Optimizing Pattern Weights with a Genetic Algorithm to Improve Automatic Working Memory Capacity Identification

Jason Bernard[1(✉)], Ting-Wen Chang[2], Elvira Popescu[3], and Sabine Graf[1]

[1] Athabasca University, Athabasca, Canada
c.j.bernard@ieee.org, sabineg@athabascau.ca
[2] Beijing Normal University, Beijing, China
tingwenchang@bnu.edu.cn
[3] University of Craiova, Craiova, Romania
popescu_elvira@software.ucv.ro

Abstract. Cognitive load theory states that improper cognitive loads may negatively affect learning. By identifying students' working memory capacity (WMC), personalized scaffolding techniques can be used, either by teachers or adaptive systems to offer students individual recommendations of learning activities based on their individual cognitive load. WMC has been identified traditionally by dedicated tests. However, these tests have certain drawbacks (e.g., students have to spend additional time on them, etc.). Therefore, recent research aims at automatically detecting WMC from students' behavior in learning systems. This paper introduces an automatic approach to identify WMC in learning systems using a genetic algorithm. An evaluation of this approach using data from 63 students shows it outperforms the existing leading approach with an accuracy of 85.1 %. By increasing the accuracy of automatic WMC identification, more accurate interventions can be made to better support students and ensure that their working memory is balanced properly while learning.

Keywords: Working Memory Capacity · Student modeling · Genetic algorithm

1 Introduction

Working memory capacity (WMC) is a cognitive trait that influences the learning process, in terms of learning speed, memorization of learned concepts and effectiveness of skill acquisition [1]. WMC enables us to keep active a limited amount of information (7 ± 2 items) for a brief period of time [2]. Exceeding the WMC limit can reduce students' learning performance, reduce transfer of learning or increase the amount of time needed to learn [3, 4]. By identifying WMC, cognitive load can be individualized to the student which benefits the learning process. For example, an adaptive recommendation system could provide personalized suggestions for learning activities to students [5]. Furthermore, simple awareness of WMC supports students in making

A. Micarelli et al. (Eds.): ITS 2016, LNCS 9684, pp. 334–340, 2016.
DOI: 10.1007/978-3-319-39583-8_38

better choices for self-regulated learning and teachers may factor in WMC when making interventions for their students.

Traditionally, WMC is measured by asking students to take a specific multitasking test such as operation span task (OSPAN) [6]. OSPAN is considered a stable and reliable test [7] and several online versions of this test have been created, such as WebOSPAN [8]. Although such tests are effective, they have the notable drawbacks of requiring additional time and effort from learners to do the test and the risk of inaccuracies due to factors such as the perceived importance of the test by the students, stress or fatigue [9].

To overcome these drawbacks, automatic approaches have been investigated which analyze students' behavior to identify WMC automatically while students are learning in a learning system. As a basis for such automatic approaches, several studies investigated and found relationships between WMC and other student characteristics as well as their relation to student behavior [e.g., 1, 10, 11]. To the best of our knowledge, only one automatic approach for identifying WMC is proposed so far. DeWMC (Detecting Working Memory Capacity) [12, 13] calculates WMC using six patterns which all contribute equally to the identification of students' WMC.

This paper presents a tool for automatic WMC identification called WMCID-GA. WMCID-GA is based on DeWMC [12] and extends it through the use of a genetic algorithm which optimizes the weights of patterns impacting the WMC calculation in order to improve the precision of identifying WMC.

The remainder of this paper is structured as follows. Section 2 introduces the proposed WMC identification approach. Section 3 describes the evaluation of WMCID-GA and Sect. 4 concludes the paper.

2 WMCID-GA

In this section, we start with introducing DeWMC, followed by presenting WMCID-GA (WMC IDentifier-Genetic Algorithm) and how its genetic algorithm was built.

