Revealing Hidden Impression Topics in Students' Journals Based on Nonnegative Matrix Factorization

Yuta Taniguchi
Kyushu University, Japan
Email: yuta.taniguchi.y.t@gmail.com

Daiki Suehiro
Kyushu University, Japan
Email: suehiro93@gmail.com

Atsushi Shimada
Kyushu University, Japan
Email: atsushi@ait.kyushu-u.ac.jp

Hiroaki Ogata
Kyoto University, Japan
Email: hiroaki.ogata@gmail.com

***Abstract*—Students' reflective writings are useful not only for students themselves but also teachers. It is important for teachers to know which concepts were understood well by students and which concepts were not, to continuously improve their classes. However, it is difficult for teachers to thoroughly read the journals of more than one hundred students. In this paper, we propose a novel method to extract common topics and students' common impressions against them from students' journals. Weekly keywords are discovered from journals by scoring noun words with a measure based on TF-IDF term weighting scheme, and then we analyze co-occurrence relationships between extracted keywords and adjectives. We employs nonnegative matrix factorization, one of the topic modeling techniques, to discover the hidden impression topics from the co-occurrence relationships. As a case study, we applied our method on students' journals of the course "Information Science" held in our university. Our experimental results show that conceptual keywords are successfully extracted, and four significant impression topics are identified. We conclude that our analysis method can be used to collectively understand the impressions of students from journal texts.**

Keywords-component; journal writing; teaching analytics; nonnegative matrix factorization; impression

I. INTRODUCTION

Reflection is considered to play an important role in students' learning in higher education [1], and journal writing is commonly utilized for the purpose. In reflective writing, students describe what they learned, what they wondered, and any details regarding positive and negative points in the learning experience. Therefore, students' journal writings are often used as rich textual data in learning analytics [2].

Meanwhile, from teacher's perspective, students' journal writing is one of the handful ways which enables teachers to see their teaching objectively. Therefore, students' journals are considered full of resources leading to improvements not only in learning but also in teaching.

Reading journals intensively is time consuming, and it is not always possible for teachers to understand every journal of a class because of the large number of students. It would be helpful for teachers if we could automatically extract from journal texts to *what* topics students paid attention during class and *how* students felt about the topics. However, none of the previous researches addressed the problem to extract both *what* and *how*.

In this study, we focus on the problem to extract weekly keywords commonly discussed in journals (*what*) and students' impressions against them (*how*). Firstly, we extract weekly keywords from students' journals, most of which represent distinct concepts taught in a particular week. Secondly, related adjective words of the keywords are discovered as impressions based on co-occurrence relationships. Lastly, we make the obtained impressions into abstract impressions for easier understanding of the result. As a case study, we apply our method on the journals for the course "Information Science" held in Kyushu University last year, and discuss the result.

II. RELATED WORK

Many work proposed methods for detecting students' inner states with various approaches [3]–[7]. However, limited number of work focused on reflective writings. Chen et al. [8] proposed a method for automatic analysis and evaluation of reflective writings based on topic modeling technique. Their analysis shows only *what* topics students mentioned in journal writings. Nwanganga et al. [9] employed text mining technique to measure students' emotion from their reflections focusing on only *how* students felt. In contrast, we aim to discover both *what* and *how*, extracting co-occurrence relationships between weekly keywords and adjectives.

III. METHODS

A. Data

We analyze students' journals written for one of the courses of "Information Science" in Kyushu University, which was held for the first grade students during the first semester 2016. The course extensively covers fundamental topics of information science as shown in Table I. Students were instructed to write a journal entry per week after a class with the content including their impression after class, what they learned, what aspects they found interesting, and so on.

Our dataset consists of journal texts of about a hundred students in the class for 14 weeks, resulting in 1,664 entries. We stripped out pre-filled instruction texts, and run the Japanese morphological analyzer [10] on the resulting texts to tokenize sentences into words and infer their part-of-speeches.

