Solving the Individualized Instructional Content Delivery Problem Using Ontology and Metaheuristics on the Set Covering Problem: An Experimental Analysis

Clarivando F. Belizário Júnior[1], Fabiano A. Dorça[1], Luciana Assis[2], Alessandro Vivas[2]

[1]Faculdade de Computação – Universidade Federal de Uberlândia (UFU)
Uberlândia – MG – Brasil

[2]Faculdade de Ciências Exatas
Universidade Federal dos Vales do Jequitinhonha e Mucuri (UFVJM)
Diamantina – MG – Brasil

{clarivando,fabianodor}@ufu.br, {alessandrovivas,lpassis}@ufvjm.edu.br

Abstract. *Intelligent Tutoring Systems (ITSs) based on a step-by-step problem-solving approach are limited in terms of compatible content. On the other hand, recommendation systems can suggest various content types but lack the granularity of concepts found in step-by-step approaches. This study addresses this challenge by proposing a method to recommend instructional content from diverse knowledge domains while incorporating the refined concepts of ITSs. To tackle this issue, the instructional content delivery problem (LORP) is formulated as a set covering problem, classified as NP-hard. We show that a PSO-based algorithm is a good candidate to solve LORP in a better runtime than the exact algorithm and with better solutions than the greedy heuristic. By leveraging collaborative filtering and an ontology that models students' knowledge, learning styles, and search parameters, the approach offers more individualized content.*

1. Introduction

Intelligent tutoring systems (ITSs) are computer systems that use techniques from Artificial Intelligence and Cognitive Psychology to provide feedback to students without the need for human intervention [Bernacki et al. 2014]. One of the challenges that these systems face is the Individualized Instructional Content Delivery Problem (LORP), which is addressed in the literature through several techniques such as content-based filtering, collaborative filtering, and hybrid recommendation. However, these techniques can be affected by the rating sparsity [Zhao et al. 2015], a problem that emerges when only a small number of students have rated a particular instructional content or learning object (LO) and there is no overlap in the classification preferences, and by the cold-start problem [Adomavicius and Tuzhilin 2005] that occurs when it is impractical to provide dependable recommendations because there are no initial evaluations available for new students or educational resources.

ITSs are effective in providing step-by-step feedback to students in problem-solving tasks [VanLehn 2006], but this approach is not suitable for most content [Soofi and Ahmed 2019]. On the other hand, recommendation systems (RSs) can recommend content from different areas but may overlook the specific concepts that students need to learn, which limits the personalization of recommendations.

The main contribution of this work to the e-learning RSs is an approach that combines ontology-based recommendation and collaborative filtering techniques for the delivery of LOs based on concepts and the reuse of web content reducing the rating sparsity and cold-start problems. The ontology [Gruber 1993] models LOs and the students' knowledge level and profile, and it implements inference rules to aid the recommendation process. In addition, we formalize the LORP as the Set Covering Problem (SCP) [Garey and Johnson 1979] and we adapt four algorithms to solve it, providing a more personalized delivery of LOs that cover the concepts that the student needs to master.

The remainder of the paper is organized as follows: Section 2 presents the background; in Section 3, we discuss the related work; Section 4 details the proposed approach; Section 5 presents experiments and results; and finally, Section 6 outlines the discussion of the results, conclusions and future work.

2. Background

We provide an overview of the Semantic Web in Section 2.1, which is a technology that use ontologies to semantically represent the vast content of the traditional Web. In this work, the ontology is used to store information about students and LOs; educational standards used to structure this information are presented in Section 2.2. Section 2.3 introduces the main filtering and recommendation techniques that support RSs. Our approach uses the ontology-based recommendation technique, in which the domain model and the learner model are structured in an ontology. Our RS uses these models to solve the LORP.

2.1. Semantic Web

The Semantic Web (SW), which was introduced by Berners-Lee in 2001 [Berners-Lee et al. 2001], is an extension of the traditional Web that includes semantic information using XML (eXtensible Markup Language), RDF (Resource Definition Framework), and OWL (Web Ontology Language). OWL is widely used for knowledge representation and implementation of ontologies, which can be compared to non-relational databases that can be queried using SPARQL, a language similar to SQL. Ontologies have the advantage of facilitating new knowledge discovery through inference rules expressed in the SWRL (Semantic Web Rule Language) [Horrocks et al. 2004], as it provides a means of representing knowledge using logic and reasoning.

