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

Clarivando F. Belizário Júnior[1] , **Fabiano A. Dorça**[1] , **Alessandro Vivas**[2] , **Luciana Assis**[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. *A limitation of recommending systems is that they do not guarantee that the recommended learning objects (LOs) cover all concepts to be learned by the student. To deal with this challenge, in this paper the Learning Object Recommendation Problem (LORP) is formulated as the Set Covering Problem and the Maximum Covering Problem, both NP-Hard class problems. Greedy and evolutionary metaheuristics were properly adapted, resulting in a promising approach to the LORP solution. It resulted in more personalized content for students using an ontology that models their knowledge, needs, learning styles and search parameters.*

Resumo. *Uma limitação dos sistemas de recomendação é não garantir que os objetos de aprendizagem (OAs) recomendados cobrem todos os conceitos que o estudante precisa aprender. Para lidar com esse desafio, neste artigo, o Problema de Recomendação de Objetos de Aprendizagem (PROA) é formulado como o Problema de Cobertura de Conjuntos e um problema similar ao Problema de Máxima Cobertura, ambos da classe NP-Hard. Adaptam-se meta-heurísticas gulosas e evolucionárias mostrando que elas são apropriadas para a solução do PROA. Este artigo contribui com o estado da arte com uma recomendação de OAs mais personalizada usando uma ontologia que modela as necessidades de conhecimento do estudante e o seu estilo de aprendizagem.*

1. Introdução

Os Sistemas Tutores Inteligentes (STIs) são sistemas de computadores que utilizam técnicas de Inteligência Artificial e da Psicologia Cognitiva para dar *feedback* aos alunos sem a intervenção humana. Um desafio inerente a esses sistemas é a Recomendação Personalizada de Objetos de Aprendizagem. Esse problema, doravante denominado PROA, é tratado na literatura por diversas técnicas, sendo as mais usadas a filtragem por conteúdo, a filtragem colaborativa e a combinação de ambas as técnicas (a filtragem híbrida). Essas técnicas sofrem do problema *cold-start* e de escassez de dados (*data sparsity*). No contexto do *e-learning*, o *cold-start* ocorre quando não é possível fazer recomendações

confiáveis devido à falta de avaliações (*ratings*) iniciais para novos estudantes ou recursos educacionais. A escassez de dados ocorre quando poucos estudantes têm avaliado o mesmo objeto de aprendizagem e não há sobreposição nas preferências de classificação.

Além disso, nesses sistemas, o conteúdo dos Objetos de Aprendizagem (OAs) recomendados é direcionado a um único assunto e não é garantido que eles cobrem múltiplos conceitos. Desconsiderar cada conceito que o aluno precisa aprender é uma limitação desses sistemas que impede uma recomendação mais personalizada ao conhecimento do estudante.

Neste artigo, propõe-se uma abordagem de recomendação híbrida que ameniza esses problemas usando uma ontologia para modelar o conhecimento sobre os estudantes e os recursos educacionais contribuindo com o estado da arte. A ontologia também modela os conceitos que o estudante domina e os conceitos que cada objeto de aprendizagem cobre, propiciando uma recomendação de Objetos de Aprendizagem (OAs) que cobrem os conceitos que o estudante ainda não domina. As principais contribuições deste trabalho para os sistemas de recomendação do *e-learning* são:

1. Uma abordagem que utiliza as filtragens baseada em conteúdo e colaborativa para a recomendação de objetos de aprendizagem baseada em conceitos e no reúso de conteúdo da Web;
2. A abordagem proposta ameniza os problemas *cold-start* e de escassez de dados por meio da recomendação baseada em conhecimento, que utiliza uma ontologia para modelar o conhecimento dos estudantes e dos recursos educacionais;
3. O problema de recomendação de OAs é formalizado como o Problema de Cobertura de Conjuntos reformulado e um problema similar ao Problema de Máxima Cobertura [Garey and Johnson 1979]. As meta-heurísticas utilizadas para resolver esses problemas são adaptadas para a solução do problema de recomendação do contexto do *e-learning*.

