論文
3-7

BBS의 토의활동 측정
Measuring Discussion Activities in BBS

문교식(Gyo Sik Moon)[1]

Abstract

Electronic BBS (bulletin board system) has been widely recognized as an appropriate medium for exchanging ideas and sharing information asynchronously. The communication ability of BBS is the main reason for utilizing it as a tool for collaborative learning. Researchers in the community reported a number of findings regarding the educational utilization of BBS recently. In this paper, we propose a qualitative method to measure communication activities using BBS so that the complex discussion behaviors of participants can be understood analytically. We propose characteristic vectors to describe discussion behaviors of groups and individuals, which can be conveniently used for characterizing and comparing discussion groups as well as individuals. The interactivity model representing interactive activities shows graphically the degree of interactivity of discussion groups as well as individuals. Also, time dependent measurements are investigated to analyze discussion activities with time. Experiments on the proposed measurements conducted on the Web-based discussion project using BBS demonstrate how measurements can be carried out, how characteristic vectors and interactivity model can be constructed and used.

요약

전자게시판은 비동기적으로 생각을 교환하고 정보를 공유하는 적절한 매체로써 널리 인식되어왔다. 게시판의 의사소통 능력은 협력학습을 위한 도구로 사용하는 주된 이유이다. 최근 이 분야의 연구자들은 게시판의 교육적 활용에 대한 연구결과를 다수 보고하고 있다. 이 논문에서는 게시판에서의 의사소통 활동을 측정함으로써 토의 참여자들의 복잡한 토의 행동을 정량적으로 이해할 수 있는 분석적 방법을 제안한다. 토의 집단이나 개인들의 토의 행동을 기술하기 위한 특성 벡터를 제안한다. 이 특성 벡터는 토의 집단과 개인에 대한 특성화와 비교를 편리하게 할 수 있도록 한다. 상호작용 활동을 표현하는 상호작용 모델은 토의 그룹과 개인에 대한 상호작용 정도를 시각적으로 나타낸 준다. 또한, 시간에 따른 토의 활동을 분석하기 위해 시간 종속적인 측정을 탐구한다. 전자 게시판을 사용한 웹 기반 토의 프로젝트를 통하여 제안한 측정이 어떻게 이루어지는지, 또 특성 벡터와 상호작용 모델이 어떻게 구축되고 사용되는지를 실험을 통하여 보여준다.

Keywords:BBS, discussion behavior, interactivity model, CMC, asynchronous communication

논문접수 : 2004. 3. 10.
심사완료 : 2004. 3. 25.

1) 정회원 : 대구교육대학교전산교육과 교수

1. Introduction

BBS has been recognized as a powerful tool for asynchronous communication, thereby providing a fitting environment for collaborative learning [1],[2],[3],[4],[5]. Numerous efforts to make improvements to traditional BBS have been made to achieve better performance in terms of operability, user-friendliness, efficiency, and interactivity [6],[7]. For the analysis of information entities and their relationships social network analysis has been widely investigated in the community [8]. Content analysis of social presence in asynchronous text-based computer conference can be found in [9]. In this paper, we focus on measuring discussion activities in BBS and propose a formal method to diagnose communication activities in BBS so that complex discussion behaviors of participants can be understood formally.

Objectives of the paper can be briefly stated as the following;

(1) to identify statistics to be measured for the comprehension of communication behavior and for the formal analysis,

(2) to develop characteristic vectors describing discussion activities of groups as well as individuals formally,

(3) to develop interactivity models to represent interactive activities in a graphical form and to show clearly the degree of interactivity of discussion groups as well as individuals,

(4) to take time dependent measurements to analyze discussion activities with time, which reveals dynamic nature of the statistics, and lastly,

(5) to take actual measurements on a real discussion project conducted at a teacher's college in Korea and discuss the results along with difficulties and problems.

2. Measuring Discussion Activities

The measurement of discussion activities provides valuable statistical information on discussion activities to diagnose communication activities, to produce data for grading participants' performance, and to adjust the system for improvement. The discussion activities can be classified into two categories - group activities and individual activities.

The data for discussion activities are gathered throughout the project from measurable instances caused by participants' requests for posting, replying, viewing, and evaluating as well as the operations mechanically performed by the system to manage discussion groups and to ensure favorable environment for high interactivity among participants.