DeWMC [12] uses six patterns (five behavior patterns and one pattern related to learning styles based on the Felder-Silverman learning style model [14]) to calculate WMC. The five behavior patterns consider behaviors including linear navigation, constant reverse navigation, performing simultaneous tasks, recalling learned material, and revisiting passed learning objects. Each of the six patterns has been selected based on detailed investigations and evidence from literature that there exists a relation between the respective pattern and WMC [12]. To calculate WMC, DeWMC first extracts student data from a learning system's database and computes the respective patterns considering student behavior and their learning styles. For each pattern, a high or low value is associated to high or low WMC, based on existing studies from literature [12]. Then, for each learning session of a student, a WMC session value is calculated building the average of all pattern values. Subsequently, the overall WMC value is calculated by building a weighted average over all WMC session values, considering the amount of available behavior data per learning session as a weight.

WMCID-GA is based on DeWMC. It uses the same patterns and a similar concept to calculate WMC from these patterns, with the only difference that WMCID-GA is using a weight for each pattern when building the WMC session value (instead of assuming that all patterns contribute equally to the WMC session value). To find the optimal weight for each pattern, WMCID-GA uses a genetic algorithm (GA) [15, 16] which is an optimization algorithm that utilizes concepts from evolutionary biology to solve optimization problems.

A GA represents solutions as genomes, where each genome consists of a set of numbers representing genes. To find the optimal weights of patterns, each genome consists of six genes (each for one pattern) where each gene has a range of values (representing the weight of the respective pattern) from 0.01 to 1.0 in increments of 0.01. A value of 0 is excluded since according to literature [12], each of these patterns has at least a small contribution to the WMC identification. To calculate the fitness/quality of a genome, the error between the actual WMC and the calculated WMC for each student in a given dataset is calculated and the average error over all students is used as fitness value. The calculated WMC is computed from the six patterns, as described above, using the genome's gene values as pattern weights. The GA starts by initializing the population (P) with random values for each genome as no information is available on the potential quality for any weight value. In each generation, $P/2$ genome pairs are selected for crossover using the roulette wheel technique and uniform crossover is used where each gene has a chance of being swapped equal to the crossover weight (C). Then, uniform mutation is used on each new offspring where each gene has a chance of being mutated equal to the mutation weight (M). After crossover and mutation, the new genomes are merged into the population and the genomes with the lowest fitness are culled until the population is size P again. Once the new population is built, a new generation starts. To promote finding the optimal solution, the generation number of the best solution (G_{best}) is recorded and the GA stops only after another G_{best} generations passed without finding a new best solution. To prevent early termination, a minimum of 10,000 generations must pass before WMCID-GA can terminate.

3 Evaluation

In this section, the evaluation of WMCID-GA is reported, starting with presenting the dataset and describing the evaluation design and performance metrics. Subsequently, the optimization of parameters and overfitting reduction strategies are explained, followed by a discussion of the results.

To evaluate WMCID-GA, data from 63 undergraduate students on the five behavior patterns and the learning style pattern (identified by the Index of Learning Styles questionnaire [17]), and WMC (identified by WebOSPAN [8]) was used.

The evaluation consists of three parts. First, to find the optimal values for the parameters of the GA, an iterative experimental process was used. Second, an experimental process was also used to test overfitting reduction strategies and find optimal parameters for those strategies. Third, the optimal GA parameters and the optimal overfitting reduction strategies were then used to run WMCID-GA and get final results. In order to ensure generalizability to any datasets, 10 fold cross validation was used for each part of the evaluation.

To evaluate the performance of WMCID-GA in each part of the evaluation, three metrics were used: ACC measures the difference between a student's actual WMC and the WMC identified by WMCID-GA. An ACC value is computed for each student and an average ACC is built, which provides details on the overall accuracy of WMCID-GA. LACC is the lowest ACC value in the assessment set and measures the worst case scenario for an individual student. %Match measures the percentage of students who were identified with reasonable accuracy. A threshold for reasonable accuracy of ACC > 0.7 was calculated by considering the range of WMC values in the dataset and assuming that ACC has to be at least higher than half of this range.