Na Seção 2 a seguir, discute-se sobre os trabalhos relacionados a este estudo. Na Seção 3, apresenta-se a abordagem proposta com duas formulações para o PROA. Os algoritmos adaptados para a resolução desse problema, os experimentos e os resultados são apresentados na Seção 4. Por fim, traçam-se as considerações finais e os trabalhos futuros na Seção 5.

2. Trabalhos Relacionados

Ao contrário da filtragem baseada em conhecimento (KB), as filtragens baseada em conteúdo (FBC) e colaborativa (FC) sofrem dos problemas *cold-start* e *data sparsity*. A KB agrega conhecimento sobre o estudante e os materiais de aprendizagem para aplicá-los no processo de recomendação. A recomendação baseada em ontologia é um tipo de KB que utiliza a ontologia para a representação desse conhecimento. Diversas pesquisas combinam essas técnicas com ontologias e/ou a Web, inclusive a Wikipédia, para a recomendação de recursos educacionais. Por exemplo, [Limongelli et al. 2015] criaram um módulo em um sistema para recomendação colaborativa de páginas *wiki* usadas por professores na criação de cursos. O professor-alvo se beneficia da recomendação realizada no passado para outros docentes que possuem um estilo de ensino similar ao seu.

[Belizário Júnior and Dorça 2018], por sua vez, apresentam uma abordagem que recomenda conteúdo *wiki* direto para o estudante-alvo sem usar o professor como in-

termediário. Essa abordagem seleciona as páginas *wiki* de melhor qualidade usando as classes de qualidade atribuídas a elas pelos usuários. As seções (dentro dessas páginas) que cobrem os conceitos que o estudante-alvo precisa aprender são recomendas a ele. A abordagem utiliza uma ontologia para a modelagem de estudantes e OAs. O problema de recomendação de OAs é formalizado como um Problema de Cobertura de Conjuntos e é resolvido por um Algoritmo Genético (AG).

Esse mesmo problema é resolvido por um algoritmo mais rápido considerando uma heurística gulosa, conforme demonstram [Falci et al. 2019]. A intuição subjacente à heurística é que os OAs que atendem ao estilo de aprendizagem do estudante enquanto cobrem mais conceitos tendem a entregar candidatos melhores para a solução final. O algoritmo que implementa essa heurística é mais rápido que o AG, principalmente, para instâncias com milhares de OAs, para as quais o AG pode se tornar impraticável dado o espaço de busca exponencial e o alto número de cálculos da função *fitness*.

Além de ontologias, a Web Semântica (WS) também conta com tecnologias que têm sido exploradas por alguns autores para a recomendação de OAs. [Ouf et al. 2017] desenvolveram uma ferramenta para um ecossistema de aprendizagem inteligente usando ontologias e regras SWRL (Semantic Web Rule Language) [Horrocks et al. 2004]. As ontologias são utilizadas para modelar os estudantes e para adequar os componentes do processo de aprendizagem aos alunos, tais como OAs, atividades de aprendizagem preferidas e métodos de ensino-aprendizagem relevantes.

[Pereira et al. 2018] criaram uma infraestrutura para a recomendação de recursos educacionais com base em informações como o perfil do usuário e o contexto educacional, extraídas da rede social Facebook. Tecnologias da WS e técnicas de extração de informações são utilizadas para extrair, enriquecer e definir os perfis e interesses dos usuários. A estratégia de recomendação é baseada em dados ligados, repositórios de OAs e vídeos, beneficiando-se do tempo que o usuário gasta na Web.

Nota-se que os trabalhos da literatura correlata utilizam-se da Web para o reúso de conteúdo (inclusive de repositórios de OAs) e/ou a utilização de tecnologias da WS, mas nenhum deles combina ontologias com a criação de novos recursos educacionais. O avanço desta pesquisa em relação a outros trabalhos é a criação de OAs pelo reúso de conteúdo da Web, mais a recomendação desses OAs com o auxílio de uma ontologia em conformidade com os padrões IEEE-LOM e sua extensão CLEO [CLEOLab 2003]. Além disso, utiliza-se uma estratégia baseada em conceitos para a recomendação de OAs com conteúdo mais personalizado às necessidades do estudante.

