전자게시판의 토의 활동에 대한 계산 모델
A Computational Model for Discussion Activities on Electronic BBS

문교식(Gyo Sik Moon)[1)]

요 약

전자게시판 기반의 컴퓨터 매게 통신 (CMC: Computer-Mediated Communication)은 인터넷의 출현 이래로 다양한 교육적 목적을 위하여 활용되어 왔다. 교수자는 게시물의 내용 분석을 통하여 교수목적을 성취할 수 있고 토의 활동에 대한 분석적 정보로부터 기술적 도움을 받을 수 있다. 내용 분석과 사회성에 대한 연구는 사회과학 분야에서 전통적으로 이루어져 왔다. 그러나 그 현상에 대한 수리논리적 분석에 대한 연구는 드물다. 이러한 목적을 위하여 본 논문에서는 커뮤니케이션 활동을 논리적으로 표현하고 분석하기 위하여 계산 모델을 제안한다. 모델에는 토의 활동에 대한 형식 표현을 제공하는 상호작용 벡터와 그것의 그래픽 표현인 상호작용 차트로 이루어져 있다. 이러한 모델은 토의 그룹을 계량적으로 분석하고 비교하는데 유용한 도움을 줄 수 있다.

Abstract

Computer-Mediated Communication (CMC) based on electronic bulletin board system (e-BBS) has been widely used for various educational purposes since the advent of the Internet. Instructors can achieve pedagogical goals by analyzing the contents of postings. They can be benefited technically from analytic information on discussion activities. Studies on content analysis and social presence in communication have been traditionally done by social studies. However, researches on computational and logical analysis of the phenomenon are rare. To do that, we propose a formal model to represent and analyze communication activities logically. The model consists of interactivity vector providing a formal representation of discussion activities and interactivity chart providing a graphical representation of interactivity vector. The formal approach can be used as a useful technical assistance to analyze and compare discussion groups computationally.

논문접수 : 2005. 5. 23.
심사완료 : 2005. 6. 20

1) 정회원 : 대구교육대학교 컴퓨터교육과

1. Introduction

Computer-Mediated Communication (CMC) has been widely recognized as an effective medium for collaborative learning (Eugenia, 2002; Levin & Waugh, 1997; Oliver et at., 1997; Yeung, 2002). Results on educational effectiveness of CMC using e-BBS and the improvements of the medium can be found in the literature (Hewitt, 2003; Nagai et al., 2002; Sato et al., 2001). Two areas of social science social network analysis, content analysis have contributed significantly to the study of CMC. Social network analysis that studies information entities and their relationships recently extends its specialty to CMC (Krebs, 2003). Content analysis investigates systematic ways to analyze messages for various purposes (Rourke et al., 1999). In this paper, we develop a formal representation of interactivity to describe formally discussion activities on e-BBS, which includes the interactivity vector and the interactivity chart.

2. Interactivity Vector

We use the notion of discussion tree, *T*, as usual in which nodes represent postings and edges express the connection of a posting and its immediate replies. Let *w*(*X*) be the writer of node *X*. A thread of a discussion tree is a sequence of nodes along the path from the root to its terminal node. We assume that, without loss of generality, (1) *T* consists of at least two nodes including the root, (2) a node *A* is allowed to be inserted as a child of *X* if *A* is a reply to *X* and *W*(*A*)*W*(*X*), (3) for any two nodes *A* and *B*, *W*(*A*)*W*(*B*) if they are siblings, and (4) the order of children nodes does not matter (Moon, 2004).

An interactivity vector with respect to a discussion tree, *T*, denoted by *I(T)*, can be formally represented by a 10-tuple (*P*, *S*, *W*, *D*, *Uni*, *Dom*, *Act*, *Att*, *Con*, *Dep*) where *P*: total number of participants (population), *S*: total number of postings (size), *W*: width, *D*: depth, *Uni*: uniformity, *Dom*: dominancy, *Act*: activeness, *Att*: attractiveness, *Con*: continuity, *Dep*: depth dependency. The vector can be a convenient computational tool to measure interactivity of a discussion tree.

(1) Definitions of *P*, *S*, *W*, and *D* *P* (population) and *S* (size) are defined above. *W* (width) is the maximum number of nodes at same level of a tree. *D* (depth) is the number of nodes of the longest thread. Figure 1 shows an example of a discussion tree where *m*1, *m*2, *m*3, are nodes and *a*, *b*, *c*, are participants. So, we have *P*=4, *S*=10, *W*=5 (*m*5, *m*6, *m*7, *m*8, *m*9) and *D*=4 (*m*1, *m*3, *m*8, *m*10).

