IMPROVEMENT OF QUANTITATIVE LEARNER'S MOTIVATION METHOD BY OPTIMIZING THE COMBINATION OF THE ELEMENTS

Shun TERASHITA[1], Fumito MASUI[2] and Michal PTASZYNSKI[3]

[1]Master's Course in Computer Science, Graduate School of Engineering, Kitami Inst. of Tech.
E-mail: m2052400044@std.kitami-it.ac.jp
[2]Professor, Dept. of Computer Science, Kitami Inst. of Tech.
E-mail: f-masui@mail.kitami-it.ac.jp
[3]Associate Professor, Dept. of Computer Science, Kitami Inst. of Tech.
E-mail: michal@mail.kitami-it.ac.jp

(Koencho 165, Kitami, Hokkaido 090-8507, Japan)

Nobuta et al.[1)] proposed a method to estimate learner's motivation by analyzing results of class evaluation questionnaires. Through experiments, they confirmed that their method performed effectively. On the other hand, they also pointed out that the method will be improved by optimizing the combination of the three elements of which learning motivation consists.

In this paper, we extended Nobuta's method by optimizing the three elements combination. Specifically, we investigated the degree to which these elements affect the estimation of learning motivation. We applied principal component analysis to the elements and identified the optimal combination. Using the optimal combination, we conducted an experiment on the estimation of learning motivation. The results showed that it significantly outperformed the former method.

Furthermore, it was shown that the proposed method could estimate the motivation for learning stably even when the group of students in each course, the class format, and the class design were different.

Keywords : *information engineering, course analysis, course evaluation questionnaires, learning motivation, educational technology, ARCS model*

1. INTRODUCTION

Nobuta et al.[1)] proposed quantitative learner's motivation method by analyzing results of course evaluation questionnaires to support a more accurate implementation and utilization of course questionnaires. Inspired by the ARCS model[2)3)], this method defines three elements: (1) interest in the course, (2) usefulness of the course for the future, and (3) expectation (satisfaction) of the course. The results of the questionnaire are used to quantitatively estimate the motivation to learn. Furthermore, experiments have confirmed that the motivation to learn can be estimated with a certain level of performance. However, they also pointed out that the method will be improved by optimizing the combination of the three elements of which learning motivation consists.

We extended the method of Nobuta et al.[1)] by optimizing the three elements combination[4)]. Specifically, we investigated the degree to which these elements affect the estimation of learning motivation. The element combinations were optimized using factor analysis and principal component analysis. To evaluate an element optimization effectiveness for the learning motivation estimation performance, the optimal solution was applied to Nobuta's method. The results showed that it significantly outperformed the former method. However, the general-purpose effectiveness of our approach is not confirmed yet[4)].

In this study, we evaluate the general-purpose effectiveness of our method by utilizing the questionnaire results on multiple courses.

Specifically, the element combinations was optimized by using principal component analysis. The obtained optimal solution was used to estimate the learning motivation.

In this paper, Section 2 describes related research, and Section 3 describes the proposed method. Next, Section 4 describes the experiments to evaluate the effectiveness of the proposed method. Finally, in Section 5, conclusions and future perspectives are given.

2. RELATED WORK

There have been a number of studies that attempted to model learners' motivation in a systematic way.

Keller[2)3)] proposed the ARCS model to provide guidelines for designing attractive courses and teaching materials. This model is based on the analysis of a large amount of data collected by psychological research and factorizes learning motivation into four categories: (1) Attention, (2) Relevance, (3) Confidence, and (4) Satisfaction. However, the ARCS model is a qualitative model of learning motivation, and it cannot quantitatively express the degree of motivation and its change.

Nobuta et al.[1)] applied the ARCS model to quantitatively express learning motivation in terms of three elements: (1) interest in the course, (2) usefulness of the course for the future, and (3) expectation (satisfaction) with the course. By quantifying and categorizing these elements using ratings from course surveys, they estimated learner motivation in three levels. They assume that these three elements have an equal impact on learning motivation. In general, however, it is not equal.

