Intelligent Agent System by Self Organizing Neural Network

Young Im Cho
Department of Computer Science, The University of Suwon, Seoul, Korea
(Tel: +82-31-229-8214; E-mail: ycho@suwon.ac.kr)

Abstract: In this paper, I proposed the INIntel1gent Agent System by Kohonen's Self Organizing Neural Network (INTAS). INTAS creates each user's profile from the information. Based on it, learning community grouping suitable to each individual is automatically executed by using unsupervised learning algorithm. In INTAS, grouping and learning are automatically performed on real time by multiagents, regardless of the number of learners. A new framework has been proposed to generate multiagents, and it is a feature that efficient multiagents can be executed by proposing a new negotiation mode between multiagents.

Keywords: multiagents, self organizing neural network, e-learning

1. INTRODUCTION

As many researches have shown, the dropout rate in e-learning is higher than that in traditional face-to-face learning due to its low degree of continuity. In order to lower this dropout rate, many researches have been done to heighten the degree of learners' satisfaction and to provide them with motivation. Tinto [2] argues that the formation of a strong learning community can solve this problem. A strong learning community can eliminate the feeling of isolation, facilitate interaction with other learners and assist learning. In e-learning, learning is accomplished when education of true value which satisfies the demands of the members of a learning community is done through interaction among one another and common goals can be met among communities. Rovai [3] argues that the seven factors which positively correlate to a sense of community are transactional distance, social presence, social equality, small group activities, group activities, group facilitation, teaching style and learning stage and community size.

The Agent system, which began to appear in the 1990s, is a system which is automatically managed and self-operative. It is a very intelligent concept which can manage the information of each learner in the e-learning system, and recommend and search information that fits the inclination of each individual [4]. By applying the concept of the agent to the e-learning system, we can develop the next generation's technology which will contribute to the increase of the degree of satisfaction of learners, as well as the degree of learning achievement, by analyzing the inclination of each individual learner and reflecting its result in each group.

In this paper, as I recognize the importance of a learning community and intend to form a learning community which is strong and at the same time, the most feasible, I will develop through the questionnaire called the inclination test, Intelligent Agent based Learning Community Grouping e-learning System, with the method of artificial intelligence based agent which will reflect inclination and characteristic of an individual learner.

The structure of this paper is as follows. Chapter 2 will explain theoretical research concerning the online learning community, motivation theory and agent theory, and will explain evaluating the list of the inclination test. Chapter 3 will explain the system overview, module feature and algorithm of INTAS (INTElligent Agent System). Chapter 4 will build and evaluate the INTAS. And Chapter 5, the final chapter, will draw a conclusion.

2. BACKGROUNDS

3.1 Agent system

The agent system is an independent area that has been studied in the field of artificial intelligence for a long period of time. It has been recognized as a separate field of research since the early 1990s. It signifies a system which perceives the environment through its own sensors and takes actions against environments with its effectors [4, 17]. The agent system is an autonomous process which solves tasks wanted by a user automatically, as a substitute for the user. It operates mainly in distributed environments. It is a software system with an independent function which can perform its own task by itself.

Distributed Artificial Intelligence (DAI) is a concept that first appeared near the end of the 1970s. Its researches are concentrated on the fusion of multi-agent systems and distributed problem-solving systems. [18,19] In multi agent systems, each agent performs a job in part and the overall goal is achieved through interaction among the agents. The most representative paradigm of such an interactive adjustment is the FA/C(Functionally Accurate Cooperative paradigm) [20].

In the area of e-learning, individual learning achievement should be regarded as an important goal. Therefore, the e-learning area is not a generalized, homogeneous system, but should be regarded as a field where individual learning achievement is reflected in a very sensitive manner. It is necessary to construct a personalized learning system on the basis of a strong learning community and the motivation theory as explained before.

3.2 E-learning system

For the grouping of learning communities, one should log in online through an inclination test. The degree of the satisfaction of learners will be evaluated inside the system and its result will be reflected on subsequent learning. As for the questionnaire of the inclination test list, users have themselves taken the pilot-test, the purpose of which was to examine whether the contents are properly understood and to determine concrete criteria for evaluation. After this, we have gone through the process of correction and supplement.

A member of a general online community will join a chatting room, a mass game or a café with a mailing list or a subject for discussion, in virtual space characterized by anonymity and openness. But in case of a learning community, a user should in general log in and register directly at the
learning site. Therefore, it is in most cases characterized by actuality and approximately.