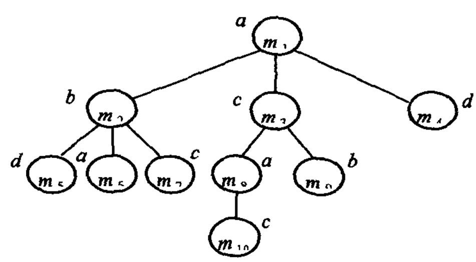


Figure 1: A discussion tree with 10 nodes

(2) Definition of *Uni* *Uniformity* is to measure the distribution of postings among participants, which can be done by simply computing the reciprocal of one plus the standard deviation. The value lies in between 0 and 1.

(3) Definition of *Dom* A small fraction of participants can be dominant in postings. Let *d* be the least number of participants whose postings take the majority. Then, *Dom* = 1 *d*/*P*. The larger the value of *Dom*is, the fewer participants take the majority of

postings. Example: Let *TA* and *TB* be two trees. Both have same number of *P* and *S* P=5, S=15. The participants and their postings are as follows; *TA*: (*a*, *b*, *c*, *d*, *e*) = (3,3,4,2,3), *TB*: (*v*, *w*, *x*, *y*, *z*) = (3,1,2,7,2). Then, three participants (*a*, *b*, *c*) of *TA* take the majority of postings while two (*v*, *y*) of *TB* take the majority. We have $Dom(TA)$=1-3/5=0.4, $Dom(TB)$=1-2/5=0.6. So, $Dom(TB) > Dom(TA)$.

(4) Definition of *Act* Let's define *active participants*as those participants that post no less than the average number of postings. The *activeness* of a tree is given as the number of active participants divided by *P*. Example: Let *TA*, *TB* be the two trees as defined above. Four (*a*, *b*, *c*, *e*) out of five are active in *TA*, but only two (*v*, *y*) are active in *TB*. So, we have $Act(TA)$ = 4/5 = 0.8, $Act(TB)$ = 2/5 = 0.4. The result shows that participants of *TA* are more active in general than those of *TB*.

(5) Definition of *Att* Some postings are more likely to motivate others to respond, which can contribute to the interactivity of discussion. *Attractiveness* for a nonterminal node *X*, $Att[X]$, is the number of immediate replies to *X*. *Att* of a tree is given as $XAtt[X]$ / *N*, where *N* is the number of nonterminals.

Example: Figure 2 has $Att[m1]$ = 1, $Att[m2]$ = 2, $Att[m3]$ = 1, $Att[m5]$ = 1, and Att = (1+2+1+1) / 4 = 1.25, and Figure 3 has $Att[n1]$ = 3, $Att[n3]$ = 2, and Att = (3+2) / 2 = 2.5, which indicates that postings of the latter are more likely to attract participants to respond than those of the former.

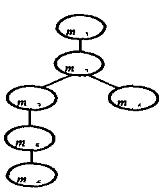


Figure 2: *Attractiveness* (*Att*=1.25)

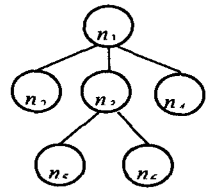


Figure 3: *Attractiveness* (*Att*=2.5)

(6) Definition of *Con* Let's define a thread, $h(n)$, be the sequence of nodes from the root to the node *n*, where *n* is the terminal node in the thread. Let the continuity of a node be $Ch(n)[X]$ = 1 if there exists a node *A* such that *A* is a proper ancestor of *X* and $w(A)$=$w(X)$ along the thread $h(n)$ and 0 otherwise. Then thecontinuity of a thread is the sum of continuities of all nodes in the thread. I.e., $C(h(n))$ =$jCh(n)[Xj]$. Furthermore, the continuity of a tree is the sum of continuities of all threads divided by the number of threads in the tree. Therefore, *Con* = $nC(h(n))$ / $n|h(n)|$. Example: Both (Fig. 4) and (Fig. 5) have equal number of *S* and *P* with same structure except the writers of nodes 4, 5, and 9. (Fig. 4): *Con* = ($C(h(m4))$ + $C(h(m8))$ + $C(h(m9))$ + $C(h(m6))$ + $C(h(m10))$) / ($|h(m4)|$ + $|h(m8)|$ + $|h(m9)|$ + $|h(m6)|$ + $|h(m10)|$) = (1+1+1+0+2) / (3+4+4+3+4) = 5/18 0.28. (Fig.5): *Con* = ($C(h(m4))$ + $C(h(m8))$ + $C(h(m9))$ + $C(h(m6))$ + $C(h(m10))$) / ($|h(m4)|$ + $|h(m8)|$ + $|h(m9)|$ + $|h(m6)|$ + $|h(m10)|$) = (0+1+0+0+2) / (3+4+4+3+4) = 3/18 0.16. The continuity of (Fig. 4) is higher than that of the other.