In this study, we extend Nobuta's method by optimizing the three elements combine and aim to improve the estimation performance of learning motivation.

3. PROPOSED METHOD

In this study, we extended Nobuta's method by optimizing the three elements combination. The following are some of the questions in the course questionnaire used to estimate the learning motivation. These questions were designed by Nobuta et al.[1)]. All these questions are based on the 5-point scale method.

(1) Did you have an interest in this course?
(2) Do you think participating in this course will help in the future (further college life or job hunting)?
(3) Were you satisfied with this course?
(4) Did you attend this course with a desire to learn?

Questions (1) through (3) above ask about the elements of learning motivation. The rating values obtained from these questions were used to estimate the motivation to learn. Question (4) is a question that asks students about the degree of their own motivation to learn (hereinafter referred to as "self-evaluation"). The accuracy of the estimation is evaluated by comparing the estimated learning motivation using each of the self-evaluation and questions on the three elements.

The procedure for estimating the learning motivation is shown below. Assume that questionnaire $Q = \{q_1, q_2, \ldots, q_i \ldots, q_m\}$ consists of m questions. The Q is then applied to a learner group $X = \{x_1, x_2, \ldots, x_j, \ldots, x_n\}$ containing n number of persons.

■Step 1 The values of evaluation $r_{i,j}$ from the answers of learners x_j for the questionnaire q_i are collected and for each question mean average μ_i and standard deviation σ_i of the respondents are calculated. In this case, μ_i is calculated according to equation 1, and σ_i is calculated according to equation 2.

$$\mu_i = \frac{1}{n}\sum_{j=1}^{n} r_{i,j} \tag{1}$$

$$\sigma_i{}^2 = \frac{1}{n-1}\sum_{j=1}^{n}(r_{i,j} - \mu_i)^2 \tag{2}$$

■Step 2. Classify $r_{i,j}$ for each of the three evaluated items using as threshold σ_i, as in equation 3 and produce scores for all elements $s_{i,j}$. Next, calculate the scores $s_{i,j}$ of each learner x_j for each question n times.

$$s_{i,j} = \begin{cases} 1 & if\ r_{i,j} \geq \mu_i + \sigma_i \\ -1 & if\ r_{i,j} \leq \mu_i - \sigma_i \\ 0 & otherwise \end{cases} \tag{3}$$

■Step 3. The principal component analysis is performed using the rated values $r_{i,j}$ of each of the three questions to obtain the principal component loadings pcl_i for each question. Then, among the three pcl_i, only the two pcl_i with relatively large values are used to calculate the average value pcl_{ave}. For example, when the size relationship of pcl_i is $pcl_3 < pcl_2 < pcl_1$, then pcl_{ave} is calculated according to equation 4.

$$pcl_{ave} = \frac{pcl_1 + pcl_2}{2} \tag{4}$$

■Step 4. Calculate the weight $weight_i$ for the element score $s_{i,j}$. The $weight_i$ is calculated according to equation 5. Note that $weight_i$ for $s_{i,j}$ with the smallest pcl_i is set to 0.

$$weight_i = \frac{pcl_i}{pcl_{ave}} \tag{5}$$

■Step 5. Sum the $s_{i,j}$ for each learner and predict the learner's motivation M_j, according to equation 6.

$$M_j = \sum_{i=1}^{n} weight_i * s_{i,j} \tag{6}$$

■Step 6. The learner's motivation M_j obtained in Step 5 has the following properties: $\{M_j \mid M_j \leq \mid m \mid, M_j \in Q\}$. Therefore, M_j is classified into the following clusters according to equation 7.

$$Learner's\ motivation\ is \begin{cases} high & if\ M_j \geq \ 1 \\ low & if\ M_j \leq -1 \\ moderate & otherwise \end{cases} \tag{7}$$

4. EXPERIMENTS

(1) Experiments

Experiments were conducted to evaluate the general effectiveness of the proposed method.

