohesion for the experimental and control groups. The result shows no significant difference in group cohesion between two groups before performing the CPBL missions (t=-0.542, p=.590>.05); however, the group cohesion of both groups achieves remarkable difference after performing the CPBL missions (t=-2.005, p=.048<.05), and the experimental group is superior to the control group. Remarkably, compared to the pretest of group cohesion, the group cohesion of the control group slightly decreases. It was logically inferred the learners in the control group were not guided well by the teacher in the collaborative learning process so that they are confused and could not concentrate on the learning.

Table 8. Independent-samples *t* test of pretest and posttest of group cohesion for the experimental and control groups

Test item	Group	Number	Mean	SD	Mean-equivalent t test	
					t	Significance (two-tailed)
Pretest of group cohesion	The experimental group	24	50.92	9.930	-0.542	.590
	The control group	25	49.28	11.137		
Posttest of group cohesion	The experimental group	24	52.20	9.495	-2.205	.048
	The control group	25	46.00	11.982		

4.4 Summary of the interview results from the experimental group

To understand how opinion leaders conveyed the teaching guidance about solving the CPBL missions from the instructor to their group members and facilitated knowledge sharing and peers' interaction, six learners with especially excellent or poor learning performance and four opinion leaders were invited from the experimental group to participate in a semi-structured interview. Five learners agreed that they can get correct and rich information through the dissemination of information by their opinion leaders while one learner questioned the correctness of the information sent by her opinion leader. However, as a whole, most learners confirmed that the opinion leaders can help them improve the CPBL performance. Moreover, the six learners had inconsistent opinions on the effect of information dissemination through the instructor's website announcement for the learners in the control group. Some learners expressed that the information announced by the instructor through the website announcement can be delivered fairly because the website announcement is normally standardized by the instructor and the correctness will be relatively high. Furthermore, several learners thought that the information from the opinion leaders is not as authoritative as instructor. Interestingly, most learners agreed that not every person is interested in looking at the information from the instructor's website announcement. They also worried about that most learners may only use the information provided in the instructor's website announcement and will not have strong motivation to seek additional information by themselves, thus losing autonomous learning ability. Importantly, three opinion leaders thought that passing information through them can make learning more effective and increase positive interaction with their group members, while one thought that it may cause other group members not willing to actively find information by themselves because they just wait for the information from their opinion leaders.

5. Discussion

Li, Ma, Zhang, Huang, and Kinshuk (2013) indicated that opinion leaders can facilitate knowledge sharing on specific topics and enhance the efficiency of collaborative learning. According to the four-stage collaborative problem-based learning performance of learners in the experimental and control groups, the experimental group shows the higher learning performance than the control group, possibly because the learners in the experimental group are influenced by opinion leaders and perform the better knowledge share and information delivery. Moreover, this study found that female learners in the experimental group with two-step flow of communication through opinion leaders outperform female learners in the control group with one-step flow of communication through teacher's website announcement on learning performance, while male learners between both groups do not appear notable difference. The possible reason is that females are more easily and positively affected by opinion leaders than males. Several previous studies supported the viewpoint. For example, Elder and Greene (2003) found that women's vote choices are more likely than men's to be influenced by the organizations to which they belong. Norrander (1997) found that women feel significantly more positive about groups in society than men. Also, the effects of different information communication models on the interactions in CPBL communities are discussed in this study. Several studies (Carson, Tesluk, & Marrone, 2007; Cascio, & Shurygailo, 2002) indicated that opinion leaders play an important role in improving communication and encouraging group members to have greater level of information exchange. The results of the study reveal that the experimental group applying two-step flow of communication shows higher degree centrality, closeness centrality, and betweenness centrality than the control group using one-step flow of communication. It is inferred that the experimental group presents better information transfer efficiency and would absorb opinion leaders' opinions to facilitate better learning interactions. More importantly, the group cohesion of the learners in the experimental group was significantly higher than the learners in the control group after performing the CPBL missions. Obviously, the opinion leaders in the experimental group successfully and effectively directed their group members to finish the CPBL missions, thus promoting the group cohesion. Conversely, the group cohesion of the learners in the control group slightly decreased after performing the CPBL missions in comparison with before performing the CPBL missions. In other words, the impact of the opinion leaders on facilitating group cohesion was higher than did the instructor. Finally, the modularity Q function was utilized in this study for community detection, and PageRank measure was used to determine the opinion leaders. By inquiring the homeroom teacher of the experimental group, the communities and opinion leaders detected by the proposed algorithm are about the same as those in the real life. The results reveal that the high accuracy of using modularity Q function as the fitness function of GA to perform community detection and using PageRank measure to seek opinion leaders.