O principal avanço em relação ao trabalho inicialmente proposto em [Belizário Júnior and Dorça 2018] é que neste artigo considera-se o PCC (Problema de Cobertura de Conjuntos) reformulado e calcula-se o custo de cada OA usando uma estratégia de recomendação híbrida que considera o contexto do estudante, como seu estilo de aprendizagem e nível de conhecimento, bem como, os metadados dos OAs. Outra contribuição inédita deste artigo é a formulação do PROA como um problema similar ao Problema de Máxima Cobertura (PMC, do inglês Maximum Coverage Problem) [McGregor and Vu 2019], melhorando a proposta inicial em diversos aspectos. A Tabela 1 sumariza uma comparação desta abordagem com outros trabalhos.

Neste trabalho, considerou-se uma heurística gulosa (CWSC) e a otimização por

enxame de partículas (PSO, do inglês Particle Swarm Optimization) para a solução do PROA. O algoritmo CWSC também foi utilizado para a solução do k-PROA. O algoritmo Concise Weighted Set Cover (CWSC) resulta da motivação dos autores [Golab et al. 2015] em encontrar uma generalização para os problemas Weighted Set Cover e Maximum Coverage. A entrada do CWSC é um conjunto de n elementos, uma coleção de pesos para esses elementos, um valor inteiro de restrição de tamanho k e uma fração de cobertura mínima s. A saída é uma subcoleção de até k conjuntos cuja união contém ao menos sn elementos e cuja soma de pesos é mínima.

Tabela 1. Comparação da literatura correlata com a proposta deste trabalho

Referência	Reúso de conteúdo da Web	Ontologia ou tecnologias da Web Semântica	Técnica de recomend. de OAs	Cobertura de OAs usando conceitos
[Limongelli et al. 2015]	Sim	Não	FBC e FC	Não
[Ouf et al. 2017]	Não	Sim	SWRL	Não
[Pereira et al. 2018]	Sim	Sim	SPARQL	Não
Nossa proposta	**Sim**	**Sim**	**SWRL, PSO, FC e KB**	**Sim**

Uma das variantes discretas da otimização por enxame de partículas é o *jumping* PSO (JPSO), apresentado por [Garcia and Pérez 2008]. [Balaji and Revathi 2016] implementaram um algoritmo baseado no JPSO, o Jpso-Scp, para resolver o Problema de Cobertura de Conjuntos. Em cada iteração, o Jpso-Scp utiliza três etapas. Na primeira fase, uma solução viável é construída por meio de um atrator que é selecionado com base na cobertura e custo das colunas. Essa solução é refinada na segunda fase e, na última fase, as colunas redundantes são removidas da solução. Esse método melhora as soluções logo nas primeiras iterações a um baixo custo computacional.

3. Abordagem Proposta

A ontologia inicialmente proposta por [Belizário Júnior and Dorça 2018] é utilizada na abordagem proposta para modelar os OAs conforme o padrão IEEE-LOM e sua extensão CLEO e armazenar conhecimento sobre os estudantes. Além disso, a ontologia implementa regras SWRL para inferir os tipos de OAs adequados ao estilo de aprendizagem do estudante e para realizar a pré-seleção dos OAs que apresentam alguma similaridade em relação aos parâmetros de busca do usuário (ver Fig. 1).

No contexto do ensino-aprendizagem, imagine uma situação em que um estudante precise aprender quatro conceitos pertencentes ao conjunto finito $X = \{C_1, C_2, C_3, C_4\}$. Considere uma coleção de subconjuntos de X dada por $F = \{O_1, O_2, O_3, O_4\}$, sendo $O_1 = \{C_1, C_3\}$, $O_2 = \{C_3, C_4\}$, $O_3 = \{C_1\}$ e $O_4 = \{C_2, C_3\}$. Os conjuntos O_1, O_2, O_3 e O_4 têm, respectivamente, custos 5, 3, 2 e 2. O custo representa o grau de dissimilaridade entre o OA e os parâmetros de busca do usuário. Cada elemento de F é um objeto de aprendizagem que cobre um conjunto de conceitos. O OA O_1, por exemplo, cobre os conceitos C_1 e C_3. Nesse cenário, há dois objetivos relevantes:

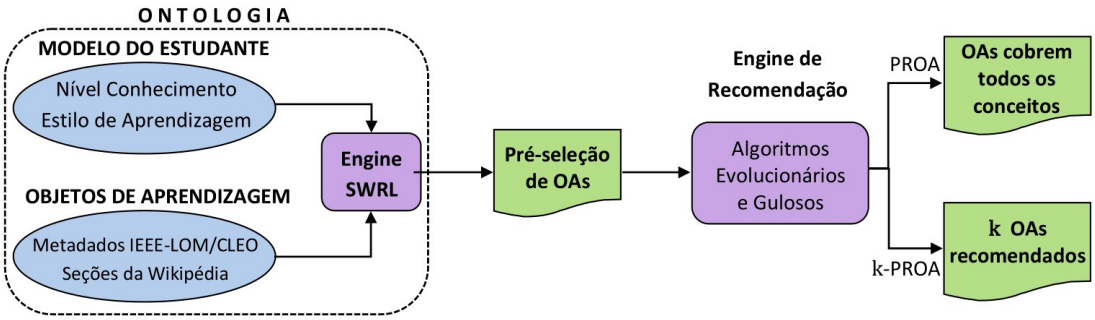

Figura 1. Modelo de Recomendação de OAs na abordagem proposta

1. Encontrar um conjunto de OAs que juntos cobrem todos os conceitos (elementos de X) a um custo mínimo;
2. Encontrar um conjunto de k OAs que maximizam o número de conceitos cobertos a um custo mínimo.

O primeiro objetivo equivale ao PCC apresentado na Eq. (1) e o segundo objetivo é similar ao PMC. Para o primeiro objetivo tem-se a solução $\{O_2, O_3, O_4\}$ com custo 7.

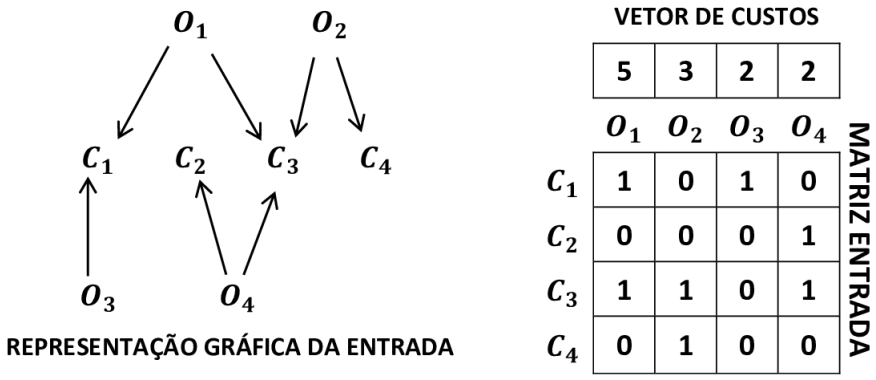

Figura 2. Matriz e vetor de custos da entrada

No processo de busca, os conceitos e os parâmetros de busca da entrada do usuário são utilizados pelas regras SWRL na pré-seleção de objetos de aprendizagem. Os OAs pré-selecionados cobrem ao menos um conceito da entrada. É gerada uma matriz binária a_{ij} (ver Fig. 2) na qual cada linha corresponde a um conceito da entrada e cada coluna está associada a um OA resultante da pré-seleção. Tem-se $a_{\text{ij}} = 1$, se a coluna j (O_{j}) cobre a linha i (C_{i}), e $a_{\text{ij}} = 0$, caso contrário.

A cada O_{j} da matriz é atribuído um custo c_{j} calculado pela dissimilaridade entre os parâmetros de busca do usuário (O_{ideal}) e o O_{j}. Tais parâmetros são detalhados por [Belizário Júnior and Dorça 2018]. As características de cada OA, tais como o tipo (exercício, vídeo etc) e as palavras-chave, são comparadas com os parâmetros de busca equivalentes configurados nos metadados do O_{ideal} da ontologia. Para cada objetivo há uma maneira de definir o PROA. As definições formais para o problema são apresentadas nas seções a seguir.