For the experiment, it used 2,577 questionnaire responses in 40 groups of students from first-year to third-year undergraduates who took courses at Kitami Institute of Technology from 2014 to 2016. These questionnaire responses include four types of courses: (1) lectures, (2) exercises, (3) experiments, and (4) lectures with exercises.

In the experiment, learning motivation was estimated for each group using the proposed method and the previous method, and the estimation performance was evaluated.

When using the previous method, the operations except for Step 3 and Step 4 of the proposed method described in the previous section are performed. Note that the $weight_i$ in Step 5 is ignored.

The following evaluation indices are used to evaluate the performance of the estimation of learning motivation: precision (P), recall (R), and F-measure (F). These metrics are calculated according to Equation 8, Equation 9, and Equation 10, respectively.

Precision is the number of responses that are consistent with the student's self-evaluation among the number of responses estimated using the three components of learning motivation. The recall is the number of responses that are consistent with the student's self-evaluation out of the total number of responses estimated by the student's self-evaluation. F-measure is the harmonic mean of precision and recall and is a performance evaluation index that has been used in the fields of information retrieval and natural language processing.

$$P = \frac{n}{A} \tag{8}$$

n:Number of predictions based on three elements matching self-evaluation

A:Number of all responses predicted using three elements

$$R = \frac{n}{B} \tag{9}$$

n:Number of predictions based on three elements matching self-evaluation

B:Number of all responses inferred by self-evaluation of learners

$$F = \frac{2 * P * R}{P + R} \tag{10}$$

(2) Results and Discussion

Table 1 shows the estimation performance for each cluster. Note that "S.D." in the tables indicates the standard deviation.

The average value of F-measure represents the performance of the estimation for the entire cluster. The mean F-measure for the proposed method and Nobuta's method are 0.68 and 0.65 respectively. It shows that the overall estimation performance of the proposed method overcomes the previous method. Also, the mean of precision and recall for the proposed method were 0.69 and 0.71 respectively, showing no issues with bias or variance.

Therefore, we can confirm the general effectiveness of the proposed method.

Second, it considers the estimation performance of each cluster. The F-value for "High Motivation" and "Moderate" were higher than the previous method. The F-measure of "Low Motivation" was also lower than the previous method.

The estimation performance of the proposed method varied depending on the cluster as the F-measure of "Moderate" was 0.81, while the F-measures of "High Motivation" and "Low Motivation" were only 0.65 and 0.57 respectively. This trend is consistent with the previous method.

To improve the estimation performance, the factors that are not correctly estimated must be investigated. In the future, we plan to compare clusters of responses that failed to be estimated by both methods.

Finally, it describes the estimation performance of each group of students. The estimation performance in **Table 2** shows the average of the precision, the recall, and the F-measure for each year. From **Table 2**, it can be seen that the estimation performance for the year 2014 is the highest for both methods.

Additionally, it shows that the estimation performance is

Table 1 Evaluation results for each cluster

	Learner's Motivation	Precision	Recall	F-measure
Nobuta's method	High Motivation	0.49 ± 0.17S.D.	0.90 ± 0.11S.D.	0.62 ± 0.15S.D.
	Moderate	0.89 ± 0.11S.D.	0.63 ± 0.16S.D.	0.72 ± 0.14S.D.
	Low Motivation	0.54 ± 0.25S.D.	0.82 ± 0.14S.D.	0.61 ± 0.21S.D.
	Average	0.64	0.78	0.65
Proposed method	High Motivation	0.61 ± 0.20S.D.	0.74 ± 0.20S.D.	0.65 ± 0.16S.D.
	Moderate	0.83 ± 0.13S.D.	0.81 ± 0.14S.D.	0.81 ± 0.11S.D.
	Low Motivation	0.64 ± 0.26S.D.	0.58 ± 0.26S.D.	0.57 ± 0.21S.D.
	Average	0.69	0.71	0.68