2.2. Modelling learning objects and students

The IEEE-LOM [LTSC 2002] is a widely-used metadata standard for describing LOs. It has nine categories, with the General and Educational categories being the most important. The General category includes fields such as Entry, which stores the LO's link, and the keyword field, which can store the concepts that the LO covers. The Educational category provides pedagogical information about LOs, such as their type and degree of difficulty. Although not all fields of IEEE-LOM are widely used, some extensions like Customised Learning Experience Online (CLEO) [CLEOLab 2003] can be used to expand the vocabulary of specific fields.

Computer systems also need to model students, including their learning styles. The most suitable model for this is the FSLSM (Felder-Silverman Learning Style Model) [Felder et al. 1988]. It covers more psychological aspects than other models

[Deborah et al. 2014] and has four polar dimensions: Input (Visual and Verbal), Organization (Sequential and Global), Perception (Sensitive and Intuitive), and Processing (Active and Reflective). The Index of Learning Styles questionnaire [Soloman and Felder 2005] is one of the instruments used to assess student preferences in these four dimensions.

2.3. Filtering and recommendation techniques

Content-based filtering (CBF) [Vanetti et al. 2010] recommends objects to the target user based on the content characteristics of objects that the user has liked in the past. However, the disadvantage of CBF is that it only recommends LOs similar to the user's past experience. In contrast, collaborative filtering (CF) [De Medio et al. 2020] considers the recommendation history of other students to suggest new LOs for the target student. CF uses object evaluations (see Figure 1) to calculate the similarity of users or objects and make recommendations. Both CBF and CF techniques suffer from rating sparsity and cold-start problems while knowledge-based (KB) recommendation aggregates [Tarus et al. 2017] knowledge about the student and learning materials to alleviate these problems. Ontology-based recommendation [Tarus et al. 2017] is a type of KB recommendation that uses an ontology to represent this knowledge.

KB recommendation in Figure 1 predicts the score that the target student L_1 would assign to the instructional content O_2, based on the grade given by other students at the same level as L_1, whereas CF only takes the ratings of the LOs into account. KB is a type of CF that aggregates contextual information about students, which helps to reduce the rating sparsity and cold-start issues in CF.

Learner	O_1	O_2	O_3
L_1	4	?	5
L_2	3	3	4
L_3	4	4	3
L_4	1	5	3
L_5	5	4	5

Rating matrix of LOs for collaborative filtering

Learner	LO	Level	Rating
L_1	O_2	Beginner	2
L_2	O_2	Advanced	3
L_3	O_2	Beginner	2
L_1	O_2	Intermediate	?
L_3	O_2	Intermediate	5

Rating matrix of LOs for knowledge-based recommendation

Figure 1. Rating matrix of CF and KB recommendation

3. Related Work

The research in educational resource recommendation often combines recommendation techniques with ontologies and the Web, including Wikipedia, as shown in Table 1. Wiki content can be recommended to the teacher to create courses [Limongelli et al. 2015] or recommended to the student [Belizário Júnior and Dorça 2018]. In this previous work, we defined LORP as a SCP and solved it using a Genetic Algorithm (GA). Later in [Pereira et al. 2020], we also solved it using Prey-predator algorithm (PPA) [Tilahun and Ong 2015] and Particle Swarm Optimization (PSO). In [Falci et al. 2019], this problem was addressed using a greedy heuristic algorithm that selects LOs based on the student's learning style while covering a wide range of concepts. The heuristic algorithm is faster than GA, particularly for larger instances with thousands of LOs. However, GA-based solutions, like the Compatible Genetic Algorithm (CGA) proposed in [Christudas et al. 2018] for LO delivery, can be good when LORP is not based on SCP.

As shown in Section 4.3, we have improved the LORP definition to use collaborative filtering [Belizário Júnior et al. 2020] and hint-type LOs [Belizário Júnior et al. 2023]. By doing so, we can take into account the specific concepts that each student needs to learn. While previous works [Ouf et al. 2017, Pereira et al. 2018] have used the Web for content reuse (including LO repositories) and/or SW technologies, they do not combine the recommendation of fine-grained concepts with content from different areas of knowledge. Therefore, we proposed an approach for recommending LOs from diverse knowledge areas, while considering fine-granularity concepts.