6. Conclusions and Future Works

The research findings of the study show that learners in the experimental group with two-step flow of communication through opinion leaders outperform those in the control group with one-step communication through website announcement on learning performance, peer interaction, and group cohesion. Moreover, under CPBL environment, the two-step flow of communication through opinion leaders provides more benefit in terms of promoting the learning performance of female learners than male learners. Moreover, by inquiring the homeroom teacher of the experiment group, the detected communities and opinion leaders are about the same as the real life conditions. It reveals that such a method on the CPBL system could effectively detect communities and opinion leaders.

Several issues warrant further study. First, learners in this study enhance the problem-solving capability with the CPBL system. Future research should consider apply the two-step flow of communication through opinion leaders to other collaborative learning systems for promoting learning performance. Moreover, elementary school Grade 4 students are the research objects in this study. Future research could expand the learners in different age groups of junior high schools, senior high schools, colleges, and graduate schools to further compare the learning performance and interactions so as to complete the research results. Finally, learners in both groups in this study are not tested the cognitive styles. The effect of two-step flow of communication, under CPBL system, on the learning performance of learners with distinct cognitive styles could be further discussed in the future.

Statements on open data, ethics and conflicts of interest

To consider the research ethics of the designed experiment, written informed consent was obtained from the research subjects following full explanation of the experiment. The informed consent letter contains about the specific nature of the research, including the data that collect from them is only for the research, their name will never appear on any data collected and that instead we will provide a unique identification number on their data and that this information will remain secure such that only the principal investigator of this study will have access to it, the collected data that are no longer needed will be destroyed, and how participation will make a contribution to our research's goals. Access to the database will be provided by the first author on the request of the interested party. Solicitations should contain information about the aim of the research and the type of analysis that researchers want to do. Finally, we certify that there is no conflict of interest in this paper.

Appendix

Group cohesion scale used in the study

Dimension	Item
Social cohesion	Our team members can always accept the opinions of others
	Our group members often express their support for the ideas from other members
	Our team members communicate well
	Our team members interact well
	It is always a pleasant time for our team members to discuss
	Our team members are very friendly to each other
	Our team members respond friendly to the questions from other members
Task cohesion	Our team members are very dedicated to the completion of the assignment
	Our team members are very dedicated to achieving the job goal
	Our team members show an interest in doing the job
	Our team members work together to complete the assignment
	Our team members will engage in each other's success in actual work or competition
	Our team members can work together and share work when they actually work

References

Barrows, H. S. (1996). Problem-based learning in medicine and beyond: A brief overview. *New Directions for Teaching and Learning*, *1996*(68), 3–12. http://doi.org/10.1002/tl.37219966804

Bass, B. M. (1990). *Bass & Stodgill's handbook of leadership*. New York: Free Press.

Choi, S. (2015). The two-step flow of communication in Twitter-based public forums. Social Science Computer Review, *33* (6), 696-711.

Carson, J. B., Tesluk, P. E., & Marrone, J. A. (2007). Shared leadership in teams: An investigation of antecedent conditions and performance. *Academy of Management Journal*, *50*(5), 1217-1234.

Cascio, W.F. & Shurygailo, S. (2002). E-Leadership and virtual teams. *Organizational Dynamics*, *31*(2), 362-376.

Chintalapudi, S. R., & Krishna Prasad, M. H. M. (2017). Community detection in large-scale social networks: A survey. *Graph Theoretic Approaches for Analyzing Large-Scale Social Networks*, 189-206.

Elder, L., & Greene, S. (2003). Political information, gender and the vote: The differential impact of organizations, personal discussion, and the media on the electoral decisions of women and men. *The Social Science Journal*, *40*, 385-399.

