A Technique for Tracking the Reading Rate to Provide Students' Learning Feedback

Fuzheng ZHAO[a] & Chengjiu YIN[b*]
[a]*Graduate School of System Informatics, Kobe University, Japan*
[b]*Information Science and Technology Center, Kobe University, Japan*
*yin@lion.kobe-u.ac.jp

Abstract: Reading rate is the number of words a user reads per unit of time, also known as reading speed. As online learning has been continuously growing, research on tracking learning behavior is undertaken, and the reading rate was beginning to be applied to measure the learning process. However, current calculating methods are based on the traditional learning environment of a paper-based course, and further research is needed to determine whether they can be used to track students' learning in an online learning environment and integrate with their online reading habits. To this end, this study proposed a new method based on online learning used to calculate reading rate as a measure to track students' learning status, after analyzing students' reading behavior in the e-book learning system. This method not only can identify students with learning difficulties through reading rate outliers, but also provide urgent feedback to students on their reading status.

Keywords: Reading rate, learning tracking, urgent feedback

1. Introduction

Accurately and timely understand students' learning status, discover abnormal learning habits and intervene, which puts forward higher requirements for evaluating the learning process. Many studies have proved that reading comprehension is correlated with learning outcomes (Crawford, Tindal, & Stieber, 2001), and the reading rate is one of the important indicators to measure reading comprehension. However, frequently used tests like classroom oral reading rate in the classroom (Deno & Mirkin, 1977) cannot detect students' learning status in real-time (Skinner et al., 2009).

With the continuous application of online learning environments and learning methods, the popularity of digital teaching materials has further accelerated the development of online reading (Zhao, Hwang, & Yin, 2021). Besides, studies have shown that students who use online reading achieve better reading comprehension than paper reading (Huang, 2014). Especially, online reading is capable to provide a reliable reading tool and learning environment that gives timely feedback. But, Vonderwell (2003) reported that students' greatest disappointment with online learning was the lack of communication and a sense of community, and Rovai (2002) proved that when students encounter difficulties, this sense of community is rewarded if teachers detect and help them.

To this end, this study will use the characteristics of the online learning environment to design a reading rate calculation method, which is used to detect students' reading rate behaviors in real-time, help teachers understand students' reading status, and find abnormal reading behaviors.

2. Literature Review

2.1 Tracking Learning by Reading Rate

Reading rate refers to the number of words read per unit time, also known as reading speed (Carver, 1992), which has been used as an independent variable to determine students' learning skills. While, various reading rate is correlated with reading errors and knowledge comprehension (Breznitz & Berman, 2003). The conventional view is that humanity's reading speed varies with the difficulty of the material and the purpose of the reader (Brysbaert, 2019). Besides, reading rate is often quantified by

reading speed, which is the number of words read correctly per minute or the length of time it takes a reader to complete an article (Hudson, Lane, & Pullen, 2005). Notably, the basic measure of reading rate is done by reading at regular intervals. Specifically, during the test, students have to complete the reading materials arranged by the teacher, to judge the number of words they read correctly and the number of errors in a certain period.

2.2 Relationship between Reading Rate and Learning Comprehension

The main function of reading is comprehension. Carver (1992) considered the reading rate to reflect the ability of word recognition and reading comprehension. Within this context, the reading rate is used as a diagnostic tool for dyslexia. Since slow reading leads to diminished comprehension, many interventions focus on improving reading speed (Bernacki, Byrnes, & Cromley, 2012). On the other hand, it is important to keep abreast of students' reading rates. Although a measure of reading speed is highly correlated with reading comprehension, it does not contain a direct measure of comprehension. Crawford, Tindal, and Stieber (2001) used reading rate to predict student performance on a national test consisting of reading and mathematics, and through correlation analysis, only validated a good way to monitor student progress by measuring reading rate.