Table 1. Comparison of related literature with the proposal of this work

Reference	Web content reuse	Ontology or Semantic Web technologies	LOs recommendation technique	LOs coverage using fine-grained concepts from different areas of knowledge
[Limongelli et al. 2015]	Yes	No	CBF and CF	No
[Belizário Júnior and Dorça 2018]	Yes	Yes	GA	No
[Falci et al. 2019]	Yes	Yes	Greedy alg.	No
[Belizário Júnior et al. 2020]	Yes	Yes	CF, SWRL, PSO	No
[Pereira et al. 2020]	Yes	Yes	GA, PPA, PSO	No
[Christudas et al. 2018]	No	No	CGA	No
[Belizário Júnior et al. 2023]	Yes	Yes	CF, SWRL, Exact and Greedy alg.	Yes
[Ouf et al. 2017]	No	Yes	SWRL	No
[Pereira et al. 2018]	Yes	Yes	SPARQL	No
Our proposal	*Yes*	*Yes*	*CF, SWRL, SPARQL, GA, PSO, Exact and Greedy alg.*	*Yes*

The main advantage of this work compared to previous works [Belizário Júnior et al. 2020, Belizário Júnior et al. 2023] is to consider GA and a PSO-based algorithm to solve LORP defined in terms of hint-type LOs. Furthermore, our approach implements SPARQL queries to find more content from different areas of knowledge and inference rules to aid the recommendation process. It also addresses the shortcomings of previous approaches and makes advancements in the field by alleviating issues related to the rating sparsity and cold-start problems. Additionally, the ontology is employed to represent refined LOs and the corresponding concepts encompassed by each LO, providing a personalized delivery of LOs that cover the specific concepts that a student needs to learn.

4. Proposed Approach

The proposed RS shown in Figure 2 utilizes a hybrid recommendation approach that combines collaborative filtering and ontology-based recommendation techniques. Wikipedia serves as an excellent digital encyclopedia, with millions of articles available in several languages and under a Creative Commons BY-SA license, which can be copied and modified. The RS interface, such as a chatbot, captures users' search parameters, including the concepts they want to learn and their preferences. These search parameters are stored as ideal LOs in the ontology, representing the ideal characteristics expected in recommended LOs.

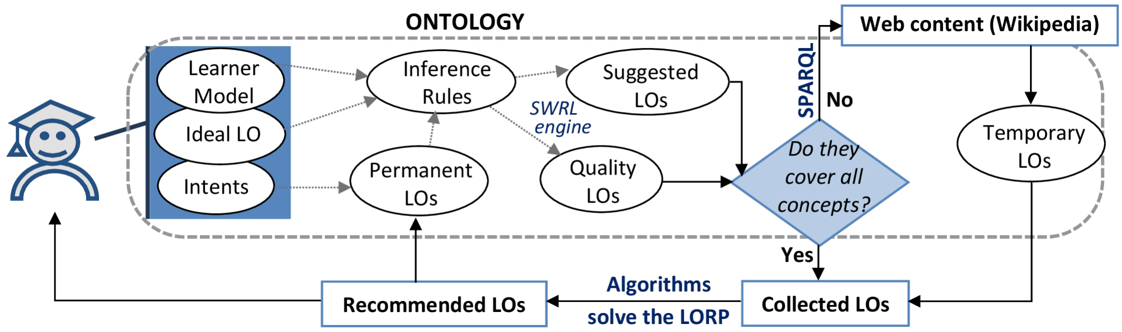


Figure 2. Overall of the proposed recommendation system

Using inference rules, the ontology suggests LOs to the student, and if necessary, web content, such as Wikipedia pages discovered through SPARQL queries, is transformed into temporary LOs in the ontology to cover any uncovered concepts. These temporary LOs are combined with the suggested and quality LOs to form the set of collected LOs, which are used as input for algorithms that solve the LORP, and the best solution found is recommended to the student, with the temporary LOs becoming permanent LOs in the ontology, which can be reused in future recommendations. In the following sections, we present the improvements made to the ontology in Section 4.1. Section 4.2 formalizes the LORP using the SCP, and Section 4.3 describes how the cost of LOs is calculated.

4.1. Ontology improvements

The ontology utilized in this study was initially introduced in [Belizário Júnior and Dorça 2018] and is designed to store information about students and rather than containing the LOs themselves, the ontology stores their metadata according to the nine categories of the IEEE-LOM standard and its CLEO extension. We made four main improvements to this ontology to enhance the delivery process of LOs.

The first improvement was the creation of the QualityLOs class in the ontology. Quality and suggested LOs are inferred LOs that cover at least one of the concepts that the student needs to master. The difference is that quality LOs receive a bonus so that they are more likely to be recommended to the student since they were specifically created to answer/solve the learner's specific doubts. The second improvement was the incorporation of the hint type into the ontology, which are permanent LOs that can be deduced as instances of the QualityLOs class. The way in which a link is established between the hint-type LOs and intents (students' doubts) is shown in Figure 3. The intent_001, recognized by a chatbot, has two hints (LO_2 and LO_3) that are instances of the PermanentLOs class and are connected to the intent via the hasLearningObject property. Each hint has a unique identification with a URI, such as http://localhost/hint_1 and http://localhost/hint_2 for LO_2 and LO_3, respectively.