We propose characteristic vectors to describe the discussion behaviors of groups as well as individuals. The vectors are succinct and efficient representations of complex discussion behaviors showing characteristic features and they are conveniently used for characterizing and comparing discussion groups as well as individuals.

2.1 Measuring Group Activities

Data for measuring group activities can be gathered throughout a discussion project. A discussion group consists of an original writing and its replies and their replies too and so on. The original writing is called the root of the group and a group can be considered as a multi-level tree in which the original writing corresponds to the root and replies to the root correspond to children nodes. A node can be either a posting or a reply.

A characteristic vector with respect to a group G, denoted by V_G, can be formally represented by a 7-tuple (N, S, L, V, A, T_E, T_M), where

(1) N: total number of participants (*population*),
(2) S: total number of nodes (*size*),
(3) L: average length of writings (in terms of the number of characters),
(4) V: total number of views,
(5) A: average value of agreement,
(6) T_E: elapsed time in minutes,
(7) T_M: meantime between replies (*reply time gap*).

N (*population*) is the number of participants participated in a discussion group G, while S (*size*) is the number of nodes in G created by the participants. $S \geq N$ because each participant of G creates at least one node in G. Meantime between replies is the average time between when messages are posted and when they get replied. Bigger meantime means wider gaps of time between messages, suggesting that communication is inactive. The vector can be constructed directly from observing group activities. Furthermore, we can derive a set of new statistics out of the vector for identifying critical variables to construct a group interactivity model. The derived statistics include;

(1) *view ratio*: $v = V / S$,
(2) *maximum depth*: δ_M,
(3) *average depth*: δ_A,
(4) *average participation*: $\pi = S / N$,
(5) *frequency*: $\phi = S / T_E$.

View ratio is the average number of messages read per writing. Reading sufficient number of messages is essential to grasp the flow of discussion and to have a balanced opinion. So, the higher the ratio is, the better chance to have a high quality opinion.

Both maximum depth and average depth refer to how deep a discussion tree is. Maximum depth determines the depth of a discussion tree as a whole. Average depth can be easily computed by dividing the sum of depth of each leaf node by the total number of leaves. They indicate the depth of interaction in a discussion group. The depth of interaction can be labeled as "deep" or "shallow" depending on the values of δ_M and δ_A. As the number of leaf nodes with high values of δ_M and δ_A increases, the group is likely to lead to deep interaction.

The average participation tells us the average number of writings per participant. The frequency is the average number of writings per unit time.

2.1.1 Group Interactivity Model

We propose a group interactivity model to represent interactive activities in a graphical form. The purpose of the model is to show succinctly the degree of interactivity of a discussion group. Using the model, we can easily understand the behavior of a discussion group as well as compare behaviors of different groups. Some discussion groups are "strong" while others are "weak" in terms of the degree of interactivity.

We define the model which consists of seven dimensions of discussion activities as the following; N(*population*), S(*size*), v(*view ratio*), δ_M(*maximum depth*), δ_A(*average depth*), T_M(*reply time gap*), and ϕ(*frequency*). The model consists of horizontal lines each of which corresponds to a dimension mentioned as above. The interactivity increases as the value of a dimension goes to the left and decreases as the value goes to the right. Figure 1 shows an example of using the model illustrating graphically the interactivity of three discussion groups. Both group 1 and group 2 have nearly same values of population and size. However, the two groups show noticeable differences in other dimensions. Group 3 has low population and size, but it shows fairly high values in maximum depth, average depth, and reply time gap.

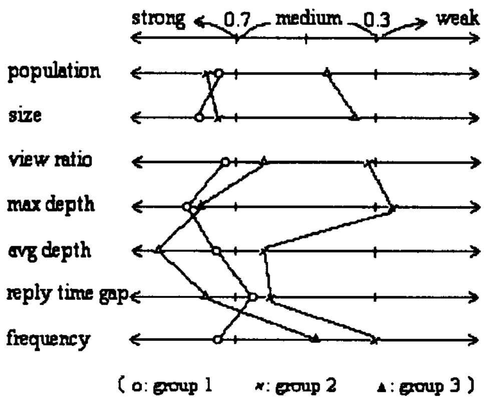

Figure 1. Example of a Group Interactivity Model

The actual computation of interactivity involves five dimensions - view ratio, maximum depth, average depth, reply time gap, and frequency. Both population and size are excluded in the actual computation of interactivity because they are not directly related to interactivity, but the two remain present in the figure due to their significance in understanding the nature of a discussion group. Each of these values has a scale of 0 to 1. The value of each dimension is relative to the entire discussion project consisting of many groups. The scale of the reply time gap is normalized to the range from 0 to 1 such that shorter gaps are located near 1.