In this paper, members of a learning community can be compared to teachers of elementary, junior and senior high schools, and the questionnaire was designed on the assumption that they enroll in the learning site for teachers' training. It is supposed that at least fifty thousand people register at the learning community for training. We intended to practice the grouping characteristics of each learner through an agent so that a learning community can complete learning effectively and successfully by maintaining participation motivation and continue the learning without dropping out.

In order to describe concrete social and cultural contexts which are important factors in constituting a cyber learning community, we include in this paper, gender, online training experience, region(area) and the duration of computer use in the category of diversity. In the category of homogeneity, we include major, teaching career, sports and hobby, favorite food and favorite color, in order to constitute a group with a homogeneous inclination.

The dropout rate of e-learning is higher than that of traditional face-to-face learning by 10 to 20%. The reason is that it puts more emphasis on grades than on learning, and also that the participants are physically apart from one another. Therefore, in order to lower dropout rates and increase learning effects, a social environment should be continuously provided to the learners, and they should be induced to participate in learning activities through provision of a strong sense of community.[21]

On the basis of these researches, this paper presents a list of homogeneity and heterogeneous items in inclination test.

As for a questionnaire, 10 items are included which are considered to be adequate for grouping, according to the characteristics of each category. First, among 6 items for homogenization, the major category signifies a teacher's major, which are then classified into such sub-items as Korean language, mathematics, English, science, social studies, etc. Teachers in junior and senior high schools are supposed only to select the class which they are currently teaching. As for elementary school teachers, their own major should be selected. The teaching career category is classified into less than one year, one to five, six to ten, eleven to fifteen, sixteen to twenty and more than twenty years. As for favorable sports, they are classified into mountain climbing, workout, golf, tennis, swimming and others. The hobby categories are movie, Paduk, fishing, reading, web surfing, game and etc. And favorite food and favorite color are also classified into sub-items. But those learners who have similar inclination may show cultural diversity and have diverse opinions. So we have added diversity by selecting four items such as online training experience, residential area, gender and duration of computer use.

In each category of the inclination test, priority is determined depending on the characteristics of test content. The advantage of prioritizing the characteristics for homogenization is to form a strong learning community, as this learning seeks to accomplish, and eliminate defects of online learning. That is to say, the degree of familiarity can be increased at the initial stage through working together with those who share similar inclinations and eliminate feeling of isolation that may be caused by online learning. Also, collaboration and communication, which are the foundation of a learning community, can be promoted, and the common denominators among learners can be found and suggested as a learning project. We also have included diversification categories in order to supplement diverse learning experience problems that may be caused by homogenization.

I have suggested the major category as the first standard because it is appropriate to perform project centered around majors, due to the official appropriateness of the major classification itself and also because it is an area which learners can usually find subjects in common. And we have selected 'teaching career' as the second standard for homogenization because teachers with a similar teaching career tend to share many common topics for conversation. And the order of other categories have been determined according to their proximity to 'major' as having the highest priority.

Thus the inclination test list have been presented to learners and teachers, following the general format of a questionnaire survey. Such questions as teaching career, gender and region deemed as general characteristics are asked first and then the other questions are presented in order.

As for regional items, I have included them in this paper in the category for diversification in order for opinions of diverse regions to communicate with one another. As for gender, women and men exist in diverse ways, and we have made it a standard for diversification items in order for men and women to share opinions and ways of thinking that can be originated uniquely from a certain gender. I have put duration of computer use and online training experience in diversification category because skillful computer users may help those who are not as experienced and also because learners with online training may help those without it.

### 3. INTELLIGENT AGENT SYSTEM

#### 3.1 Overview

Information from individual learners through security and certification procedure as seen in Figure 1 is inputted to the system, Intelligent Agent based e-learning System(INTAS), to be proposed in this paper, and INTAS creates each user's profile from the information. Based on it, learning community grouping suitable to each individual is automatically executed by using Self Organizing Feature Map (SOM) learning algorithm via multi agents.

In INTAS, grouping and learning are automatically performed on real time by multi-agents, regardless of the number of learners. A new framework has been proposed to generate multi agents, and it is a feature that efficient multi agents can be executed by proposing a new negotiation mode between multiagents.