The third improvement was the storage of user's search parameters in the ontology's IdealLO class, where the intent name is a significant search parameter for identifying the concepts the student has doubts about and the LOs, particularly hint-type LOs,

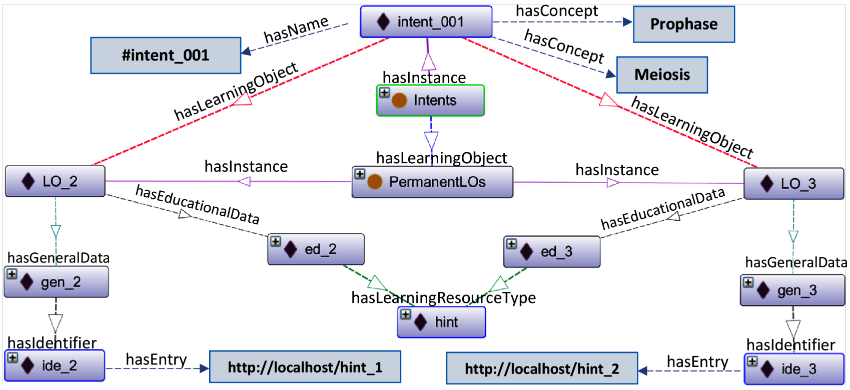


Figure 3. Relationship (hasLearningObject) between intent and hint-type LOs

created for that particular intent. The ideal LO also takes into account the student's learning style, and uses SWRL rules with two different purposes. First, some rules are used to infer the types of LOs appropriate to the student's learning style based on the theory described by [Graf et al. 2010], who address which types of LOs should be recommended for each type of student profile associated with the FSLSM. Second, other rules are used to select the LOs that are similar to the user's search parameters.

The fourth improvement includes two new rules to suggest helpful LOs, including hints, that are related to the collected LOs. The intent name stored in the idealLO is used to identify the specific topic that the student is struggling with. The first new rule uses the concepts associated with this intent, such as Prophase and Meiosis in Figure 3, to suggest LOs in the ontology that cover at least one of these concepts. Meanwhile, the second new rule ensures that all hints associated with this intent become instances of the QualityLOs class. Our recommendation system is able to infer these suggested and quality LOs using these new rules, respectively. Only if the suggested and quality LOs are insufficient to cover all of the concepts the student needs to learn, are temporary LOs created with web content.

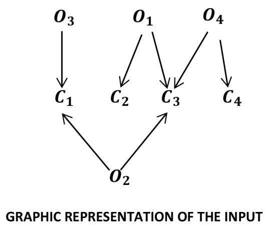
4.2. LORP defined as the SCP

The SCP formalized in Eq. (1) is defined as the task of covering all rows of a zero-one matrix a_{ij} with a subset of columns at the lowest possible cost. The column j has cost $c_{\mathrm{j}} > 0$ and is part of the solution if x_{j} = 1, otherwise $x_{\mathrm{j}} = 0$.

$$\begin{aligned} &\mathit{Minimise} \ \sum_{j=1}^{n} c_j x_j \\ \mathit{Subject\ to} \quad 1 \leq \sum_{j=1}^{n} a_{ij} x_j, \quad i = 1, ..., m, \quad & x_j \in \{0, 1\} \end{aligned} \tag{1}$$

The LORP is a problem that seeks to identify the minimum-cost coverage of LOs to cover all concepts, and it corresponds to the SCP formulated by Eq. (1). The value of c_j is calculated in Section 4.3. The matrix a_{ij} is filled with user input concepts and LOs gathered by the RS, where each row i corresponds to an input concept C_i, and each column j is linked to an LO O_j resulting from the set of collected LOs. If O_j covers C_i, $a_{ij} = 1$, otherwise $a_{ij} = 0$.

Figure 4 shows a small LORP instance, with the input matrix, its graphical representation and the cost vector. The LOs O_1, O_2, O_3 and O_4 have costs of 2, 5, 2, and 3, respectively. The solution for this example is $\{O_1, O_3, O_4\}$ with cost 7.