A dimension is called "strong" in terms of group interactivity if its value is no less than 0.7 and "weak" if the value is no greater than 0.3 and "medium" if it lies in between. For example, Figure 2 shows that all dimensions of group 1 except the reply time gap are strong while group 2 finds no strong-strength interactivity. Three dimensions of group 3 are strong.

We may introduce a single indicator called "the degree of interactivity" out of the five dimensions of interactivity for a group by computing an arithmetic mean of all values of the five dimensions. The single indicator can be a convenient tool for characterizing a discussion group in terms of interactivity. Table 1 shows the computation of the degree for each group illustrated in Figure 1, revealing that both group 1 and group 3 are strong, but group 2 is medium-strength interactivity.

Dimension	Group 1	Group 2	Group 3
population	0.75	0.77	0.44
size	0.81	0.75	0.36
view ratio	0.73	0.32	0.62
max depth	0.83	0.25	0.80
avg depth	0.74	0.30	0.47
reply time	0.76	0.60	0.92
frequency	0.65	0.60	0.78
degree	0.74	0.42	0.72
results	strong	medium	strong

Table 1. Computation of the Degree of Interactivity

2.1.2 Time Dependent Measurements

We have discussed static measurements to analyze group activities. However, dynamic analysis of discussion activities may give us better understanding of group behaviors. For instance, measurements of population and size can only provide static information on the number of participants and that of messages throughout the whole session. Sometimes we might be interested in seeing the pattern of population growth with time. Time dependent measurements can be done daily/hourly depending on the type of analysis sought. Daily/hourly measurements may reveal the participants' tendency towards particular activities. Table 2 shows appropriate types of measurements for various statistics.

Daily Measurements	population, size, depth, frequency
Daily & Hourly Measurements	length, views, replies, postings, reply time gap, view ratio, reply ratio

Table 2. Daily and Hourly Measurements

Furthermore, we can measure the degree of interactivity on a daily basis to see how interactivity of a group changes with time, which might give us valuable insight for the behavioral analysis of a discussion group.

2.2 Measuring Individual Activities

Analysis of individual activities is just as important as that of group activities. Individual behaviors may vary widely from person to person due to personal characteristics. Yet we can design an objective measuring model that clearly shows an individual's behavior.

We define a characteristic vector for an individual as well as individual interactivity model and then we define the degree of interactivity as we did in group activities. We investigate why some individuals have high degree of interactivity while others do not. We also seek to find out which factors affect most the degree of interactivity for an individual.

A characteristic vector with respect to an individual I, denoted by V_I, can be formally represented by a 10-tuple (S, L, V, R, R', R'', P, E, T_E, T_M), where

(1) S: total number of nodes (*size*),
(2) L: average length of writings (in terms of the number of characters),
(3) V: total number of views,
(4) R: total number of replies,
(5) R': total number of primary replies,
(6) R'': total number of secondary replies,
(7) P: total number of postings,
(8) E: total number of evaluations,
(9) T_E: elapsed time in minutes,
(10) T_M: meantime between writings (*writing time gap*).

The construction of the vector can be done similarly as done in a group vector. Also, we can derive a set of new statistics out of the vector for identifying critical variables to construct an individual interactivity model. The derived statistics include;

(1) *view ratio*: $\upsilon = V / (R + P)$,
(2) *reply ratio*: $\rho = R / (R + P)$,
(3) *attraction*: $\alpha = R' / (R + P)$,
(4) *responsiveness*: $\omega = R'' / R'$,
(5) *frequency*: $\phi = S / T_E$.

View ratio, reply ratio, and frequency have the same meaning as in group statistics except that all statistics are concerned about individual activities rather than group ones.