3.1. O Problema de Cobertura de Conjuntos reformulado

[Bilal et al. 2013] reformularam o Problema de Cobertura de Conjuntos para que a função objetivo contivesse a principal meta do problema. Dada uma matriz binária a_{ij} com m

linhas e n colunas, o objetivo é cobrir todas as linhas usando um subconjunto de colunas com custo mínimo. Para cada linha i coberta é associado um ganho g_{i} com um possível custo. Como todas as linhas devem ser cobertas, o ganho atribuído a cada linha deve ser maior que o custo de ao menos uma das colunas que cobrem a linha. Caso contrário, não há nenhum benefício em cobrir essa linha. Seja $c_{\text{min}}(l_{\text{i}})$ o custo da coluna de menor custo dentre as colunas que cobrem a linha l_{i}. O ganho $g_{\text{i}} = c_{\text{min}}(l_{\text{i}}) + \epsilon$ é atribuído a cada linha l_{i}, sendo ϵ uma constante positiva pequena. Sejam x_{j} = 1, se a coluna *j* (com custo $c_{\text{j}} > 0$) está na solução e $x_{\text{j}} = 0$, caso contrário. Sejam y_{i} = 1 se a linha l_{i} (com ganho $g_{\text{i}} > 0$) é coberta na solução e y_{i} = 0, caso contrário, então o PCC é reformulado como:

$$Maximize \quad \sum_{i=1}^{m} g_{\text{i}} y_{\text{i}} - \sum_{j=1}^{n} c_{\text{j}} x_{\text{j}} \tag{1}$$

$$Sujeito\ a \quad y_{\text{i}} \leq \sum_{j=1}^{n} a_{\text{ij}} x_{\text{j}}, \quad i = 1, ..., m \tag{2}$$

$$x_{\text{j}}, y_{\text{i}} \in \{0, 1\}, \quad i = 1, ..., m; \quad j = 1, ..., n \tag{3}$$

3.2. O PROA definido como o PCC

O PROA, cujo objetivo é encontrar uma cobertura de OAs que cobre todos os conceitos a um custo mínimo, corresponde ao PCC formalizado pela Eq. (1). No cálculo do custo c_{j}, considera-se o grau de dissimilaridade entre o O_{ideal} e o O_{j}, bem como as filtragens colaborativa e baseada em conhecimento:

$$c_{\text{j}} = diss(O_{\text{ideal}}, O_{\text{j}}) + dissContext(E_{\text{alvo}}, O_{\text{j}}) \tag{4}$$

em que $diss(O_{\text{ideal}}, O_{\text{j}})$ é inversamente proporcional ao grau de similaridade entre o O_{ideal} e o O_{j} e $dissContext(E_{\text{alvo}}, O_{\text{j}})$ é inversamente proporcional à relevância que o O_{j} tem para o estudante-alvo E_{alvo}, para o qual é direcionada a recomendação dos OAs. Essa relevância é medida por técnicas de filtragem colaborativa e baseada em conhecimento.

O resultado de $diss(O_{\text{ideal}}, O_{\text{j}})$ depende da proximidade entre o O_{ideal} e o O_{j}. Os parâmetros do O_{j}, tais como grau de dificuldade, densidade semântica e tipo de recurso de aprendizagem, são comparados com os parâmetros correspondes do O_{ideal}. Os valores atribuídos aos parâmetros do O_{ideal} são dados pelo usuário.

Formalmente, seja α_{i} o valor do parâmetro i. O cálculo da dissimilaridade entre o O_{j} e os parâmetros de busca do usuário é dado pela equação 5:

$$diss(O_{\text{ideal}}, O_{\text{j}}) = \sum_{i=1}^{k} (\alpha_{\text{i(ideal)}} - \alpha_{\text{i(j)}}) \tag{5}$$

em que k é o número de parâmetros, $\alpha_{\text{i(ideal)}}$ é o valor do parâmetro i do O_{ideal} e $\alpha_{\text{i(j)}}$ é o valor do parâmetro i do O_{j}.