2.3 Reading Rate Calculation

Current reading rate calculation methods include oral and non-verbal reading rates. The oral reading rate is divided into correct word rate, question-answer correct rate, and cloze correct rate, depending on the calculation method. Specifically, correct word rate is measured as the number of words read correctly per minute (Skinner et al., 2009). In the measurement, students are asked to read aloud a given passage within 1 minute, and teachers rate the accuracy of reading each word. Thus, the correct word rate is a speed measure that includes accurate speaking and reading speed (Good & Kaminski, 2002). In addition, to measure comprehension accuracy during reading, students answer pre-designed comprehension questions after reading a text passage aloud and measure the percentage of text per minute reading comprehension, which is referred to as the question response rate (Daly, Chafouleas, & Skinner, 2005). It is calculated by dividing the correct question rate by the number of seconds it takes to read the text and then multiplying by 60 seconds. Finally, the correct completion question rate is measured by asking students to select the correct answer from the options provided after reading aloud (Ridge & Skinner, 2011). Although the completion test is still widely recommended in instructional methods texts, it is also dependent on writing skills and can be time-consuming and frustrating. By contrast, a measure of non-verbal reading rate is a measure of reading speed, i.e. the number of words a student reads in one minute. This method has been used as a technique to track student reading and indirectly assess student reading comprehension (Liang & Huang, 2014).

3. Reading Rate Calculation Based on Online Learning Behavior

3.1 Online Reading Behavior

Currently, online reading is done using a learning management system, such as Moodle, using a dedicated e-book reading system, or a hardware device such as a reading kinder. Unlike paper-based reading, learning strategies or styles vary due to changes in a learning environment and reading tools, such as reading back during reading (Yin et al., 2019), skipping during the reading review (Yin et al., 2015), and repetitive manipulation of reading tools (Zhao, Hwang, & Yin, 2021). At the same time, these behaviors in the online learning environment are related to academic performance, learning style, and learning pressure. For this reason, this study will utilize reading behaviors in the online reading environment, such as page-turning, view manipulation, highlighting, underlining, and making memos, as one of the main bases for reading calculation methods. This choice mainly takes into account the e-book reading system, which can provide a wealth of reading tools such as turning pages, changing views, and annotating. In addition, the e-book reading system can record students' reading behavior in

the form of a log for easy data analysis. From the convenience of data collection, this study uses the e-book system developed by Yin et al.(2019) to collect data on students' reading behavior.

3.2 Calculation Method

The main idea of the reading rate calculation method is to select the features that best represent the occurrence of reading behavior, after which the core elements of reading are extracted from these reading behavior features, the core elements include time, word count, and page count. The specific reading rate calculation method is shown in Figure 1. The reading rate calculation method proposed in this study focuses on calculating the speed of reading material and reading content: calculating the speed of reading material means calculating the reading speed of students when using the page-turning tool; calculating the speed of reading content is obtained by calculating the reading speed of students when using underlined/highlighted annotated text.

The formulae were categorized based on the three learning strategies: read all the words in the material (RAW), read some of the words in the material (RSW), and did not read any words (RNW). These three learning methods are judged by comparing the number of words read by the student with the total number of words per page in the reading material. If the difference is equal to zero, then the student has read all of the words on that page; if it is less than zero, then the student has read some of the words on that page; if the difference is equal to the total number of words on no page of the reading material, then the student has not read any words.

RAW calculation. When the student has read all the words, the reading rate is equal to the total number of words on each page of the reading material divided by the time spent.

RNW calculation. When no text was read, the reading rate was zero.

