An E-Report Scoring Method based on Student Peer Evaluation using Groupware

Yuanyuan Wang
Kyoto Sangyo University
Motoyama, Kamigamo
Kita-ku, Kyoto
603-8555 Japan
yuanw@cc.kyoto-su.ac.jp

Setsuko Miyamoto
University of Hyogo
1-1-12 Shinzaike-honcho
Himeji, Hyogo
670-0092 Japan
miyamoto@shse.u-hyogo.ac.jp

Yukiko Kawai
Kyoto Sangyo University
Motoyama, Kamigamo
Kita-ku, Kyoto
603-8555 Japan
kawai@cc.kyoto-su.ac.jp

Kazutoshi Sumiya
University of Hyogo
1-1-12 Shinzaike-honcho
Himeji, Hyogo
670-0092 Japan
sumiya@shse.u-hyogo.ac.jp

L@S 2015, March 14–18, 2015, Vancouver, BC, Canada.
ACM 978-1-4503-3411-2/15/03.
http://dx.doi.org/10.1145/2724660.2728674

Abstract
Nowadays, many universities utilize groupware support for students to post and share their e-reports, and the students can browse and vote other students' reports in e-learning. Teachers then need to evaluate all students' reports, but this will require a great deal of time and effort for a fair evaluation of the reports. Therefore, we propose an e-report scoring method based on student peer evaluation by considering the relationship between voting and posting time of the e-reports, to promote the quality of the votes and prevent unfair votes. Then, the method can provide scores of reports based on a voting graph by analyzing students who vote the reports. In this paper, we perform a student peer evaluation using groupware based on voting with a "Like" button in a course practice, and compare the results with teachers' evaluation.

Author Keywords
e-reports; groupware; student peer evaluation; voting

ACM Classification Keywords
K.3.2 [Computers and Education]: Computer and Information.

Introduction
Web-based report systems, such as Bulletin Board Systems (BBSs) and groupware, are now one of the most

frequently used tools in e-learning currently. Students then post and share their reports at anytime and from anywhere in a given period, i.e., after a lecture and before the next lecture, and the students can browse and vote other students' reports through these online systems. However, teachers need to evaluate all students' reports, but this will require a great deal of time and effort for a fair and multi-faceted evaluation of the reports.

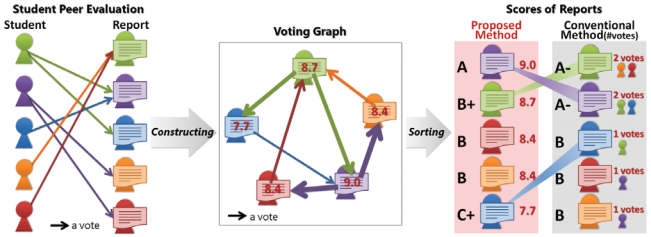

Figure 1: Conceptual diagram of our report scoring method.

As depicted in Figure 1, we propose a report scoring method to enable students instead of teachers to evaluate their reports by voting with each other. It provides scores of reports by analyzing the relationship between voting and posting time of the reports based on a voting graph of the reports, to promote the quality of the votes and prevent unfair votes. For this, the voting graph is constructed by the votes between a student and his or her voted reports. In this paper, we perform a student peer evaluation using groupware based on voting with a "Like" button in a course practice. Students can vote on the others' reports by pressing a "Like" button, when they think the report is good [3]. It can reduce the students' burden of evaluating others' reports without specific points. During a course using the groupware system, the following steps are followed: 1) posting reports after lectures in a certain period, 2) voting reports written by other students and voting with a "Like" button; and 3) receiving votes for their own reports.

The next section describes our proposed report scoring method based on student peer evaluation. In Section 3, we summarize the results of the proposed method in a course practice. Finally, in Section 4, we conclude this paper with suggestions for future work.

Report Scoring Method using Groupware

Voting Graph Construction

A voting graph is first constructed. The nodes of the directed graph consist of students' reports, and the links can be considered as votes from students for the others' reports. In our previous work, we built a system that evaluates users who browse the Web pages based on their links between a user and his or her browsing pages [5]. In this work, in order to evaluate the reports based on student peer evaluation; we focused on the students who vote on the others' reports. If one student u_i is voting another student u_j's report, then, a link from u_i's report to u_j's report (arrows in left and center parts of Figure 1), and the element of its corresponding adjacency matrix (u_j, u_i) is set to 1 (left matrix of Figure 2).

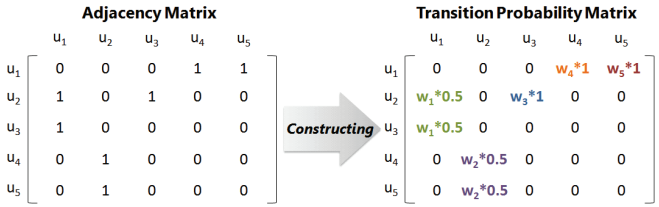

Figure 2: Adjacency matrix and transition probability matrix.

We next describe the construction of transition probability matrix from the adjacency matrix. For example, it can be

transformed to a transition probability matrix as shown in the right matrix of Figure 2. In general, before the students post their reports, they often refer to previous others' reports, then, we should reduce the scores of the last posting reports. Then, w_i is the weight assigned to each student's report by considering its posting time.