COST VECTOR

2	5	2	3

INPUT MATRIX

	O_1	O_2	O_3	O_4
C_1	0	1	1	0
C_2	1	0	0	0
C_3	1	1	0	1
C_4	0	0	0	1

Figure 4. Input matrix and cost vector of the LO Covering Problem

4.3. Improvements in cost calculation

The calculation of the cost as $c_j = diss(O_{ideal}, O_j)$ was initially proposed in [Belizário Júnior and Dorça 2018]. The $diss(O_{ideal}, O_j)$ value is inversely proportional to the degree of similarity between O_{ideal} and O_j. Six parameters of O_j (title, interactivity type, learning resource type, interactivity level, semantic density and difficulty) are compared with the corresponding parameters of O_{ideal} given by the user.

Later, the cost c_j was reformulated in [Belizário Júnior et al. 2020] as: $c_j = diss(O_{ideal}, O_j) + (1 - P_j^L)$. This prediction P_j^L represents the relevance that O_j has for the target student L. This relevance is calculated using collaborative filtering.

In this paper, we improve this cost calculation to make refined LO recommendations using hint-type LOs for this. The new cost is formally defined as:

$$c_j = diss(O_{ideal}, O_j) + (1 - P_j^L + 1 - H_j) * \max_{j \in \{1,\ldots,n\}} diss(O_{ideal}, O_j) \tag{2}$$

where the max operator is a weight given to P_j^L and H_j to assign them the same importance as $diss$.

The RS has two delivery modes. If the student has doubts when studying some content or solving an exercise, then **the more hints the better** ($H_j = 1$ if the LO O_j is of the hint type, and $H_j = 0$ otherwise) to provide a more fine-grained recommendation. On the other hand, if the student has no doubts and needs to learn new concepts, then **the less hints the better** to recommend ($H_j = 0$ if the LO O_j is of the hint type, and $H_j = 1$ otherwise). In this case, the RS should recommend other types of LOs, such as lectures and exercises. Our two research questions derived from Eq. (2) are:

- **Research question 1**: Does the use of collaborative filtering (variable P_j) contribute to the delivery of the LOs with the best rating for the student?
- **Research question 2**: Does the use of H (hint: fine-grained LOs) in the calculation of c_j in the objective function improve the quality of LOs recommendation in relation to the number of hints expected by students?

The value P_j^L in Eq. (2) represents the importance of a LO for a student. This value ranges between 0 and 1 and is calculated using the k-Nearest Neighbours approach proposed in [Tarus et al. 2018], which finds the k most similar students to the target student, based on their ratings of the same LO. The similarity calculation only considers students with similar characteristics to the target student, such as knowledge level or learning style. This KB recommendation method can help address the rating sparsity and cold-start problems. We found that using only the collaborative filtering approach was sufficient for experimental tests. However, the KB recommendation method may be useful in a real learning context.

5. Experimental analysis

The algorithms were implemented in Python and executed on a notebook running Windows 10 OS, with an AMD Quad-Core A10-9600P processor operating at 2.40 GHz and 8GB of RAM. The experiments followed a randomized complete block design (RCBD) where problems were treated as blocks. To avoid assuming normality, we used the non-parametric Wilcoxon test. Post-hoc analysis for effect size estimation was performed using Tukey's test [Tukey 1949].

The dataset consists of 24 instances containing symbolic data. Each instance has an input matrix, a cost vector (see Figure 4), either 2, 6, 10, 25, 40 or 55 rows (concepts), and either 100, 500, 2000 or 10000 columns (LOs). However, in our dataset, the cost vector is interpreted as a dissimilarity vector (*diss* in Eq. (2)), which is utilized to calculate the cost vector. The input matrix and dissimilarity vector in the instances were designed to simulate a real-world scenario. In addition to "diss", the determination of P_j and H_j is necessary to calculate the cost using Eq. (2). To achieve this, we used a rating matrix proposed in [Belizário Júnior et al. 2023] that simulates a real-life scenario with ratings provided by students for the LOs they evaluated.

The exact algorithm from Python Pulp Library [Mitchell et al. 2011] and the greedy algorithm [Golab et al. 2015] were implemented as in [Belizário Júnior et al. 2023]. The third algorithm used to solve the LORP is an adaptation of the Particle Swarm Optimization, named Jumping Particle Swarm Optimization (JPSO) [Balaji and Revathi 2016]. Ten particles were used to carry out the tests. In general, JPSO converges in the first few iterations.