![Figure 1. INTAS Configuration](image-url)
3.2 INTAS Module Features and Algorithm

3.2.1 User Profile

Grouping agent is generated by learner's drawing up a distribution map related to items using Kohonen's SOM learning algorithm, based on inputted information for homogeneity and heterogeneity. As in Figure 3, when homogeneous and heterogeneous items are inputted respectively, input vector is generated in order pairs, each. Then, learning grouping is automatically executed with the weight provided by drawing up a categorization map on real time through SOM network [22].

The explanation of user profile drawing method regarding each number of the above Figure 3 is as follows:

1. Input vector is generated with regard to learner's input value for homogeneous and heterogeneous categorization criteria.
2. A distribution map is drawn up by providing weight to detailed homogeneous and heterogeneous items via SOM network.
3. The SOM network configured in this paper is a 2 layer structure, which consists of n input nodes expressing n dimensional input data and k output nodes to express k categorization sectors. All the input nodes are connected to all the output nodes and have connection weight. Learning algorithm for SOM categorization is as follows:

   - Step 1: Initialize connection weight \( W_0 \).
   - Step 2: Put input vector in the input node \( X \).
   - Step 3: Select a WIN in the output node.
   - Calculate the distance \( D \) between input vector and connection weight using the following formula:
     \[
     D(X, WIN) = \frac{1}{1 + \frac{1}{2} \sum_{i=1}^{n} (X_i - WIN_i)^2}
     \]
   - Step 4: Learn connection weight.
   - Learn the WIN and the node within the neighboring radius as seen in the following formula:
     \[
     WIN_{new} = WIN_{old} + \eta \cdot \Delta W
     \]
   - After learning, reduce a value.
   - Step 5: Go back to step 2 and repeat the above process. If \( a = 0 \), finish learning.

The WIN of input vector in the SOM is decided by the extent of similarity between input vector and weight vector. In the SOM, neighborhood function is used to renew nodes around the WIN; Gaussian function was used in this paper.

- After forming a categorization map using 2 dimensional SOM network from the distribution of homogeneous and heterogeneous items, draw up a 3 dimensional categorization map by mapping the values excluding concerned heterogeneous categorization map value regarding concerned homogeneous categorization map value with regard to input vector I.
- When the number of learners of each grouping exceeds prescribed number of group, divide them adequately. The number of each group needs to be comprised for a manager to freely input at early stage.

The update process of user profile used in the INTAS system is executed by the user profile update agent:

- After generating user profile after real time grouping by using initial input items, update the user profile when requests of a learner comes in additionally or when the learner's work information has been updated.

3.2.2 DIMAF Multi Agent Framework

When an agent is generated, DIMAF consists of the negotiation algorithm between the agent name server (ANS) providing agent ID, an agent manager controlling and monitoring generation, execution and movement of agent and multi agents.

The negotiation algorithm is greatly required for suitable grouping from grouping list by searching learned user profile with user input item inputted at an early stage in the grouping agent. In the grouping list, the user (ID), group number (G), satisfaction degree (SD) and team information (TI) are recorded. TI is recorded as a value among maintenance (M), don't care (D) and break (B). Maintenance (M) is the case where satisfaction degree of previous group members is very high, which means the value is required to be maintained constantly, not desiring to break. Don't care (D) is the value meaning that it may be changed, according to learners' responses in the normal position. Break (B) means the group to be regrouped, after breaking existing group, since satisfaction degree of the previous group members is very low.

The detailed negotiation process by negotiation algorithm proposed in this paper is as follows:

- Step 1: Grouping agent searches concerned individual (ID), group number (G), satisfaction degree (SD) and team information (TI) from grouping list by inspecting user profile from the homogeneous items categorized primarily. If a concerned ID's TI value is M, the concerned grouping is maintained without executing 2nd step and you need to move to step 4. If TI value is D or B, you need to move to step 2 and continue.
- Step 2: From the table saved in the temporary storage, grouping agent (GA) calculates G and SD which performed grouping by SOM learning algorithm by using user input homogeneous item. If TI value was B, you need to move to step 4, beyond step 3.
- Step 3: When the grouping result value performed by GA in step 2 and the G value of grouping list are different, concerned grouping should be maintained in the user ID with priority in the result of grouping list in the user profile. However, concerned user should judge by showing the group member list to concerned user (ID).
- Step 4: Show grouping information and member list to each learner.