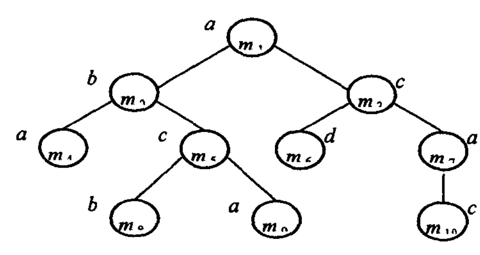


Figure 4: *Continuity* (*Con*=0.28)

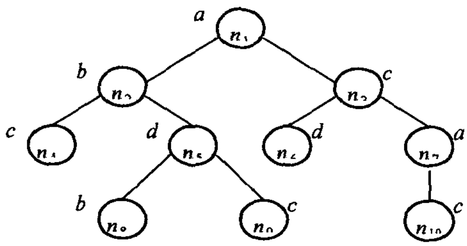
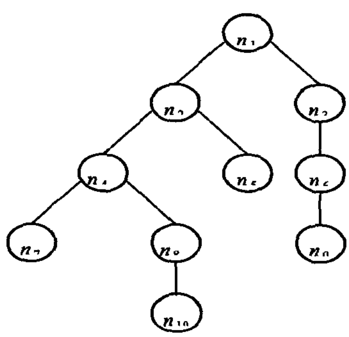


Figure 5: *Continuity* (*Con*=0.16)

(7) Definition of *Dep* We define *max threadas* a thread containing largest number of nodes, which is deepest in a tree. *Depth dependency*, *Dep*, is (the sum of the number of nodes of a max thread + the number of their immediate replies) / *S*. Example (Figure 6): In this figure, max thread is (*n*1, *n*2, *n*4, *n*8, *n*10) and their children nodes excluding those already counted in the max thread are *n*3, *n*5, and *n*7. So, *Dep*= 8/10, which shows fairly high dependency on the max thread.



gure 6: *Depth dependency*(*Dep*=0.8)

Example of interactivity vector: Let's take two trees, *TA* and *TB*, as examples for showing interactivity vectors (Figure 7, Figure 8). Both have same values of *P* and *S*. (Table 1.) shows the result of the computation and the comparison of the two trees. (Table 2.) provides a succinct interpretation to help achieve better understanding of discussion activities includingstructural analysis as well as participants' behavioral characteristics.

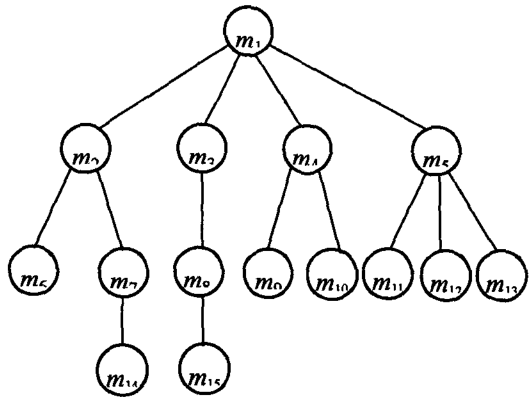


Figure 7: *Interactivity Vector* for *TA*

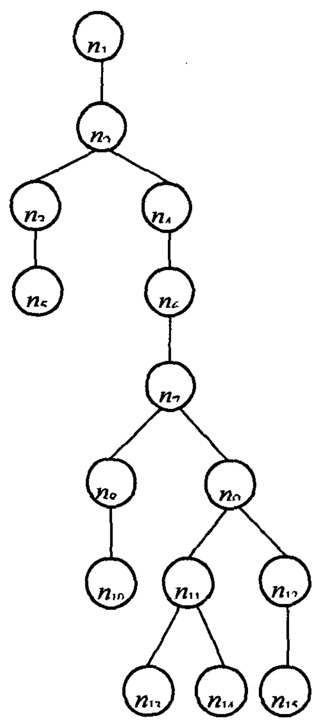


Figure 8: *Interactivity Vector* for *TB*

Table 1: Computation of two interactivity vectors

	P	*S*	*W*	*D*	*Uni*	*Dom*	*Act*	*Att*	*Con*	*Dep*
TA	5	15	8	4	0.58	0.4	0.8	2.0	0.12	0.47
TB	5	15	3	9	0.30	0.6	0.4	1.4	0.37	0.87