RSW calculation. When the student reads some of the text, calculating the reading rate further involves determining whether the student is using underlining/highlighting when reading. In the case of underlined/highlighted reading, the reading rate for a single reading operation is first calculated as the total number of words marked in a single reading divided by the time spent. The average reading rate is then calculated as the sum of the single reading rates divided by the total number of reading operations. In the case of no underlined/highlighted reading, the reading rate is equal to the total number of untagged words divided by the total time spent reading minus the difference in time spent reading without underlining/highlighting. Finally, the size of the reading rate in these two cases is determined. If either reading rate is greater than 50%, it is considered to be the dominant reading rate.
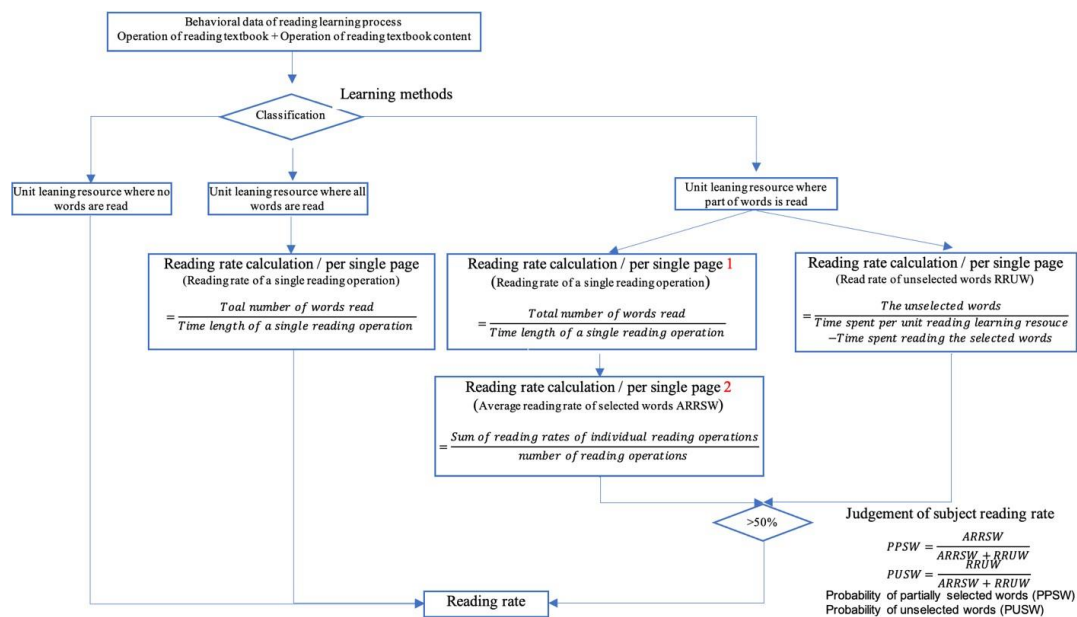

Figure 1. Reading Rate Calculation Flowchart

3.3 Data Collection and Processing

The reading rate calculation method proposed by this study requires the students' operation behavior number, operation behavior time, number of words read, and number of pages of reading materials. First, the data collection is done by the e-book system, and the data is stored in the database in the form of log data, as shown in Table 1. "ID" is the student's number, and "OperationName" represents the student's operation behavior: including specific operation behaviors, such as adding underscores; "id: 23" indicates the number of the operation behavior; "Makertext" indicates the number of words involved in the operation behavior, such as students using underscores to read a total of two words "more attention". "Operation data" represents the time when the student performed the operation behavior. "E-book" is a reading material. "Device" is the device used by the student. "Pages" indicate the total number of pages of the reading material.

Table 1. *Log Data of Reading Behavior Collected from E-book System*

ID	Operation name	Operation date	E-book	Device	Pages
1	ADD Highlight id:23 page:5 markertext: more attention	2021-05-19 11:00:48.329	23	pc	49
1	ADD Underline id:24 page:5 markertext: address some problems	2021-05-19 11:00:55.061	23	pc	49
1	ADD Memo id:46 page:5 markertext: make reflections	2021-05-19 11:01:44.326	23	pc	49

Data processing and analysis use Python, including Xlrd, Xlwt, Openpyxl, and Pandas modules. First, use the Pandas module to remove invalid and missing data. Second, use Xlrd, Xlwt, and Openpyxl to split the data of "OperationName", the split data types include the specific name of the reading operation and its number, the page number where the reading operation occurred, and the number of words read, and the total number of words. Finally, the Datetime module calculates the duration of the read operation behavior. The final results are shown in Table 2. Among them, "BehaviorN" is the name of the reading operation, and "PageN" is the page number where the reading operation is located. "MarkingT" is the number of words read, and "TextS" is the sum of the number of words read. "Time" is the start time of the read operation. "Timedelta" is the duration of the read operation. "WordN" is the sum of all the word counts of the page number on which the read operation occurs.