B. Weekly Keywords Extraction

From journal texts, we extract *weekly keywords* which are commonly used by students in a particular week. Since different topics are discussed every week, weekly keywords are considered to be highly related to the topic of the week.

We propose a method based on the TF-IDF term-weighting measure [11] for the purpose. We consider a concatenated text of journal entries for a particular week as a *document*, and computer a weight for each noun in documents based on *raw term frequency* and *classical idf weight*. The proposed construction of a document makes our term-weighting measure take account of the week-specificity of words, and thus words with highest weights can be considered to be weekly keywords. All noun words are ranked according to their TF-IDF weights for each week, and a collection of top ten words from every week is used in the later analysis.

C. Co-occurrence Analysis

Adjectives are the most descriptive words when reading journals to know what impressions students have. We regard adjectives which are tightly associated to a noun as impressions for the noun. We employ normalized point-wise mutual information (NPMI) [12] to compute a degree of associations between two words, considering co-occurrences of words in a sliding window [13] spanning adjacent ten words.

NPMI is a normalized version of point-wise mutual information (PMI) [14], and it is defined as follows:

$$pmi(w_1, w_2) = \ln\left(\frac{p(w_1, w_2)}{p(w_1)p(w_2)}\right), \tag{1}$$

$$npmi(w_1, w_2) = pmi(w_1, w_2) / - \ln p(w_1, w_2), \tag{2}$$

where $p(w_1, w_2)$ is the joint probability that the co-occurrences of words w_1 and w_2 happens in a sliding window, and $p(w_i)$ is the probabilities that a word w_i occurs in a window. NPMI value ranges from -1 to 1; when two words are never observed together, $npmi(w_1, w_2) = -1$; when two words occur independently, $npmi(w_1, w_2) = 0$; when two words are always observed together, $npmi(w_1, w_2) = 1$.

D. Abstracting Impressions

We perform nonnegative matrix factorization (NMF) [15] on an NPMI matrix to obtain *abstract impressions*. For N weekly keywords and M adjectives, let $A = [a_{ij}]$ be a $N \times M$ real matrix, where $a_{ij} = pmi(w_i, w_j)$ and w_i is a weekly keyword and w_j is an adjective. We use NMF to approximates such a matrix by a multiplication $A \approx WH$, where W and H are low-rank matrices of $N \times K$ and $K \times M$, respectively. Giving a smaller value than N and M as the parameter K, we can obtain K abstract impressions. Matrices W and H describe the associations between weekly keywords and abstract impressions, and between abstract impressions and adjectives, respectively. Performing stability analysis [16], we chosen $K = 4$ which gives the highest stability among possible parameter values between 2 and 20.

IV. Results & Discussion

A. Weekly Keywords Extraction

Table I shows actual topics and extracted weekly keywords for each week, where only top five words given highest TF-IDF scores are shown in order. Words were translated from Japanese to English by the first author preserving the original meanings as much as possible.

Comparing the keywords with actual topics, we can see many keywords that are specific to course contents were successfully extracted. For example, in the first week, because the topic was the introduction of the field, relatively general words are obtained. In contrast to that, in the sixth week, we can see many terms related to cryptography are extracted.

B. Co-occurrence Analysis

Table II shows the associated words and NPMI scores for the adjective “difficult” and the weekly keyword “cipher”. We can see what weekly topics are frequently mentioned in students’ journals when students were writing about something *difficult* for them. For example, *Irfanview* is the image viewing software which we use in the course. We can guess that installation and/or use of the software were hard for them.

The table also shows what impression students had when they were writing about “cipher.” We can say that students had *troublesome* feeling, but we cannot say they think it is *difficult* or not since “cipher” and “difficult” occurs almost independently. Rather, it seems like that they had *fun* and *amazing* feelings more strongly with “cipher” in class.

C. Abstracting Impressions

Table III shows significant adjectives from the matrix H for every abstract impression. The association between weekly keywords and abstract impressions from the matrix W are not show due to limitations of space. In the table, rows contains some words multiple times because of translation.