In the first part of the evaluation, the GA parameters are optimized in the following order: population size (*P*), crossover weight (*C*) and mutation weight (*M*). For each parameter, suitable parameter ranges or principles were investigated based on existing literature [15, 16], resulting in a set of possible values for each parameter. For the first parameter, WMCID-GA was executed iteratively for each value in the set while using a mid-range value for the remaining parameters. The parameter value which produces the best result is considered the optimal choice and used for all subsequent executions. This process is repeated for each parameter with the resulting optimal parameter settings shown in Table 1.

Table 1. Optimal parameter settings

Population	Crossover weight	Mutation weight
25	0.80	0.001

Table 2. Optimal overfitting reduction settings

Stratification	FEP	min_{gen}
On	On	25

With genetic algorithms, overfitting is a potential problem. This problem was addressed in the second part of the evaluation where the benefit of using two overfitting reduction techniques, stratification [18] and future error prediction (FEP) [19], was assessed through experimentation. For FEP, the optimal setting of an early termination parameter (min_{gen}) was also investigated. Table 2 shows the optimal overfitting reduction settings.

In the third part of the evaluation, the optimal parameter and overfitting reduction settings were used to obtain a final result. The results for the three performance metrics are shown in Table 3, together with the respective results from DeWMC.

Comparing the results shows that WMCID-GA has outperformed DeWMC in every metric; thereby, showing that optimizing the pattern weights improves the overall accuracy of WMC identification as well as provides solutions that are fairer for each single student. By conducting a closer examination of the results for each individual student, it could be seen that WMCID-GA improved identification accuracy (ACC) for every individual. Additionally, students with WMC between 0.4 and 0.7 (60.3 % of students in the dataset) are identified better (average ACC = 0.898) than students below 0.4 (average ACC = 0.820) and above 0.7 (average ACC = 0.762). These results still compare favorably to the corresponding results for DeWMC with an average ACC of 0.818 and 0.684 respectively. Most likely, this is caused by the GA not

Table 3. Result comparison between WMCID and DeWMC (top result bolded)

Approach	ACC	LACC	%Match
WMCID-GA	**0.851**	**0.694**	**0.893**
DeWMC [12, 13]	0.809	0.442	0.809

Table 4. Minimum, maximum, and average weights and percentage of activated learning sessions per pattern

Pattern	Min	Max	Average	Activated
Linear navigation	3	13	7	89.98 %
Constant reverse navigation	50	99	82	78.62 %
Performing simultaneous tasks	81	100	97	8.25 %
Recalling learned material	10	33	22	58.86 %
Revisiting passed learning objects	36	84	62	60.19 %
Learning styles	2	17	10	100.00 %

having enough data for students with very high and very low WMC, as 28.6 % of students have a WMC higher than 0.7 and 11.1 % have a WMC below 0.4. Therefore, a larger sample size could help improve the results of WMCID-GA even further.

For each pattern, the minimum, maximum and average weights across all folds are shown in Table 4. Additionally, Table 4 shows the percentage of learning sessions in which a pattern was activated. These results indicate that constant reverse navigation, performing simultaneous tasks and revisiting passed learning objects are more predictive of WMC than other patterns, however, further investigations have to be done with respect to performing simultaneous tasks as such behavior was only found in 8.25 % of the learning sessions.

4 Conclusions

This paper has introduced WMCID-GA, an approach for identifying students' working memory capacity (WMC) from their behavior in learning systems. WMCID-GA extends the rule-based approach DeWMC by optimizing the weights of patterns through the use of a genetic algorithm. An evaluation with data from 63 students shows that WMCID-GA is outperforming DeWMC in all investigated metrics and therefore, can provide more accurate WMC results for more students. The results also indicate that different patterns have different impact on the WMC identification.

By improving the precision of WMC identification and making it possible to identify WMC automatically while students learn, learning environments can be personalized, providing students with individualized recommendations for learning activities that help balancing the cognitive load to their WMC. By optimizing the cognitive load, students can have better learning outcomes and may require less time to learn [3, 4]. Furthermore, more accurate WMC information can help students make better choices for self-regulated learning by taking their WMC into account while teachers may make better individualized suggestions to their students to help them learn.