Student Peer Evaluation for their Reports
We next calculate the scores of students' reports by the following formula based on the concept of PageRank [2] and ObjectRank [1].

$$S(r) = (1-d) + d \times (S(v_1)/T(v_1) + \cdots + S(v_i)/T(v_i)) \quad (1)$$

- r: a student's report, i.e., the report of u_1, u_2, u_3, u_4, or u_5
- $v_1, \cdots, v_i$: the set of the votes of r, i.e., the votes of u_1's report by u_4 and u_5
- $T(v_i)$: the number of votes from a student, i.e., $T(v_1)$=2
- w_i: the weight of posting time of a report, i.e., w_1 is the weight of u_1's report. There are two methods:
  1. w_i=1/n_i, n_i is the posting number of u_i's report in the posted order
  2. w_i=(m-n_i+1)/n_i, m is the total number of students in a course
- d: a damping factor adjusts the derived value downward. Various studies have tested different damping factors, but it is generally assumed that the damping factor is approximately set as 0.85

Initially, the weight of each vote is 1, if a student votes multiple report, the weight distributes through each vote evenly, e.g., if u_1 votes reports of u_2 and u_3, the weight of each vote becomes 0.5. Finally, we normalize the score of each report between 0–10.0.

Results of Student Peer Evaluation

In this section, we present our findings from the results of our proposed report scoring method based on student peer evaluation in a course practice. This is a course of Applied Informatics, which consists of 10 lectures on different topics, and 23 students who participated in this course. Using the "Like" button as a vote through an online groupware, A) we calculated the score of each report based on the sum of the number of "Like" from students; B) we calculated the score of each report based on our proposed method by Equation 1 without the weight w_i considering the quality of votes only; C) we calculated the score of each report based on our proposed method by Equation 1 with the weight w_i(=1/n_i); and D) we calculated the score of each report based on our proposed method by Equation 1 with the weight w_i(=(m-n_i+1)/n_i) considering both the quality of votes and the posting time.

For each report scoring method, we then calculated the Spearman's rank correlation coefficient [4] between the score rankings by the teacher considering the posting time and the above methods of Lecture #3 in the same criteria. The correlation value ranges from -1 to 1, where -1 indicates that two rankings are completely reversed, whereas 1 indicates that the rankings are exactly the same. The correlation results are listed in Table 1, and the results can be explained as follows:

Table 1: Correlation results

Report Scoring Method	Correlation Value
A)	0.417
B)	0.465
C)	**0.488**
D)	**0.475**

- The correlation values of all methods (i.e., A), B), C), and D)) were not very close to 1.
- The correlation values of proposed methods (i.e., C) and D)) and the teacher's evaluation, become higher than those of baseline A) and our previous method B).

Although our proposed methods did not reach a very high correlation value, this experiment indicated that our proposed report scoring methods have the potential to support teachers easily and efficiently evaluate students' reports based on student peer evaluation using groupware in Japan. Since our proposed methods by considering both the quality of the votes and the posting time of the reports (strategy aspect by utilizing the cultural psychology of Japanese), achieved a good performance compared with the conventional report scoring methods by counting the total number of votes or considering the quality of votes only.

Future work will deeply analyze the correlation between our proposed methods based on student peer evaluation and teachers' evaluation with large datasets. In the teachers' evaluation, we need to adopt different methods by considering the posting time of reports or not. In order to verify the reliability of our proposed methods, we should do an interview of the students who participated in the student peer evaluation of their reports.

Conclusions

In this paper, we proposed a report scoring method based on student peer evaluation using groupware. In a course practice, students performed a peer evaluation for their reports by voting for valuable reports using a "Like" button. Therefore, it is not only a total number of votes for evaluating the reports, but also considering both the relationship between voting and posting time of the reports. It can lead to a new method rooted in indigenous culture of review by the student peer evaluation.

In the future, we need to measure inter-rater reliability of our proposed report scoring method based on student peer evaluation by combining teachers' evaluation and the student peer evaluation.

Acknowledgements

This work was supported by the Strategic Information and Communications R&D Promotion Programme (SCOPE) of the Ministry of Internal Affairs and Communications of Japan, and JSPS KAKENHI Grant Number 26280042.

References

[1] Balmin, A., Hristidis, V., and Papakonstantinou, Y. Objectrank: Authority-based keyword search in databases. In *Proc. VLDB 2004* (2004), 564–575.

[2] Brin, S., and Page, L. The anatomy of a large-scale hypertextual web search engine. *Computer Networks 30*, 1–7 (1998), 107–117.

[3] Iturrioz, J., Azpeitia, I., and Díaz, O. Generalizing the like button: empowering websites with monitoring capabilities. In *Proc. SAC 2014*, ACM (2014), 743–750.

[4] Myers, J. L., and Well, A. D. *Research Design and Statistical Analysis*. Lawrence Erlbaum Associates, 2003.

[5] Shiraishi, Y., Zhang, J., Kawai, Y., and Akiyama, T. Simultaneous realization of page-centric communication and search. In *Proc. CIKM 2012* (2012), 2719–2721.