The abstract impression 1 can be interpreted as the impression relevant to understanding concepts, and they were strongly associated with cryptography, algorithms, and information theory. This suggests that these concepts are controversial for students, and teachers may need to care about students’ understanding of these concepts during teaching. For the rest of abstract impressions, we can say the second one is for computation, algorithm, and time complexity; the third one is for information theory; and the forth one is for midterm- and final exams.

V. Conclusions

We proposed a novel method for analyzing students’ journal writings, which could lead to improvements of teaching. Our method successfully extracted major topics discussed for each week. Furthermore, co-occurrence analysis revealed students’ hidden abstract impressions against them. Those obtained insights are considered to be helpful for teachers to improve teaching in future lectures. Future work may explore also journals from other classes of the same course together, which will enable us to compare different teaching styles.

TABLE I.
ACTUAL TOPICS AND TOP TEN WEEKLY KEYWORDS FOR EVERY WEEK.

Week	Actual Topics	Extracted Weekly Keywords
1st	Positioning the lecture	Internet, orientation, university, search, binary, remaining, system, method, life, study
2nd	Introduction to information science	Morse, signal, topic, quiz, homework, Japanese, thought, English, remaining, like
3rd	Information quantity and entropy	code, entropy, prefix, encoding, average, word, combine, length, symbol, multiple
4th	Entropy	entropy, mutual, expectation, value, quantity, log, computation, with, information, condition
5th	Channel coding	correction, error, automated, detection, Hamming, distance, communication, encoding, doodlebug's pit, example
6th	Cryptography, computer science	cipher, encryption, key, Caesar, mod, public, RSA, secret, decryption, high school
7th	Computation, algorithm, time complexity	coin, Euclid, mutual division, balance scale, fake, algorithm, method, mathematics, GCD, rectangle
8th	Midterm exam	exam, midterm, one question, perfect score, final, right answer, question, miss, two questions, effect
9th	Stack and queue, bubble sort	notation, Polish, queue, stack, infix, sort, order, bubble, principle, first
10th	Heap sort, merge sort	sort, heap sort, merge, tree, binary, comparison, algorithm, binary tree, drawback, practice
11th	Bucket sort, binary search	sort, bucket, search, binary, search, binary, dictionary, Google, comparison, heap sort
12th	Digital images	image, app, install, usage, download, exercise, Irfanview, rose, file, organism
13th	Image processing, character recognition	image, recognition, processing, next week, letter, unistroke, edge, final, report, single stroke
14th	Final exam	exam, final, small, perfect score, report, first semester, two questions, three questions, plan, final

TABLE II.
EXAMPLE LISTS OF ASSOCIATED WORDS AND NPMI VALUES FOR "DIFFICULT" AND "CIPHER".

Target	Associated Weekly Keywords / Adjectives
difficult	Irfanview (0.379), exercise (0.271), Japanese (0.227), infix (0.217), notation (0.199), Caesar (0.193), merge (0.117), comparison (0.117), encoding (0.117), code (0.100)
cipher	troublesome (0.247), detailed (0.208), fun (0.118), good (0.117), amazing (0.111), easy (0.088), interesting (0.068), difficult (0.022)

TABLE III.
FOUR ABSTRACT IMPRESSIONS OBTAINED BY NONNEGATIVE FACTORIZATION OF THE NPMI MATRIX.

ID	Descriptive Adjectives
1	difficult, easy, fun, troublesome, interesting, difficult, interesting, new, amazing, few
2	heavy, light, easy, interesting, few, involved, interesting, long, difficult, difficult
3	many, difficult, smart, big, well, difficult, interesting, rare, short, right
4	good, regretful, amazing, interesting, detailed, wonderful, happy, casual, terrible, near

ACKNOWLEDGMENT

The research is supported by "Research and Development on Fundamental and Utilization Technologies for Social Big Data" (178A03), the Commissioned Research of the National Institute of Information and Communications Technology (NICT), Japan; Grant-in-Aid for Scientific Research (S) No. 16H06304; and the Education Enhancement Program (EEP) of Kyushu University.