Future work will deal with investigating other optimization algorithms and hybrid algorithms for the given problem, to overcome some of the weaknesses of GAs.

Acknowledgement. The authors acknowledge the support of this research by Alberta Innovates Technology Futures, Alberta Innovation and Advanced Education, Athabasca University and NSERC. This work was also supported by a grant of the Romanian National Authority for Scientific Research and Innovation, CNCS – UEFISCDI, project number PN-II-RU-TE-2014-4-2604.

References

1. Graf, S., Liu, T.-C., Chen, N.-S., Kinshuk, Yang, S.J.H.: Learning styles and cognitive traits – their relationship and its benefits in web-based educational systems. Comput. Hum. Behav. **25**(6), 1280–1289 (2009)
2. Miller, G.A.: The magical number seven, plus or minus two: some limits on our capacity for processing information. Psychol. Rev. **63**(2), 81 (1956)
3. Kirschner, P.A.: Cognitive load theory: implications of cognitive load theory on the design of learning. Learn. Instr. **12**(1), 1–10 (2002)
4. Teigen, K.H.: Yerkes-Dodson: a law for all seasons. Theory Psychol. **4**(4), 525–547 (1994)
5. Chang, T.-.W, Kurcz, J., El-Bishouty, M.M., Graf, S, Kinshuk: Adaptive recommendations to students based on working memory capacity. In: Proceedings of the International Conference on Advanced Learning Technologies, Athens, Greece, pp 57–61, July 2014. IEEE (2014)
6. Turner, M.L., Engle, R.W.: Is working memory capacity task dependent? J. Mem. Lang. **28** (2), 127–154 (1989)
7. Klein, K., Fiss, W.H.: The reliability and stability of the turner and engle working memory task. Behav. Res. Meth. Instr. Comput. **31**(3), 429–432 (1999)
8. Lin, T.: Cognitive trait model for adaptive learning environments. Dissertation, Massey University, Palmerston North, New Zealand (2007)
9. Gohar, A., Adams, A., Gertner, E., Sackett-Lundeen, L., Heitz, R., Engle, R., Haus, E., Bijwadia, J.: Working memory capacity is decreased in sleep-deprived internal medicine residents. J. Clin. Sleep Med. **5**(3), 191 (2009)
10. Ford, N., Chen, S.Y.: Individual differences, hypermedia navigation, and learning: an empirical study. J. Educ. Multimedia Hypermedia **9**(4), 281–311 (2000)
11. Graf, S., Lin, T., Kinshuk, : The relationship between learning styles and cognitive traits – getting additional information for improving student modelling. Comput. Hum. Behav. **24** (2), 122–137 (2008)
12. Chang, T.-W., El-Bishouty, M.M., Graf, S., Kinshuk: An approach for detecting students' working memory capacity from their behavior in learning systems. In: Proceedings of the International Conference on Advanced Learning Technologies, Beijing, China, pp 82–86, July 2013. IEEE (2013)
13. Chang, T.-W., El-Bishouty, M.M., Kinshuk, Graf, S.: Identifying students' working memory capacity in learning systems. Technical report (2016)
14. Felder, R.M., Silverman, L.K.: Learning and teaching styles in engineering education. Eng. Educ. **78**(7), 674–681 (1988)
15. Grefenstette, J.J.: Optimization of control parameters for genetic algorithms. IEEE Trans. Syst. Man Cybern. **16**(1), 122–128 (1986)
16. Srinivas, M., Patnaik, L.M.: Genetic algorithms: a survey. Computer **27**(6), 17–26 (1994)

17. Felder, R.M., Solomon, B.A.: Index of learning styles North Carolina State University (1998). http://www.engr.ncsu.edu/learningstyles/ilsweb.html. Accessed 1 Jan 2016
18. Kohavi, R.: A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Mellish, C.S. (ed.) Proceedings of the 14th International Joint Conference on Artificial Intelligence, vol. 2, pp 1137–1145, August 1995. Morgan Kaufmann Publishers Inc. (1995)
19. Mitchell, T.: Machine learning, vol. 45. McGraw Hill, Burr Ridge (1997)