3.2.3 Multiagents

Four multiagents are generated basically in the INTAS system. Grouping agent (GA) is the agent that generates user profile using 1st categorization criteria (homogeneous) and 2nd categorization criteria (heterogeneous). GA is in charge of grouping as explained in 3.2.1. User profile update agent (UA) is the agent that saves user history and helps grouping performance, while consulting GA.

Evaluation agent (EA) evaluates learning satisfaction degree of user and grouping members and decides whether to
maintain, don't care or break this group, according to satisfaction degree value. In the INTAS system, a learner evaluates the group set by grouping with 5 patterns (very satisfactory, satisfactory, general, unsatisfactory, very unsatisfactory).

EA is the agent playing a role to provide a learner with guideline on whether to leave the concerned group, based on group evaluation data to which the learner belongs and his/her own input data. And, EA automatically moves the learner to a new group suitable and not exceeding prescribed number of the group, when the user decides to leave the former group. The presentation of guideline to leave the concerned group from EA is made up by using error back propagation learning algorithm of neural network.

Monitoring agent (MA) is the agent to identify state of a learner (user) by monitoring the number of grouping, number per group and satisfaction degree per group graphically through monitoring of learners' learning status.

4. INTAS PERFORMANCE EVALUATION

4.1 Implementation of INTAS

The INTAS system developed in this paper has configured the Web server using Linux server, and database has been built using Mysql. Web program was comprised using ASP and PHP, and the agent's source code was configured with Java using DIMAF framework.

In INTAS, grouping forms group via grouping agent. The grouping agent indicates homogeneous and heterogeneous distribution from homogenous and heterogeneous items selected by a user through the use of SOM algorithm. Automatic grouping is made by learner's input with this distribution.

Among homogeneous items, user selects detailed items (i.e. major subject is Korean) regarding each item. In the INTAS, input nodes and random $10^5$ output nodes are provided for a learner to learn using learning algorithm of the SOM network regarding each detailed input value of the user-input homogeneous value. Here, the reason why 100 output nodes are provided is because maximum number of cases in which homogeneity can be generated is limited to 100.

With regard to input value, values were randomly generated in order of major subject, favorite sports, etc, giving priority to each item by valuing homogeneous values numerically.

For example, ID: yicho1234, major subject: Korean, teaching experience: 1-5 years, favorite sport: swimming, hobby: movie, favorite food: Chinese food, favorite color: yellow were selected, they are expressed in the following data structure order by with priorities in order. The data structure of concerned ID input node is as follows:

<table>
<thead>
<tr>
<th>Input Node</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>yicho1234</td>
</tr>
<tr>
<td>3</td>
<td>Korean</td>
</tr>
<tr>
<td>4</td>
<td>1-5</td>
</tr>
<tr>
<td>5</td>
<td>swimming</td>
</tr>
<tr>
<td>6</td>
<td>movie</td>
</tr>
<tr>
<td>7</td>
<td>Chinese food</td>
</tr>
<tr>
<td>8</td>
<td>yellow</td>
</tr>
</tbody>
</table>

User satisfaction degree is performed by evaluation agent. User(learner) satisfaction degree has 5 input values and the user satisfaction degree is evaluated in three such patterns as maintenance (M), don't care (D) and break (B) through error back propagation (EBP) learning algorithm.

EBP is comprised of three layers (input layer, middle layer and output layer), and the learning is performed by changing connection weight through generalized delta rules. Learning through teacher is carried out by receiving the output layer data as input data with regard to input layer data. The input layer indicates satisfaction degree of users as values. EBP has connection weight (W) in each layer, and the data inputted into input layer (learning data: input) is delivered to middle layer in a forward direction, and provides result value in the output layer.

Perform learning, while adjusting connecting weight (W) in the direction of reducing difference by comparing the result value and expected data value (learning data: output), and by controlling connection weight (W) of own layer in the lower layer, based on back propagation in the upper layer.

The status of incurred value from this stage (maintenance, don't care, break) is updated by each user via the user profile update agent, and the information is maintained and reported to each user.

4.2 Performance Evaluation

Through homogeneous and heterogeneous categorization maps, a total of 151 groups, which reflected homogeneity and heterogeneity by the number of learners per group (size of
learning community. In this paper, it is 10), were generated via homogenous and heterogeneous categorization maps as well as via result to be grouped in the IN TAS in consideration of number of learners per group (size of community). The reason why 151 groups were generated, despite members were 1,000, is because they are comprised of only members within homogenous categorization map (user with the same homogeneity). In this way, groups reflected homogeneity and heterogeneity were formed.