We also used the genetic algorithm proposed in [Belizário Júnior and Dorça 2018]. In it, each individual is a vector of integers with m positions (one for each row). The integer value at a given position i corresponds to the column covering row i. The tournament selection is used to choose 4 individuals (two pairs). In each iteration, two new individuals are generated after applying a fitness-based crossover operator in each pair. The mutation is applied with a probability of 10% to replace a randomly chosen LO (integer value) with another one. Two individuals with above-average fitness (less fit) are randomly chosen to be replaced by the two new individuals.

To evaluate the CF (prediction) used in the proposed approach, we implemented two versions of each selected algorithm to solve the LORP. The difference between them relates to how the c_j cost is calculated. In the first, the cost is calculated by Eq. (3) without using the prediction variable P, while in the second, the cost is calculated by Eq.

(2) using the prediction.

$$c_j = diss(O_{ideal}, O_j) + (1 - H_j) * \max_{j \in \{1,\ldots,n\}} diss(O_{ideal}, O_j) \tag{3}$$

Each instance is run 10 times resulting in 24 average ratings. As the data do not have a normal distribution, Table 2 presents the median of these 24 values. The two versions of each algorithm are compared by the columns *No* (c_j is calculated by Eq. (3)) and *Yes* (c_j is calculated by Eq. (2)). These values indicate the significance of the LO to the student. The higher the value, the more important the LO is to the student. Statistically significant differences (*p-value* < 0.05) exist between the *Yes* and *No* variables in each analyzed algorithm, indicating that incorporating the predictive variable into the cost function has a notable impact in the two delivery modes: More(M)/Less(L) hints the better. The magnitude diff demonstrates that the use of P (CF) in the calculation of the OF (c_j) increases the average rate of the LORP solution, improving the quality of the LOs recommended to the learners. So the algorithms utilized for solving the LORP with prediction generate solutions consisting of LOs with higher ratings.

Table 2. Comparison of the average ratings in the solutions with (Yes) and without (No) the P variable in solving the LORP

		Median (*No*)	Median (*Yes*)	p-value (Wilcoxon)	Magnitude diff (Tukey)
Exact	M	0.715	0.778	$< .001$	0.1069
	L	0.697	0.728	$< .001$	0.0428
Greedy	M	0.710	0.796	$< .001$	0.1189
	L	0.700	0.722	$< .001$	0.0420
JPSO	M	0.701	0.773	$< .001$	0.1088
	L	0.701	0.728	$< .001$	0.0502
GA	M	0.697	0.715	0.009	0.0628
	L	0.702	0.711	0.006	0.0188

Note: If *p-value* < 0.05, then there is a statistically significant difference between the variables. *Legend*: More(M)/Less(L) hints the better.

To evaluate the H (hint) used in the proposed approach, we also implemented two versions of each selected algorithm to solve the LORP. In the first, the cost is calculated by Eq. (4) without using the hint variable H, while in the second, the cost is calculated by Eq. (2) using the hint variable.

$$c_j = diss(O_{ideal}, O_j) + (1 - P_j^L) * \max_{j \in \{1,\ldots,n\}} diss(O_{ideal}, O_j) \tag{4}$$

From Table 3, it can be seen that the *Yes* variables have a lower median value than *No* variables in all algorithms when the less hints the better. In this case, the magnitudes are negative, demonstrating that the number of hints returned when H is applied to the OF is smaller. On the other hand, *Yes* variables have a higher median value than *No* variables in all algorithms when the more hints the better. In this recommendation mode, the magnitudes are positive, demonstrating that the number of hints returned when H is applied to the OF is higher. So the use of H (hint: fine-grained LOs) in the calculation of c_j in the OF improves the quality of LOs recommendation in relation to the expected number of hints.

Table 3. Comparison of the average number of hints in the solutions with (yes) and without (no) the H variable in solving the LORP

		Median (*No*)	Median (*Yes*)	p-value (Wilcoxon)	Magnitude diff (Tukey)
Exact	M	2.0	4.5	0.001	2.4167
	L	2.0	0.0	$<.001$	-2.6667
Greedy	M	3.0	6.0	$<.001$	3.1667
	L	3.0	0.0	$<.001$	-2.8333
JPSO	M	2.5	5.4	0.001	2.5583
	L	2.5	0.0	$<.001$	-3.2958
GA	M	1.0	1.8	0.001	1.1458
	L	1.5	0.0	$<.001$	-1.7500

Note: If *p-value* < 0.05, then there is a statistically significant difference between the variables. *Legend*: More(M)/Less(L) hints the better.