Table 2. *Processed Behavioral Data for Calculating Reading Rate*

ID	BehaviorN	PageN	MarkingT	TextS	Time	Timedelta	WordN
1	AddHL	1	Digital Learning	2	11:13:19.565	0:04:53.050	72
1	DeleteHL	1	Num	0	11:13:33.919	0:00:00.994	0
1	AddUN	1	Education	1	11:13:45.628	0:00:33.725	514
1	Next	2	Num	0	11:25:58.384	0:00:12.801	104
1	AddHL	2	It was found …	20	11:30:52.492	0:00:06.665	233

3.4 Reading Rate Indicator

Based on the relationship between reading rate and learning behavior, Huang et al. (2014) summarized the reading rate and status related to reading behavior, as shown in Table 3. Depending on the reading rate, there are seven reading states: "Slowing", "Memorizing", "Learning", "Raiding", "Skimming", "Scanning", and "Flipping". Each reading state has a corresponding reading behavior. For example, when the reading rate is less than 50 words per minute, it can be considered that the student is in a "slowing" reading state, and the student's learning behavior at this time is mainly one of the following or a concentrated combination. For example, "Excessively slow", "Inefficient reading", "Disfluent", "Labored", "Inexpressive", and "Unenthusiastic rendering".

Table 3. *The Relationship between Reading Rate, Status, and Reading Behavior*

Reading status	Reading rate	Reading behavior
Slowing	1-50	Excessively slow, Inefficient reading, Disfluent, Labored, Inexpressive, Unenthusiastic rendering
Memorizing	50-100	Sustained attention, In-depth reading
Learning	100-200	Oral reading, Concentrated reading annotation
Raiding	200-400	Silent reading
Skimming	400-700	Keyword spotting, One-time reading
Scanning	700-1000	Reading selectively, Browsing and scanning, Non-linear reading
Flipping	>1000	Flip page, Glance and glimpse

4. Reading Rate and Learning Status Tracking

The students' behavior prediction is mainly based on the data collected by the e-book system and the reading rate calculation method developed in this study. Based on the existing e-book system, the Spring MVC framework and Java were used to realize the predictive analysis based on the reading rate, as shown in Figure 2. Teachers can click the reading status submenu in the main menu bar, enter the course name after entering the page, and click search to obtain the seven reading rate status distributions of the students in the course. Each learning status is sorted in descending order of reading rate.

Figure 2. Reading Rate Feedback

5. Discussion and Conclusion

Finding the best way to track students' learning status has been getting a lot of attention. As an important evaluation method, the measurement method and evaluation index of reading rate is constantly developing. With the rapid development of online education, the learning environment and methods have also changed, shifting to mobility and fragmentation. Online reading has been described as an important learning method. In such an environment, there are two problems.

Firstly, whether the traditional methods of calculating correct reading words and reading speed can be applied to online learning platforms. This method has not been well applied to online education. On the one hand, this traditional method reveals many problems, such as being time-consuming and not allowing for real-time tracking. On the other hand, learning behavior and reading habits are also different a way of learning that differs from reading on paper in traditional courses. A second problem arises: how to make effective use of behavioral models of online learning while overcoming the drawbacks of traditional methods, has become an urgent issue to be addressed. After analyzing the strengths and weaknesses of existing methods for calculating reading rates, this study starts with two aspects of reading material manipulation and reading content: firstly, the reading speed method, which

is one of the traditional methods, is used in calculating reading material manipulation, which is the number of words read divided by the number of words read. Calculated by reading time.