Childers, T. L., & Rao, A. R. (1992). The influence of familial and peer-based reference groups on consumer decisions. *Journal of Consumer Research, 19*(2), 198–211. http://doi.org/10.1086/209296

Gressick, J. & Derry, S. J. (2010). Distributed leadership in online groups. *Computer-Supported Collaborative Learning, 5*, 211-236.

Hung, P. H., Lin, I. H., Hwang, G. J., Lee, Y. H., Wu, T. H., & Chang, W. Y. (2012). The differences of collaborative learning and scientific inquiry competencies between experienced and novice learners in ubiquitous problem-based learning. *IEEE Seventh International Conference Wireless, Mobile and Ubiquitous Technology in Education (WMUTE),* pp. 277-281. http://doi.org/10.1109/WMUTE.2012.67

Hong, Y. H. (2016). The role of Facebook, Plurk, and YouTube in the two-step and N-step flows of communication and the effect on political efficacy. *Journal of International Communication, 22*(1), 42-63.

Katz, E., & Lazarsfeld, P. F. (1955). *Personal influence, the part played by people in the flow of mass communications*. New York: The Free Press.

Koschmann, T., Kelson, A. C., Feltovich, P. J. & Barrows, H. S. (1996). Computer-supported problem-based learning: A principled approach to the use of computers in collaborative learning. In T. Koschmann (Ed.) *CSCL: Theory and practice of an emerging paradigm*, Mahwah, NJ: Lawrence Earlbaum Associates.

Kobbe, L., Weinberger, A., Dillenbourg, P. et al. (2007). Specifying computer-supported collaboration scripts. *Computer-Supported Collaborative Learning*, *2*(2-3), 211-224.

Kollar, I., Fischer, F. & Hesse, F.W. (2006). Collaboration Scripts - A conceptual analysis. *Educational Psychology Review, 18*(2), 159-185.

Lazarsfeld, P. F., Berelson, B., & Gaudet, H. (1944). *The people's choice: How the voter makes up his mind in a presidential campaign*. New York: Columbia University Press.

Liu, Y. H., Lou, S. J., Shih, R. C., Meng, H. J., & Lee, C. P. (2010). A case study of online project-based learning: The beer king project. *International Journal of Technology in Teaching & Learning*, *6*(1), 43–57.

Li, Y., Ma, S., Zhang, Y., Huang, R., & Kinshuk. (2013). An improved mix framework for opinion leader identification in online learning communities, *Knowledge-Based Systems*, *43*, 43-51.

Ma, X., Wang, B., Yu, L. (2018). Semi-supervised spectral algorithms for community detection in complex networks based on equivalence of clustering methods. *Physica A: Statistical Mechanics and its Applications, 490*, 786-802.

Norrander, B. (1997). The independence gap and the gender gap. *Public Opinion Quarterly*, *61*(3), 464-476

Naidu, S., & Oliver, M. (1996). Computer-supported collaborative problem-based learning: An instructional design architecture for virtual learning in nursing education. *Journal of Distance Education*, *11*(2), 1–22.

Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E, 69*(2), 26113. http://doi.org/10.1103/PhysRevE.69.026113

Page, L., Brin, S., Motwani, R., & Winograd, T. (1998). The PageRank citation ranking: Bringing order to the Web, *Technical Report, Stanford Digital Libraries Technologies Project.*

Rogers, E. M. (1995). Diffusion of innovations. New York: Free Press.

Weinberger, A., Kollar, I., Dimitriadis, Y., Mäkitalo-Siegl, K., & Fischer, F. (2009). Computer-supported collaboration scripts. In: Balacheff N., Ludvigsen S., de Jong T., Lazonder A., Barnes S. (eds) Technology-Enhanced Learning. Springer, Dordrecht.

Seibold, G. L., & Kelly, D. R. (1988). Development of the combat platoon cohesion questionnaire. Alexandria, VA: US Army Research Institute for the Behavioral Sciences.

Zhang, Y. (2015). Study on word-of-mouth dissemination of cigarette brands based on two-step flow of communication theory. *Acta Tabacaria Sinica*, *21*(1), 112-118.

Zaccaro, S. J. (1991). Nonequivalent associations between forms of cohesiveness and group-related outcomes: Evidence for multidimensionality. *The Journal of Social Psychology*, *131*, 387-399.