Attraction is the measurement of how significantly the individual's writings contribute to the interactivity of a group discussion. Poor writings are hardly likely to attract others to respond. Messages with brilliant ideas that provoke new thoughts and opinions are very likely to motivate others to respond. Writings invoking many replies are a major source of a great contribution to the interactivity of discussion. The statistic "responsiveness" concerns with how frequent the individual replies to those messages replied to the messages written by the individual.

For example, Figure 2 shows that each

node in the graph represents a writing - either a posting or a reply. Squares represent nodes written by a particular individual and circles represent others'. Each node, n, is identified by a unique number and referenced by its unique number such as n_1, n_2, ..., and so on.

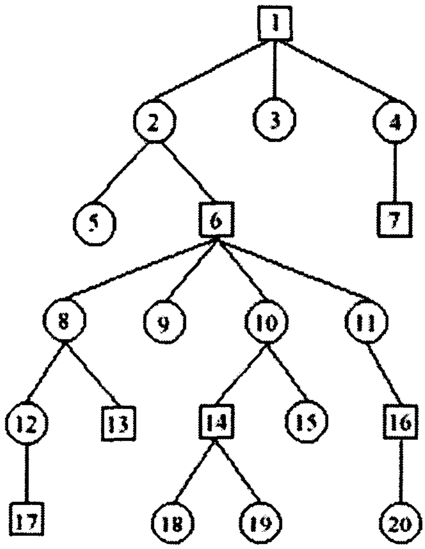

Figure 2. Example of a discussion group

Reply ratio: $\rho = |\{n_6,n_7,n_{13},n_{14},n_{16},n_{17}\}| \ / \ |\{n_1,n_6,n_7,n_{13},n_{14},n_{16},n_{17}\}| \fallingdotseq 0.86$

Attraction: $a = (|\{n_2,n_3,n_4\}| + |\{n_8,n_9,n_{10},n_{11}\}| + |\{n_{18},n_{19}\}| + |\{n_{20}\}|) \ / \ |\{ n_1, n_6,n_7,n_{13},n_{14},n_{16},n_{17}\}| \fallingdotseq 1.43$

Responsiveness: $\omega = (|\{n_6,n_7\}| + |\{n_{13},n_{14},n_{16}\}|) \ / \ (|\{n_2,n_3,n_4\}| + |\{n_8,n_9,n_{10},n_{11}\}| + |\{n_{18},n_{19}\}| + |\{n_{20}\}|) = 0.5.$

Figure 3 shows the case of a less attraction than in Figure 3, obtained from removing n_3, n_9, n_{18}, n_{19}, n_{20} from Figure 2. The attraction $a = (|\{n_2,n_4\}| + |\{n_8,n_{10},n_{11}\}|) \ / \ |\{n_1,n_6,n_7,n_{13},n_{14},n_{16},n_{17}\}| \fallingdotseq 0.71$, which is nearly half of the original.

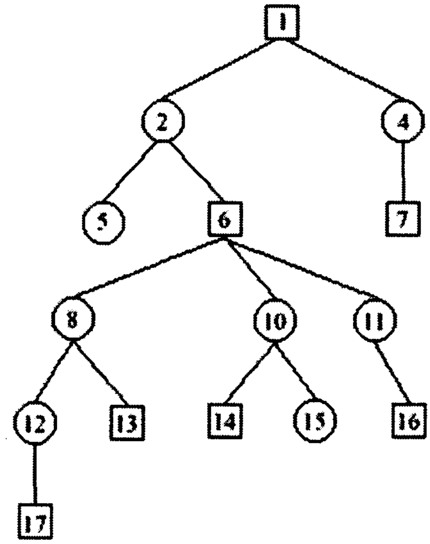

Figure 3. Example of a less attraction than in Figure 2

Figure 4 shows the case of a less responsiveness than in Figure 3, obtained from removing n_7, n_{13} from Figure 2. The responsiveness $\omega = (|\{n_6\}| + |\{ n_{14},n_{16}\}|) \ / \ (|\{n_2,n_3,n_4\}| + |\{n_8,n_9,n_{10},n_{11}\}| + |\{n_{18},n_{19}\}| + |\{n_{20}\}|) = 0.3.$