Therefore, the variables P and H employed in Eq.(2) contribute to the recommendation of solutions that consist of highly-rated LOs and an appropriate quantity of hints, respectively, based on the chosen delivery mode, confirming the hypotheses derived from research questions 1 and 2. When fewer hints are better, the Greedy, JPSO, and GA algorithms find the exact solution for 9, 16, and 5 out of 24 instances, respectively; and the more hints the better, they find the exact solution for 9, 17, and 8 out of 24 instances. Thus, JPSO presents better solutions than Greedy, but Greedy is faster than JPSO, while GA is the worst algorithm in this educational dataset. The Exact algorithm's drawback lies in its significant time consumption when dealing with larger instances. Consequently, the Exact algorithm is most suitable for solving smaller LORP instances. However, if prioritizing shorter execution time is essential, the greedy algorithm emerges as the optimal choice, but to obtain better solutions than Greedy without spending Exact's runtime, the best choice is JPSO.

6. Discussion and Conclusions

This paper proposes a RS for recommending e-learning resources and fine-grained LOs based on the student's learning style, knowledge, and search parameters. We formulate the LORP as SCP, addressing the issue of LO recommendation by incorporating the necessary concepts for student learning. Additionally, an ontology-based approach is implemented to provide more detailed LO recommendations, combining web content reuse to address the limited content diversity in ITSs and the lack of refined concepts in traditional RSs.

The findings demonstrate that our approach outperforms recommendation strategies that consider only the user's search parameters when recommending LOs [Belizário Júnior and Dorça 2018, Falci et al. 2019] and those that combine the user's search parameters with CF ([Belizário Júnior et al. 2020]). To overcome the limitations of this research, we intend to test our approach in a real educational scenario as a future work. In additon, we will explore other recommendation strategies from related works that seem promising, such as sequential pattern mining for purposes of prediction. Our forthcoming research will concentrate on the integration of the proposed RS into a learning environment, such as Moodle, with the aim of providing a more refined recommendation of LOs that closely resemble the recommendations made by ITSs.

Acknowledgments

This study was financed in part by the Coordination for the Improvement of Higher Education Personnel—Brazil (CAPES)—Finance Code 001, and it was supported by the National Council for Scientific and Technological Development (CNPq; Grant 402431/2021-9).

References

Adomavicius, G. and Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE transactions on knowledge and data engineering*, 17(6):734–749.

Balaji, S. and Revathi, N. (2016). A new approach for solving set covering problem using jumping particle swarm optimization method. *Nat. Comput.*, 15(3):503–517.

Belizário Júnior, C. F. and Dorça, F. (2018). Uma abordagem para a criaçao e recomendaçao de objetos de aprendizagem usando um algoritmo genético, tecnologias da web semântica e uma ontologia. In *Brazilian Symposium on Computers in Education*, volume 29, pages 1533–1542, Fortaleza, CE. SBC.

Belizário Júnior, C. F., Dorça, F., Andrade, A. V., and Assis, L. P. (2020). Avanços na recomendação personalizada de objetos de aprendizagem através da utilização de meta-heurísticas clássicas associadas aos problemas de cobertura de conjuntos e de máxima cobertura: Uma análise experimental. In *Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE)*, volume 31, pages 1383–1392, Natal, RN. SBC.

Belizário Júnior, C. F., Dorça, F., Assis, L. P., and Andrade, A. V. i. p. (2023). Solving the personalised learning objects recommendation problem using ontology and metaheuristics on the set covering problem: An experimental analysis. *International Journal of Learning Technology*, 18.

Bernacki, M. L., Aleven, V., and Nokes-Malach, T. J. (2014). Stability and change in adolescents' task-specific achievement goals and implications for learning mathematics with intelligent tutors. *Computers in Human behavior*, 37:73–80.

Berners-Lee, T., Hendler, J., and Lassila, O. (2001). The semantic web. *Scientific american*, 284(5):34–43.

Christudas, B. C. L., Kirubakaran, E., and Thangaiah, P. R. J. (2018). An evolutionary approach for personalization of content delivery in e-learning systems based on learner behavior forcing compatibility of learning materials. *Tel. and Inf.*, 35(3):520–533.

CLEOLab (2003). Cleo extensions to the ieee learning object metadata. *CLEO Collaborative Partners - Cisco Systems, Inc., IBM Corporation, Microsoft Corporation, Thomson NETg. CLEO White Paper*.

De Medio, C., Limongelli, C., Sciarrone, F., and Temperini, M. (2020). Moodlerec: A recommendation system for creating courses using the moodle e-learning platform. *Computers in Human Behavior*, 104:106168.