It is difficult to talk about adequate community size, since it differs according to curricular content, teachers or learners, but people from 8-9 to 20-30 are said to be adequate.

For example, there are 39 members in homogenous groups, G00, and number of people per group is 10, a total of four final learning groups (g1, g2, g3, g4) are generated reflecting heterogeneity, and the number of members of g1, g2 and g3 are 10, and that of g4 will be 9.

In this paper, a pilot test was conducted to evaluate actual users of the IN TAS system. As a result, the evaluation of user satisfaction degree per group is seen in Figure 2.

If the scale indicates 1 in satisfaction degree, it means very satisfactory, if the scale is 5, then it means very unsatisfactory. The scale from 1 to 5 with regard to 5 categories was expressed as value. Here, satisfactory means that members of a group or learning desire shows a very positive result.

![Average Satisfaction Degree per Group](image)

Figure 2. Satisfaction Degree per Group

As a result of surveying users' group satisfaction degree regarding 151 groups in the above figure, we can see the average was distributed around scale value 2 of satisfaction degree. Accordingly, when automatic grouping was performed by agent, learners were generally satisfied.

When they desired to maintain their group according to the value 2, around 51 groups of total 151 groups (34%) showed in favor of maintenance, while 54% showed don't care and less than 12% showed break.

When calculating average of the ratios of 6 items in the homogenous group. When considering only homogenous group items, major subject (68%), teacher's experience (15%), favorite sports (9%), hobbies (4%), favorite food (35) and favorite color (15) were distributed in order. Namely, major subject and teacher's experience can be important factors to decide the homogenous group.

The distribution of heterogeneous items (area, gender, computer using hours, online training experience) in relation to 449 users of 51 homogenous groups, which desire strong maintenance is indicated.

The criteria evaluating group satisfaction degree of learners have been indicated in order of online training experience, computer using hours, gender and area.

Accordingly, when learners are grouped in a group with greater online training experience or computer using hours, the learners' satisfaction degree became greatly higher; and thus it was discovered that they desired to maintain the group continuously.

5. CONCLUSIONS

In this paper, I have made an inclination test questionnaire for the formation of effective and efficient online learning community. And we have designed and realized an automatic grouping system with information of learners that appear through the questionnaire and by using an intelligent agent.

Through this automatic grouping of the learning community, first, the initial motivation of Keller can be satisfied by having friends with a similar inclination work together and gaining familiarity at the initial stage, thus eliminating the feeling of isolation that might have resulted from a lack of face-to-face directness in e-learning.

Secondly, continuity motivation can be given by promoting collaboration and communication which are the foundation of a learning community.

Third, common denominators among learners can be found and correlated into the learning projects or learning. On the other hand, the category of diversification has also been included to supplement diverse learning experiences that might be harmed by homogenization.

The results of our experiment with 1,000 people in reality by means of developing the grouping system have shown that 151 groups are automatically formed. Among them, 34% have shown very high degree of learning satisfaction and intended to maintain the groups in the future. In terms of homogenization, teachers who share the same majors and have longer teaching career have formed a group with a higher degree of homogenization.

As for diversification grouping, online training experience and longer computer use have resulted in the increase of the degree of learning satisfaction.

In this paper, with the results of the inclination test designed to form a desirable online learning community, and through grouping systems that utilize the intelligent agent system, we have found out the possibility of lessening the feeling of isolation and lowering dropout rates that may frequently occur in online learning.

As for the degree of satisfaction concerning the grouping system, within the category of hobby, the group of those who selected fishing as their favorite hobby have shown a higher degree of satisfaction (45%) than other groups.

The importance of grouping is revealed through the inclination test. When group members are satisfied, dropout rate stemming from feeling of isolation will decrease and thus maintenance and continuity motivations can be maintained.

In the future, it is necessary to improve services concerning the communication between users by supplementing the grouping system in ways which more users can use it on the web, and to continue research on whether learning can be achieved effectively and efficiently through automatic grouping.
REFERENCES

[16] D. Fulcher, and Debbie, Lock, "Distance education: The future of library and information services requirements", Distance Education. vol.20, no.2, 1999.

http://www.gurugail.com/Agent/page.html?subject=mobi leAgent.html