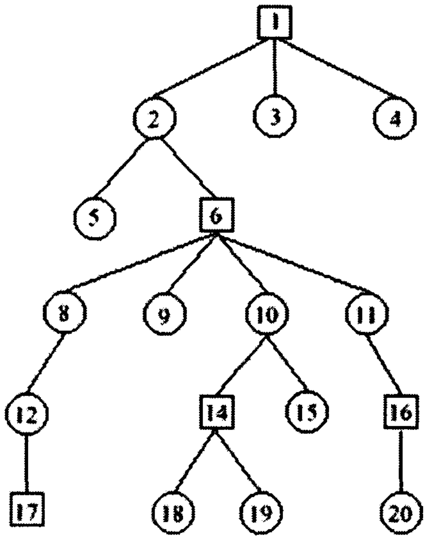

Figure 4. Example of a less

responsiveness than in Figure 2

2.2.1 Individual Interactivity Model

As we discussed previously the group interactivity model, the interactivity model with respect to an individual can be developed in a similar fashion.

We define the model which includes seven dimensions of discussion activities as the following; *S(size)*, *υ(view ratio)*, *ρ(reply ratio)*, *α(attraction)*, *ω(responsiveness)*, T_M *(writing time gap)*, and *ϕ(frequency)*. The shape and meaning of the model is similar to those of the group interactivity model. Figure 5 shows an example of the model in which two cases are illustrated. The two cases show fairly different shapes each other. One shows highly interactive activities with high degree of reply ratio, attraction, and responsiveness while the other shows very low values.

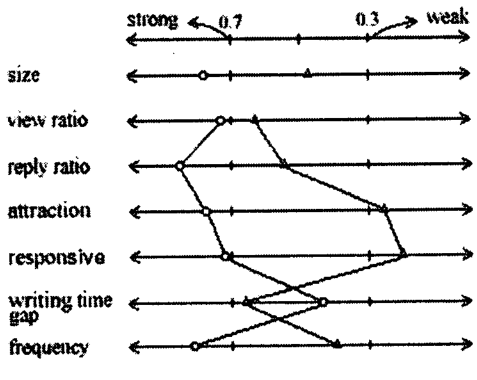

Figure 5. Example of an Individual Interactive Model

The actual computation of interactivity involves six dimensions - *view ratio, reply ratio, attraction, responsiveness, writing time gap*, and *frequency*.

The meaning of "strong" and "weak" group interactivity is similarly used here. The notion of "degree of interactivity" with respect to a group still holds for individual analysis.

2.2.2 Time Dependent Measurements

For the same reason as we need time dependent measurements to analyze group activities, individual activities need to be analyzed for the sake of understanding participants' interactive activities with time. Table 3 shows the appropriate types of measurements for various statistics.

Daily Measurements	size, frequency, attraction, responsiveness
Daily & Hourly Measurements	length, views, replies, postings, writing time gap, view ratio, reply ratio

Table 3. Daily and Hourly Measurements

Finally, we can measure the degree of interactivity on a daily basis to see how interactivity of an individual changes with time.

3. Experimental Results

A Web-based discussion project using BBS was carried out for seventy students studying at a teacher's college in Korea taking Information Society and Computer II in September 2002. The discussion topic was about computer ethics and the participants were asked to post messages for exchanging ideas on the topic. Statistics were gathered and analyzed by a set of ASP (Active Server Pages) programs working with SQL database running ona Windows NT server. The characteristic vectors and the interactivity model were obtained as the result of the project. Some of the results for a few discussion groups and individuals are presented here to demonstrate the efficacy of the formal measurements to understand the behavior of discussion activities.

3.1 Results on Group Activities

The characteristic vectors of discussion groups are shown in Table 4 and the computation of interactivity for the groups is shown in Table 5. The group G_1 has the largest population(N) and size(S). Even though the view ratio(υ) of G_1 is lower than that of G_2, maximum depth(δ_M) and average depth(δ_A) are much higher than those of G_2, indicating that interaction of G_1 is higher than that of G_2. Groups other than G_1 show "shallow" interaction and lower participation. The interactivity model can be obtained simply by normalizing the measured statistics to the range from 0 to 1 relative to the entire discussion forum.