REFERENCES

[1] N. Hatton and D. Smith, "Reflection in teacher education: Towards definition and implementation," *Teaching and Teacher Education*, vol. 11, no. 1, pp. 33 – 49, 1995.

[2] A. Gibson and K. Kitto, "Analysing reflective text for learning analytics: an approach using anomaly recontextualisation," in *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge*, 2015, pp. 275–279.

[3] B. Grawemeyer, M. Mavrikis, W. Holmes, S. Gutierrez-Santos, M. Wiedmann, and N. Rummel, "Affecting off-task behaviour: how affect-aware feedback can improve student learning," in *Proceedings of the Sixth International Conference on Learning Analytics and Knowledge*, 2016, pp. 104–113.

[4] Y. Wang, N. T. Heffernan, and C. Heffernan, "Towards better affect detectors: effect of missing skills, class features and common wrong answers," in *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge*, 2015, pp. 31–35.

[5] A. K. Vail, J. B. Wiggins, J. F. Grafsgaard, K. E. Boyer, E. N. Wiebe, and J. C. Lester, "The affective impact of tutor questions: Predicting frustration and engagement," in *Proceedings of the Ninth International Conference on Educational Data Mining*, 2016, pp. 247–254.

[6] S. Ruiz, S. Charleer, M. Urretavizcaya, J. Klerkx, I. Fernández-Castro, and E. Duval, "Supporting learning by considering emotions: tracking and visualization a case study," in *Proceedings of the Sixth International Conference on Learning Analytics and Knowledge*, 2016, pp. 254–263.

[7] M. Wen, D. Yang, and C. Rose, "Sentiment analysis in mooc discussion forums: What does it tell us?" in *Proceedings of the Seventh International Conference on Educational Data Mining*, 2014, pp. 130–137.

[8] Y. Chen, B. Yu, X. Zhang, and Y. Yu, "Topic modeling for evaluating students' reflective writing: A case study of pre-service teachers' journals," in *Proceedings of the Sixth International Conference on Learning Analytics and Knowledge*. New York, NY, USA: ACM, 2016, pp. 1–5.

[9] F. Nwanganga, E. Aguiar, G. A. Ambrose, V. Goodrich, and N. V. Chawla, "Qualitatively exploring electronic portfolios: a text mining approach to measuring student emotion as an early warning indicator," in *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge*, 2015, pp. 422–423.

[10] T. Kudo, K. Yamamoto, and Y. Matsumoto, "Applying conditional random fields to japanese morphological analysis," in *Proceedings of EMNLP 2004*, 2004, pp. 230–237.

[11] G. Salton and C. Buckley, "Term-weighting approaches in automatic text retrieval," *Information Processing & Management*, vol. 24, no. 5, pp. 513 – 523, 1988.

[12] G. Bouma, "Normalized (pointwise) mutual information in collocation extraction," *Proceedings of GSCL*, pp. 31–40, 2009.

[13] K. Lund and C. Burgess, "Producing high-dimensional semantic spaces from lexical co-occurrence," *Behavior Research Methods, Instruments, & Computers*, vol. 28, no. 2, pp. 203–208, 1996.

[14] K. W. Church and P. Hanks, "Word association norms, mutual information, and lexicography," *Comput. Linguist.*, vol. 16, no. 1, pp. 22–29, Mar. 1990.

[15] S. Sra and I. S. Dhillon, "Generalized nonnegative matrix approximations with bregman divergences," in *Advances in Neural Information Processing Systems 18*. MIT Press, 2006, pp. 283–290.

[16] D. Greene, D. O'Callaghan, and P. Cunningham, "How many topics? stability analysis for topic models," in *Proceedings of the European Conference on Machine Learning and Knowledge Discovery in Databases - Volume 8724*, 2014, pp. 498–513.