Em [Belizário Júnior and Dorça 2018], considerou-se que $c_{\text{j}} = diss(O_{\text{ideal}}, O_{\text{j}})$. Neste trabalho, estendemos a definição inicial para que o custo atribuído ao objeto de

aprendizagem contemple também o contexto do estudante em uma abordagem colaborativa. Tem-se:

$$dissContext(E_{\text{alvo}}, O_{\text{j}}) = 1 - P_{\text{E}_{\text{alvo}},\,\text{O}_{\text{j}}} \tag{6}$$

onde $\text{P}_{\text{E}_{\text{alvo}},\,\text{O}_{\text{j}}}$ é a predição da nota que o estudante-alvo daria para o novo O_{j}. Esse valor representa a importância que o OA tem para o estudante e é dado num intervalo real [0,1], em que quanto maior o seu valor maior a importância que o OA tem para o estudante. Essa predição é calculada pela abordagem kNN proposta por [Tarus et al. 2018], a qual encontra os k estudantes, dentre os que avaliaram o recurso O_{j}, mais similares ao estudante-alvo. A meta é predizer a nota que o estudante-alvo daria para o O_{j} usando as avaliações que O_{j} recebeu de outros estudantes similares (vizinhos mais próximos).

3.3. O k-PROA similar ao PMC

O segundo objetivo do PROA similar ao PMC é encontrar uma cobertura de k objetos de aprendizagem que maximiza o número de conceitos cobertos. Nesse problema, doravante k-PROA, não há a garantia de que k é suficientemente grande para cobrir todos os conceitos, então algum conceito C não coberto pode ser pré-requisito para um conceito C' coberto. Nesse caso, há uma relação de pré-requisito não satisfeita. Assim, há um impasse entre maximizar o número de conceitos cobertos e satisfazer todas as relações de pré-requisitos que precisa ser considerado no objetivo do problema.

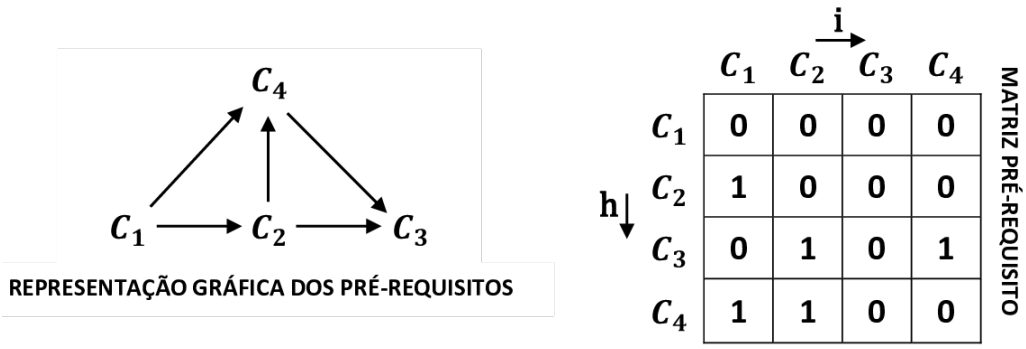

Figura 3. Matriz de pré-requisitos da entrada

Na Fig. 3, por exemplo, se uma solução cobre apenas os conceitos C_3 e C_4, então há três relações de pré-requisitos não satisfeitas, quais sejam: $C_2 \rightarrow C_3$, $C_1 \rightarrow C_4$ e $C_2 \rightarrow C_4$. Para cada relação de pré-requisito não satisfeita gera-se uma penalidade ponderada por uma constante positiva β. Formalmente, dada uma matriz binária a_{ij} com m linhas e n colunas e uma matriz de pré-requisitos binária b_{hi} de ordem m, o objetivo é encontrar uma cobertura de k OAs que maximiza o número de conceitos cobertos respeitando um número máximo de pré-requisitos não satisfeitos a um custo mínimo. Sejam $a_{\text{ij}} = 1$, se a linha i é coberta pela coluna j e $a_{\text{ij}} = 0$, caso contrário, $b_{\text{hi}} = 1$, se o conceito i é pré-requisito do conceito h e $b_{\text{hi}} = 0$, caso contrário, x_{j} = 1, se a coluna *j* (com custo $c_{\text{j}} > 0$, correspondente ao c_{j} da Seção 3.2) está na solução e $x_{\text{j}} = 0$, caso contrário. Sejam y_{i} = 1 se a linha i é coberta na solução e y_{i} = 0, caso contrário, então o k-PROA é formulado como:

$$Maximize \quad \sum_{i=1}^{m} y_{\mathrm{i}} - \beta \sum_{h=1}^{m} \sum_{i=1}^{m} (1 - y_{\mathrm{i}}) b_{\mathrm{hi}} y_{\mathrm{h}} - \frac{\sum_{j=1}^{n} c_{\mathrm{j}} x_{\mathrm{j}}}{\sum_{j=1}^{n} c_{\mathrm{j}}} \tag{7}$$

$$Sujeito\ a \quad y_{\mathrm{i}} \leq \sum_{j=1}^{n} a_{\mathrm{ij}} x_{\mathrm{j}}, \quad i = 1, ..., m \tag{8}$$

$$\sum_{j=1}^{n} x_{\mathrm{j}} = k, \quad x_{\mathrm{j}}, y_{\mathrm{i}} \in \{0, 1\}, \quad i = 1, ..., m; \quad j = 1, ..., n \tag{9}$$

4. Experimentos e Análise dos Resultados

Os algoritmos Jpso-Scp e CWSC foram adaptados para a resolução do PROA. No Jpso-Scp, a função objetivo do PCC foi substituída pela função objetivo do PCC reformulado, Eq. (1), para a avaliação das partículas. Utilizou-se 10 partículas para a execução dos testes. Em geral, o Jpso-Scp converge nas primeiras iterações.

O algoritmo CWSC foi adaptado para a solução do PROA e do k-PROA. Para resolver o k-PROA, utiliza-se $k = 80\%$ do número de colunas mínimo necessário para cobrir todas as linhas. Neste caso, a meta é cobrir o maior número possível de conceitos usando 80% das colunas necessárias para cobrir todos eles. No caso do PROA associado ao PCC reformulado, a meta é cobrir todos os conceitos e, para isso, usa-se um número de colunas no máximo igual ao número de linhas (m) da matriz de entrada. Assim, para a solução do PROA, o CWSC é configurado para $s = 1$ e $k = m$.

Esses algoritmos foram implementados em Python e os testes foram realizados em um notebook com 8GB de RAM e processador AMD Quad-Core com 2.40GHz. O *dataset* possui 24 instâncias distribuídas em 6 grupos com 4 instâncias cada, as quais possuem 100, 500, 2000 e 10000 colunas (OAs). O número de linhas (conceitos) da instância depende do índice do grupo. Por exemplo, as instâncias do grupo $C3_{10}$ possuem 10 linhas. A densidade corresponde à porcentagem de colunas que cobre cada linha da instância. As instâncias dos grupos $C1_2$ e $C2_6$ possuem densidade igual a 50%, as do grupo $C3_{10}$ têm densidade igual a 20% e as instâncias dos grupos $C4_{25}$, $C5_{40}$ e $C6_{55}$ possuem densidade igual a 10%. A cada coluna atribuiu-se um custo aleatório para simular um cenário real. O custo é um valor inteiro que varia de 1 até 10% do número de colunas da instância. Por exemplo, atribuiu-se aleatoriamente um custo no intervalo [1,50] para cada coluna das instâncias com 500 colunas. Em cada instância, um mesmo custo é atribuído a exatamente 10 colunas.

O Erro Médio Absoluto (EMA) é calculado para o PROA como a diferença entre o custo da solução exata e o custo resultante dos algoritmos CWSC e Jpso-Scp adaptados. O EMA das instâncias resolvidas é apresentado na Tabela 2. Na solução do k-PROA, a meta é cobrir o maior número de conceitos possível satisfazendo as relações de pré-requisitos. Para os testes, tem-se $\beta = 1$ e geraram-se 6 matrizes de pré-requisitos, uma para cada grupo do dataset. Cada conceito possui em média 2 pré-requisitos. Para a avaliação dos resultados, considera-se a solução igual ao número de conceitos cobertos menos o número de pré-requisitos não satisfeitos. Esse resultado é comparado com a solução exata para o cálculo do EMA na Tabela 3.