Dim	G_1	G_2	G_3	G_4
N	21	10	10	9
S	38	12	11	9
L	436	562	430	534
V	614	237	169	152
R	37	11	10	8
P	1	1	1	1
A	3.24	2.71	3.52	3.83
$T_E(m)$	2939	2449	928	427
$T_M(m)$	230	369	173	131

Table 4: Characteristic Vectors

Dim	G_1	G_2	G_3	G_4
N	21	10	10	9
S	38	12	11	9
υ	16.16	19.75	15.36	16.89
δ_M	11	3	4	2
δ_A	5.65	2.38	3	1.5
$T_M(m)$	230	369	173	131
Φ	13	4	11	9

Table 5: Computation of Interactivity

3.2 Results on Individual Activities and Future Research

Data analysis for individuals is shown in Table 6 and Table 7, illustrating characteristic vectors and the computation for interactivity model. Individuals I_3 and I_4 posted a lot more messages than the others. Also, I_3 and I_4 have higher values of V and R. I_3 has the highest value of attraction(=0.77) contributing to the interactivity of the discussion. However, the responsiveness is very low for all cases, resulting in negative effects on interactivity.

Dim	I1	I2	I3	I4
S	13	12	30	34
L	315	539	539	568
V	180	55	406	249
R	12	7	20	21
R'	4	12	23	11
R''	0	2	3	3
P	1	5	10	13
A	3.16	2.95	3.65	3.58
$T_E(m)$	8149	7066	11499	9814
$T_M(m)$	626	588	383	288

Table 6: Characteristic Vectors

Dim	I1	I2	I3	I4
S	13	12	30	34
υ	13.85	4.58	13.53	7.32
ρ	0.92	0.58	0.58	0.62
a	0.31	1	0.77	0.32
ω	0	0.17	0.13	0.27
$T_M(m)$	626	588	383	288
ϕ	1.86	2.4	3.33	4.25

Table 7: Computation of Interactivity

4. Conclusion

To formally measure discussion activities in BBS, we first identified statistics for constructing characteristic vectors (a 7-tuple for group activities and a 10-tuple for individual ones) from which we computed derived statistics. We then developed interactivity models represented by seven dimensions showing the degree of interactivity graphically as well as computationally. The interactivity model determines one of three results in terms of the degree of interactivity - strong, medium, and weak. Finally, we conducted time dependent measurements to understand the dynamic behavior of discussion activities. Experiments on a discussion project show that characteristic vectors and interactivity models were effective in analyzing discussion behaviors.

In this paper, quantitative measurements were discussed to analyze discussion activities formally. But, we have not shown the effectiveness of these measurements for collaborative learning, which may be investigated for further research.

References

[1] 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.

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

[3] Nakanishi, Y., Tatsuta, L., Ohnishi, M., Iguchi, A., Tomizawa G., & Gunji, K. (2002), Collaborative Knowledge Building through BBS for Better Cross-Cultural Communication, ICCE-2002, 404-408.

[4] Eugenia. M. W. Ng(2002), Enhancing

Collaborative Learning through Online Discussion and Peer Assessment, ICCE-2002, 262-263.
[5] Yeung, A. S. W., & Ferry, B. (2002), Developing a Peer Learning Community through the use of CMC, ICCE-2002, 371-375.
[6] Sato, K., Yokoyama, S., & Akahori, K. (2001), Development and Evaluation of an Electronic Bulletin Board System Visualizing User Interaction, ICCE-2001, 114-121.
[7] Nagai, M., Shiraki, K., Kato, H., & Akahori, K. (2002), The Effectiveness of a Web Bulletin Board Enhanced with a Knowledge Map, ICCE-2002, 268-272.
[8] Krebs, V. (2003). Introduction to Social Network Analysis, http://www.orgnet.com/sna.html.
[9] 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.

문교식

1982 경북대학교 컴퓨터공학과 공학사
1982-1986 한국과학기술원 (KAIST) 시스템공학연구소 연구원
1989 University of Oklahoma 대학원 전산학 이학석사
1995 University of North Texas 대학원 전산학 이학박사
1996-1997 동명정보대학교 컴퓨터공학과 조교수
1997-현재 대구교육대학교 전산교육과 부교수