Secondly, in terms of calculating reading content, the number of words read as annotations was divided by the number of words read as annotations by calculating students' reading behavior with annotations, i.e. calculating students' reading speed during the underlining and highlighting reading operations. time spent on reading. In this way, a method for calculating reading rates based on online reading was obtained for this study. It is worth noting that by introducing a correlation indicator between reading rate and learning behavior, the reading rate can be translated into seven reading behaviors of students that teachers can easily understand. This can help teachers to identify students' reading status on time, track their learning status, and identify which students are struggling with reading.

Acknowledgments

References

Bernacki, M. L., Byrnes, J. P., & Cromley, J. G. (2012). The effects of achievement goals and self-regulated learning behaviors on reading comprehension in technology-enhanced learning environments. *Contemporary Educational Psychology*, *37*(2), 148-161.

Breznitz, Z., & Berman, L. (2003). The underlying factors of word reading rate. *Educational Psychology Review*, *15*(3), 247-265.

Brysbaert, M. (2019). How many words do we read per minute? A review and meta-analysis of reading rate. *Journal of memory and language*, *109*, 104047.

Carver, R. P. (1992). Reading rate: Theory, research, and practical implications. *Journal of Reading*, *36*(2), 84-95.

Crawford, L., Tindal, G., & Stieber, S. (2001). Using oral reading rate to predict student performance on statewide achievement tests. *Educational Assessment*, *7*(4), 303-323.

Daly, E. J., Chafouleas, S., & Skinner, C. H. (2005). Designing and evaluating measurably effective interventions for reading problems. *New York: Guilford Publications*.

Deno, S. L., & Mirkin, P. K. (1977). Data-based program modification: A manual. Reston, VA: Council for Exceptional Children

Good, R. H., & Kaminski, R. A. (Eds.). (2002). Dynamic Indicators of Basic Early Literacy Skills (6th Ed.). Eugene, OR: Institute for the Development of Educational Achievement.

Huang, H. C. (2014). Online versus paper-based instruction: Comparing two strategy training modules for improving reading comprehension. *RELC journal*, *45*(2), 165-180.

Hudson, R. F., Lane, H. B., & Pullen, P. C. (2005). Reading fluency assessment and instruction: What, why, and how?. *The Reading Teacher*, *58*(8), 702-714.

Liang, T. H., & Huang, Y. M. (2014). An investigation of reading rate patterns and retrieval outcomes of elementary school students with e-books. *Journal of Educational Technology & Society*, *17*(1), 218-230.

Ridge, A. D., & Skinner, C. H. (2011). Using the TELLS prereading procedure to enhance comprehension levels and rates in secondary students. *Psychology in the Schools*, *48*(1), 46-58.

Rovai, A. P. (2002). Sense of community, perceived cognitive learning, and persistence in asynchronous learning networks. *The Internet and Higher Education*, *5*(4), 319-332.

Skinner, C. H., Williams, J. L., Morrow, J. A., Hale, A. D., Neddenriep, C. E., & Hawkins, R. O. (2009). The validity of reading comprehension rate: Reading speed, comprehension, and comprehension rates. *Psychology in the Schools*, *46*(10), 1036-1047.

Vonderwell, S. (2003). An examination of asynchronous communication experiences and perspectives of students in an online course: A case study. *The Internet and higher education*, *6*(1), 77-90.

Yin, C., Okubo, F., Shimada, A., Oi, M., Hirokawa, S., & Ogata, H. (2015, November). Identifying and analyzing the learning behaviors of students using e-books. In *Proceedings of the 23rd international conference on computers in education* (pp. 118-120). Hangzhou, China: Asia-Pacific Society for Computers in Education.

Yin, C., Ren, Z., Polyzou, A., & Wang, Y. (2019, July). Learning Behavioral Pattern Analysis Based on Digital Textbook Reading Logs. In International Conference on Human-Computer Interaction (pp. 471-480). Springer, Cham.

Zhao, F., Hwang, G. J., & Yin, C. (2021). A Result Confirmation-based Learning Behavior Analysis Framework for Exploring the Hidden Reasons behind Patterns and Strategies. *Educational Technology & Society*, *24*(1), 138-151.