Tabela 2. Comparação entre a acurácia do CWSC e do JPSO na solução do PROA

	Alg.	$C1_{\mathbf{2}}$	$C2_{\mathbf{6}}$	$C3_{\mathbf{10}}$	$C4_{\mathbf{25}}$	$C5_{\mathbf{40}}$	$C6_{\mathbf{55}}$	**EMA**
100	JPSO	0,00	0,00	0,00	0,00	0,56	0,00	0,09
	CWSC	0,00	0,00	3,00	0,00	4,00	2,00	1,50
500	JPSO	0,00	0,00	0,00	0,00	0,00	0,00	0,00
	CWSC	0,00	0,00	0,00	4,00	3,00	1,00	1,33
2000	JPSO	0,00	0,00	0,00	0,60	0,16	0,30	0,18
	CWSC	0,00	0,00	0,00	3,00	6,00	7,00	2,67
10000	JPSO	0,00	0,00	0,00	0,00	0,00	1,00	0,17
	CWSC	0,00	0,00	0,00	2,00	1,00	7,00	1,67

Tabela 3. Acurácia do CWSC em relação ao EXATO na solução do k-PROA

	$C1_{\mathbf{2}}$	$C2_{\mathbf{6}}$	$C3_{\mathbf{10}}$	$C4_{\mathbf{25}}$	$C5_{\mathbf{40}}$	$C6_{\mathbf{55}}$	**EMA**
100	0	0	0	4	6	11	3,50
500	0	0	1	3	3	7	2,33
2000	0	0	0	5	8	19	5,33
10000	0	0	0	2	10	17	4,83

O Jpso-Scp adaptado apresenta soluções melhores que a meta-heurística gulosa do CWSC, porém o CWSC é bem mais rápido que o Jpso-Scp. Nos testes realizados, o menor tempo de execução do Jpso-Scp é 0,15s e o maior é 1035s, enquanto que para o CWSC o menor é 0,00053s e o maior é 2,8s. O CWSC também é rápido para a solução do k-PROA e apresenta resultados ótimos para instâncias com até 10 conceitos. Os resultados mostram que ambos os algoritmos são bons candidatos para a recomendação de OAs que atendem ao contexto de cada estudante (estilo de aprendizagem e conceitos a serem aprendidos). Essa recomendação, por ser baseada em conhecimento, ameniza os problemas *cold-start* e de esparsidade de dados.

5. Considerações Finais e Trabalhos Futuros

Neste artigo, apresentou-se uma abordagem para a solução do PROA e do k-PROA que contribui com o estado da arte usando um algoritmo guloso e o JPSO. O modelo de recomendação utiliza-se de uma ontologia para modelar o conhecimento sobre os estudantes e os OAs, além de regras SWRL que pré-selecionam os OAs.

Em trabalhos futuros, visa-se encontrar maneiras de combinar a eficiência em tempo da meta-heurística gulosa com os resultados ótimos do JPSO, bem como testar a influência da quantidade de pré-requisitos entre os conceitos na qualidade dos algoritmos. Os resultados experimentais confirmam a viabilidade do modelo para ser implementado em um cenário real. Visa-se a integração desse modelo em um STI para automatizar o processo de aprendizagem diminuindo a intervenção do professor.

Agradecimentos

Os autores agradecem à CAPES e à Pró-Reitoria de Pesquisa e Pós-Graduação por todo o suporte prestado no desenvolvimento deste trabalho.

Referências

Balaji, S. and Revathi, N. (2016). A new approach for solving set covering problem using jumping particle swarm optimization method. *Natural Computing*, 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 *SBIE*, volume 29, page 1533.

Bilal, N., Galinier, P., and Guibault, F. (2013). A new formulation of the set covering problem for metaheuristic approaches. *ISRN Operations Research*, 2013.

CLEOLab (2003). Cleo extensions to the ieee learning object metadata. *CLEO Collaborative Partners*, pages 1–65.

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 2161, pages 49–53. IEEE.

Garcia, F. J. M. and Pérez, J. A. M. (2008). Jumping frogs optimization: a new swarm method for discrete optimization. *Documentos de Trabajo del DEIOC*, 3.

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 *2015 IEEE 31st Intern. Conf. on Data Engineering*, pages 879–890. IEEE.

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 *Proceedings...*, Washington, USA. ITHET, IEEE Computer Society.

McGregor, A. and Vu, H. T. (2019). Better streaming algorithms for the maximum coverage problem. *Theory of Computing Systems*, 63(7):1595–1619.

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, 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.

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.