Table 2 Evaluation results for each year

Year	Nobuta's method			Proposed method		
	Precision	Recall	F-measure	Precision	Recall	F-measure
2014	**0.66**	**0.81**	**0.68**	**0.74**	**0.76**	**0.74**
2015	0.65	0.76	0.66	0.68	0.69	0.65
2016	0.62	0.78	0.63	0.66	0.69	0.65
Average	0.64	0.78	0.65	0.69	0.71	0.68

Table 3 Results of Chi-square(χ^2) test

Course Combinations	χ^2	p-value
H vs A	0.004	0.948
H vs I	0.355	0.552
A vs I	0.426	0.514

different for each course (see Appendix A). This result suggests that the estimation performance may differ depending on the course format and the students in each course.

Therefore, we conducted a chi-square test for the estimation performance of the two courses.

For the test, it was used the responses to the questionnaire administered in 2014 for three courses: Course H (lecture format), Course A (exercise format), and Course I (experimental format).

The results are shown in **Table 3**. For example, "H vs A" means the combination of Course H and Course A. χ^2 represents the chi-square value. The p-value is the statistic value given from the chi-square test.

For all of the combinations, the p-value was higher than 0.05 so none of the combinations showed statistical significance. There are three possible reasons for this. First, the proposed method could be versatile enough. There might have also been no difference in the sets of respondents or the format of the courses.

However, since the test used responses from courses with obviously different formats, such as lectures and experiments, it can be said that both the sets of respondents and the course formats differed.

From this, we can conclude that the proposed method has general applicability.

Furthermore, **Table A** in Appendix A shows some interesting results. The results for 2014 and 2015 show that Precision, Recall, and F-measure were always the highest or lowest for one particular course. In 2014, these values were the highest for course C with both methods. In 2015, when using the previous method, these values were the highest for course A and the lowest for course B(1).

However, at this time, this factor has not been identified. In the future, we are planning to clarify the situation by further analyzing the questionnaire results.

5. CONCLUSIONS

In this study, we tried to evaluate the general-purpose effectiveness of the Learning Motivation Quantification method by using the questionnaire results from multiple courses. As a result, the estimation performance of the proposed method for the entire cluster was higher than that of the former method, and the effectiveness of the proposed method was confirmed to a certain extent. Furthermore, the results of the experiment showed that the proposed method could stably estimate the learning motivation even when the students in each course, the course format, and the course design were different.

The estimation performance of “High Motivation" and “Moderate" overcame that of the former method, while the estimation performance of “Low Motivation" was lower. Similar to the previous method, the estimation performance of “High Motivation" and “Low Motivation" was lower than that of “Moderate".

In order to further improve the estimation performance, it is necessary to clarify the factors that prevented the correct estimation.

In the future, we will compare the responses that were correctly estimated for each cluster with those that failed. We also plan to compare the responses that were consistent with the estimation results of the conventional method and those that were not.

ACKNOWLEDGEMENT We would like to express our deepest gratitude to the faculty and students of Kitami Institute of Technology for their cooperation in this research.

APPENDIX A

Table A shows the estimation performance of each course. The estimated performance in the table shows the average of the precision, the recall, and the F-measure for each course.

The F-measure of the proposed method for 25 groups overcome the previous method, while that of 12 groups was lower. The performance of the 3 groups was the same as the previous method.

REFERENCES

1) Yuka Nobuta, Fumito Masui, and Michal Ptaszynski. Modeling Learning Motivation of Students Based on Analysis of Class Evaluation Questionnaire. In *Technical Transactions of Mechanics*, Vol. 2-M(7), pp. 193–201. WYDAWNICTWO POLITECHNIKI KRAKOWSKIEJ, 2015.
2) Keller J.M. and Kopp T. Application of the ARCS model of motivational design. In C.M.Reigluth, editor, *Instructional theories in action : Lessons illustrating selected theories and models*. Lawrence Erlbaum Associates, U.S.A., 1987.
3) Keller J.M. and Suzuki K. Use of the ARCS motivation model in courseware design (Chapter 16). In D.H.Jonnasen, editor, *Instructional designs for microcomputer courseware*. Lawrence Erlbaum Associates, U.S.A., 1988.
4) Shun Terashita, Fumito Masui, and Michal Ptaszynski, editors. *Improvement of learning motivaton quantification method by introducing multivariate analysis*. The 2021 Annual Spring Conference of Japan Society for Educational Technology, 2021(In Japanese).