Deborah, L. J., Baskaran, R., and Kannan, A. (2014). Learning styles assessment and theoretical origin in an e-learning scenario: a survey. *AI Review*, 42(4):801–819.

Falci, S. H., Dorça, F. A., Falci, D. H. M., and Vivas, A. (2019). A low complexity heuristic to solve a learning objects recommendation problem. In *2019 IEEE 19th ICALT*, volume 19, pages 49–53. IEEE.

Felder, R. M., Silverman, L. K., et al. (1988). Learning and teaching styles in engineering education. *Engineering education*, 78(7):674–681.

Garey, M. R. and Johnson, D. S. (1979). *Computers and intractability*. Series of books in the mathematical sciences. W. H. Freeman & Co., New York.

Golab, L., Korn, F., Li, F., Saha, B., and Srivastava, D. (2015). Size-constrained weighted set cover. In *31st Int. Conf. on Data Engineering*, pages 879–890. IEEE.

Graf, S., Kinshuk, and Ives, C. (2010). A flexible mechanism for providing adaptivity based on learning styles in lms. In *10th ICALT*, pages 30–34. IEEE.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge acquisition*, 5(2):199–220.

Horrocks, I., Patel-Schneider, P. F., Boley, H., Tabet, S., Grosof, B., Dean, M., et al. (2004). Swrl: A semantic web rule language combining owl and ruleml. *W3C Member submission*, 21(79):1–31.

Limongelli, C., Gasparetti, F., and Sciarrone, F. (2015). Wiki course builder: a system for retrieving and sequencing didactic materials from wikipedia. In *2015 Inter. Conf. on Inform. Tech. Based Higher Education and Training*, pages 1–6. IEEE.

LTSC (2002). Standard for learning object metadata (ieee 1484.12.1). *Learning Technology Standards Committee*.

Mitchell, S., OSullivan, M., and Dunning, I. (2011). Pulp: a linear programming toolkit for python. *The University of Auckland, Auckland, New Zealand*, 65.

Ouf, S., Ellatif, M. A., Salama, S. E., and Helmy, Y. (2017). A proposed paradigm for smart learning environment based on semantic web. *Computers in Human Behavior*, 72:796–818.

Pereira, Júnior, C., Belizário, Júnior, C. F., Araújo, R. D., and Dorça, F. A. (2020). Personalized recommendation of learning objects through bio-inspired algorithms and semantic web technologies: an experimental analysis. In *Anais do XXXI Simpósio Brasileiro de Informática na Educação*, pages 1333–1342. SBC.

Pereira, C. K., Campos, F., Ströele, V., David, J. M. N., and Braga, R. (2018). Broad-rsi–educational recommender system using social networks interactions and linked data. *Journal of Internet Services and Applications*, 9(7):1–28.

Soloman, B. A. and Felder, R. M. (2005). Index of learning styles questionnaire. *NC State University. Available online at: http://www.engr.ncsu.edu/learningstyles/ilsweb.html (last visited on 12.05.2023)*, 70.

Soofi, A. A. and Ahmed, M. U. (2019). A systematic review of domains, techniques, delivery modes and validation methods for intelligent tutoring systems. *International Journal of Advanced Computer Science and Applications*, 10:99–107.

Tarus, J. K., Niu, Z., and Kalui, D. (2018). A hybrid recommender system for e-learning based on context awareness and sequential pattern mining. *Soft Computing*, 22(8):2449–2461.

Tarus, J. K., Niu, Z., and Yousif, A. (2017). A hybrid knowledge-based recommender system for e-learning based on ontology and sequential pattern mining. *Future Generation Computer Systems*, 72:37–48.

Tilahun, S. L. and Ong, H. C. (2015). Prey-predator algorithm: a new metaheuristic algorithm for optimization problems. *International Journal of Information Technology & Decision Making*, 14(06):1331–1352.

Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics*, 5(2):99–114.

Vanetti, M., Binaghi, E., Carminati, B., Carullo, M., and Ferrari, E. (2010). Content-based filtering in on-line social networks. In *International Workshop on Privacy and Security Issues in Data Mining and Machine Learning*, pages 127–140. Springer.

VanLehn, K. (2006). The behavior of tutoring systems. *International journal of artificial intelligence in education*, 16(3):227–265.

Zhao, X., Niu, Z., Chen, W., Shi, C., Niu, K., and Liu, D. (2015). A hybrid approach of topic model and matrix factorization based on two-step recommendation framework. *Journal of Intelligent Information Systems*, 44(3):335–353.