Table 2: Comparison of discussion trees

Elements	*TA*	*TB*
Width	wide	narrow
Depth	shallow	deep
Depth dependency	low	high
Continuity	low	high
Dominancy	low	high
Uniformity	high	low
Attractiveness	high	low
Activeness	high	low

3. Interactivity Chart

We propose an interactivity chart to represent the interactivity vector in a graphical form. The chart depicts all elements of the vector through a systematic arrangement of the elements taking interrelationships between them into account for clarification of discussion activities. There are four basic elements in the vector size, population, width, and depth which determine the size and shape of atree. Depth and depth dependency seem to have a strong linear relationship with each other. Continuity clearly has a positive effect on both depth and depth dependency. And it is evident that there should be a very good linear relationship between attractiveness and width. Continuity is most likely to influence positively on dominancy. It is obvious that the notion of dominancy is opposite to that of uniformity. Higher activeness tends to cause higher uniformity, and vice versa. Example: (Figures 9 and 10) depict the charts for *TA* and *TB* from (Table. 1). The value of each line is relative to the entire discussion forum consisting of multiple of discussion trees.

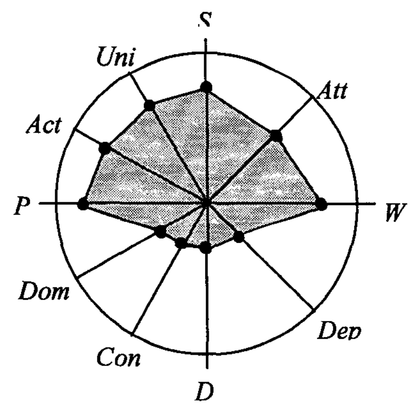


Figure 9: *Interactivity Chart* for *I*(*TA*)

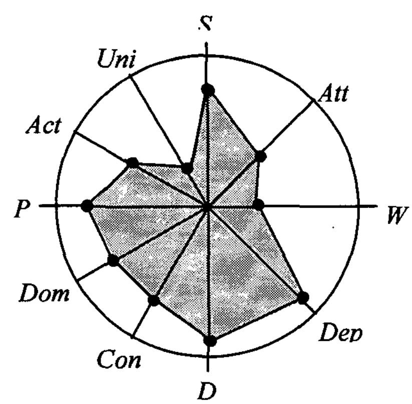


Figure 10: *Interactivity Chart* for *I*(*TB*)

4. Conclusion

We developed a 10-tuple interactivity vector for a formal representation of asynchronous discussion activities. And then we developed an interactivity chart to represent the vector graphically. The main objective of the formal approach is to provide a convenient, succinct representation of complex phenomenon of discussion activities. Therepresentation can be used as a tool for analyzing interactivity as well as participants' behavior in asynchronous communication.

References

[1] Eugenia, M. W. Ng. (2002), Enhancing Collaborative Learning through Online Discussion and Peer Assessment, ICCE2002, 262-263.

[2] Hewitt, J. (2003), How habitual online practices affect the development of asynchronous discussion threads, Journal of Educational Computing Research, 28(1), 31-45.

[3] Krebs, V. (2003). Introduction to Social Network Analysis, http://www.orgnet.com/sna.html.

[4] Levin, J., & Waugh, M. (1997), Teaching Teleapprenticeships: Electronic network-based educational frameworks for improving teacher education, Interactive Learning Environments Journal, 6(1-2), 39-58.

[5] Moon, G. S. (2004), Computation of interactivity in Asynchronous Text-Based Communication, ICCE2004, 1693-1696.

[6] Nagai, M., Shiraki, K., Kato, H., & Akahori, K. (2002), The Effectiveness of a Web Bulletin Board Enhanced with a Knowledge Map, ICCE2002, 268-272.

[7] Oliver, R., Omari, A., & Herrington, J. (1997), Exploring student interactionsin collaborative World Wide Web learning environments, T. Muldner & T. Reeves (Eds.) Educational Multimedia/Hypermedia and Telecommunications, Charlottesville, VA: AACE, 812-817.

[8] Rourke, L., Anderson, T., Garrison, D. R., & Archer, W. (1999). Assessing social presence in asynchronous, text-based computer conferencing, Journal of Distance Education, 14(3) 51-70.

[9] Sato, K., Yokoyama, S., & Akahori, K. (2001), Development and Evaluation of an Electronic Bulletin Board System Visualizing User Interaction, ICCE2001, 114-121.

[10] Yeung, A. S. W., & Ferry, B. (2002), Developing a Peer Learning Community through the use of CMC, ICCE2002, 371-375.