Table A Evaluation results for each course

Year	Course	Number of responses	Nobuta's method			Proposed method		
			Precision	Recall	F-measure	Precision	Recall	F-measure
2014	A	73	0.54	0.78	0.55	0.80	0.84	0.82
	B(1)	72	**0.47**	**0.69**	**0.42**	0.70	0.67	0.68
	B(2)	74	0.75	0.89	0.78	0.80	0.86	0.82
	C	24	**0.89**	**0.96**	**0.91**	**0.93**	**0.98**	**0.95**
	D	62	0.68	0.71	0.68	0.69	0.69	0.69
	E	69	0.71	0.75	0.71	0.64	**0.43**	**0.59**
	F	58	0.58	0.77	0.61	0.67	0.72	0.67
	G	51	0.76	0.85	0.80	0.80	0.82	0.80
	H	75	0.61	0.87	0.66	0.68	0.80	0.72
	I	60	0.82	0.84	0.83	0.88	0.82	0.85
	J	57	0.54	0.79	0.53	0.75	0.90	0.79
	K	154	0.56	0.77	0.61	**0.57**	0.64	0.59
	L	55	0.67	0.84	0.70	0.66	0.71	0.66
	Average		0.66	0.81	0.68	0.74	0.76	0.74
2015	A	74	**0.75**	**0.88**	**0.79**	0.78	**0.83**	**0.80**
	B(1)	55	**0.55**	**0.61**	**0.55**	0.61	**0.62**	0.62
	B(2)	64	0.59	0.78	**0.55**	0.57	0.78	0.56
	C	36	0.56	0.72	0.57	**0.54**	0.65	**0.55**
	D	57	0.72	0.72	0.72	0.69	0.68	0.69
	F	53	0.57	0.74	0.56	0.69	0.63	0.66
	G	44	0.70	0.75	0.71	**0.81**	**0.62**	0.59
	H	45	0.74	0.85	0.76	0.64	0.68	0.58
	J	42	0.70	0.84	0.73	0.75	0.68	0.69
	K	152	0.64	0.77	0.66	0.70	0.76	0.72
	L	67	0.64	0.70	0.65	0.74	0.68	0.69
	Average		0.65	0.76	0.66	0.68	0.69	0.65
2016	A	60	0.67	0.85	0.72	0.84	0.83	0.82
	B(1)	59	0.60	**0.67**	0.57	0.67	0.63	0.58
	B(2)	70	**0.45**	0.69	**0.35**	**0.45**	0.57	**0.42**
	C	22	0.64	0.75	0.68	0.50	**0.50**	0.75
	D	54	0.72	0.77	0.73	0.72	0.70	0.71
	E	51	0.64	0.78	0.66	0.69	0.67	0.68
	F	60	0.54	0.73	0.51	0.54	0.73	0.54
	G	45	0.72	0.87	0.75	0.66	0.74	0.64
	H	47	0.53	**0.90**	0.58	0.57	0.67	0.58
	J	66	0.58	0.89	0.63	0.84	**0.93**	**0.88**
	K	165	0.60	0.76	0.64	0.65	0.60	0.62
	L	51	0.69	0.75	0.71	**0.86**	0.61	0.60
	M	127	0.56	0.81	0.57	0.59	0.65	0.62
	N(1)	44	0.57	0.81	0.59	0.61	0.78	0.63
	N(2)	42	0.63	0.68	0.60	0.62	0.60	0.60
	O	41	**0.77**	0.78	**0.77**	0.78	0.77	0.78
	Average		0.62	0.78	0.63	0.66	0.69	0.65