

COLÔNIAS DE FORMIGAS ORIENTANDO O APRENDIZADO DE LÍNGUAS

Paulo de Almeida Afonso¹, Paulo R. Ferreira Jr.¹, Rafael Vetromille-Castro¹

¹Universidade Federal de Pelotas (UFPEL)

Rua Gomes Carneiro, 1 - Centro - CEP 96010-610 / Pelotas, RS - Brazil

{paulo.afonso, paulo.ferreira, rafael.vetromille}@ufpel.edu.br

Abstract. *This paper proposes the application of swarm intelligence to recommend language learning objects. Through the literature review, it can be seen that the existing approaches are focused to recommend general learning objects without considering the aspects that guide the learning of languages. As a solution to this problem, some essential elements should be considered during the recommendation process: the ability to combine an object with another and its reusability. In this context, the recommendation task presented here considers the creation of multiple conceptual connections between the objects and occurs through the application of the algorithm of Optimization by Ants Colony.*

Resumo. *Este artigo propõe a aplicação de técnicas de inteligência de enxames para a recomendação de objetos de aprendizagem de línguas. Através da revisão da literatura percebe-se que as abordagens existentes são voltadas para a recomendação de objetos de aprendizagem de modo genérico, sem considerar os aspectos que norteiam o aprendizado de línguas. Como solução para esse problema, alguns elementos essenciais devem ser considerados durante o processo de recomendação: a capacidade de combinar um objeto com outro e sua reusabilidade. Nesse contexto, a tarefa de recomendação aqui apresentada prevê a existência de múltiplas conexões conceituais entre os objetos e ocorre através da aplicação do algoritmo de Otimização por Colônia de Formigas.*

1. Introdução

A utilização de recursos online tem se tornado cada vez mais presente na educação, quer seja no âmbito do ensino a distância ou presencial. Diante disso, novas formas de ensino e aprendizagem tem surgido, assim como diferentes possibilidades de desenvolvimento de conteúdos por meio de Objetos de Aprendizagem (OA) [Cazella et al. 2012]. Compreendidos como entidades digitais ou recursos educacionais, podem ser acessados via internet, compartilhados, modificados e recombinados com outros, formando módulos independentes de instrução para serem utilizados on-line ou off-line [Soares 2009].

Apesar da grande disponibilidade desses recursos educacionais, estudos recentes tem demonstrado que seu desenvolvimento encontra-se mais direcionado aos aspectos tecnológicos do que à aspectos pedagógicos. Dessa forma, para o ensino e aprendizagem de línguas, faz-se necessário um embasamento teórico diferenciado, adequado à aprendizagem de línguas e com ênfase na comunicação [Vetromille-Castro et al. 2013]. Nesse contexto, surgem os Objetos de Aprendizagem de Línguas (OAL), um OA preparado para

o ensino e aprendizagem de línguas, cuja definição deve ser norteada por aspectos de design de usabilidade e usabilidade pedagógica [Vetromille-Castro et al. 2012]. Deve ter como objetivo principal a capacidade de facilitar a integração de competências gramatical, sociolinguística e estratégica, características particularmente ligadas ao ensino e à aprendizagem de idiomas.

No melhor de nosso conhecimento, não se pode encontrar na literatura trabalhos específicos relacionados à recomendação de OAL. Percebe-se que as abordagens existentes são voltadas para a recomendação de OA de modo genérico, sem considerar os aspectos que norteiam o aprendizado de línguas. Nesse sentido, foi realizado um estudo sobre os elementos essenciais que definem um OAL e as principais abordagens utilizadas em Sistemas de Recomendação (SR) de uso geral, objetivando, a partir do conhecimento adquirido, propor uma estratégia para recomendação de OAL. O modelo proposto constituiu-se na combinação de técnicas colaborativas e baseadas em conteúdo em um único sistema de recomendação híbrido, baseado em regras. A recomendação propriamente dita se dá através da aplicação do Algoritmo de Otimização por Colônia de Formigas.

O presente trabalho segmenta-se em cinco partes distintas e complementares, além desta introdução: O capítulo 2 apresenta os trabalhos relacionados; No capítulo 3 são apresentados os conceitos relacionados ao algoritmo ACO; O capítulo 4 descreve a abordagem utilizada; No capítulo 5 são apresentados os experimentos realizados e a interpretação dos resultados obtidos e, no capítulo 6 são apresentadas as conclusões.

2. Trabalhos Relacionados

Para análise da viabilidade de aplicação da solução proposta, foram estudados, além dos trabalhos ligados ao contexto de recomendação de OA, abordagens relacionadas à recomendação de objetos em diferentes contextos, visto que as técnicas adotadas por esses trabalhos podem contribuir para o desenvolvimento da solução proposta.

A exemplo do exposto, temos o trabalho de [Prabha and Rathipriya 2013], onde é apresentado um SR com base em um modelo de navegação. A recomendação ocorre a partir da análise do perfil de acesso dos usuários, obtido de forma implícita, utilizando para tanto um método baseado em inteligência de enxames. Nessa mesma linha, [Bedi and Sharma 2012] e [Bellaachia and Alathel 2016], abordam técnicas que utilizam a confiança existente entre os usuários para gerar recomendações.

Ainda na área de inteligência de enxames, [Wang et al. 2012] e [Sobecki and Tomczak 2010], apresentam um novo tipo de esquema de recomendação, com ênfase em filtragem colaborativa, baseada no comportamento das formigas. A técnica utiliza um mecanismo de transmissão de feromônio para identificar a similaridade entre usuários e objetos.

Diferente da proposta deste trabalho, as abordagens analisadas não exploram uma relação conceitual entre os objetos a serem recomendados. O mesmo ocorre com os trabalhos de [Cazella et al. 2009] e [Zaina et al. 2012], onde são apresentadas técnicas de recomendação baseadas em competências a serem desenvolvidas por alunos e suas preferências. Além disso, as recomendações são realizadas mediante a descoberta de interesses, obtida de forma explícita, seja pela atribuição de classificações por parte dos usuários ou análise de perfis pré-cadastrados.

3. Fundamentação Teórica

Ant Colony Optimization - ACO é um termo de uso geral utilizado para determinar uma classe de procedimentos metaheurísticos fundamentados no comportamento de formigas [Serapião 2009]. Em um algoritmo ACO os agentes se comunicam indiretamente através da estigmergia, cooperando na tentativa de encontrar soluções. Esta comunicação indireta, mediada pelo ambiente, pode levar a boas soluções, que são uma propriedade emergente da interação entre os agentes.

De maneira informal um algoritmo ACO pode ser representado através da interação de três procedimentos básicos [Dorigo et al. 2008]:

- Construir soluções: método aleatorizado para obtenção da população inicial;
- Aplicar busca local (opcional): utilizado para melhorar a solução através da aplicação de busca local com base em soluções obtidas por um ou mais agentes;
- Atualizar feromônio: considera o índice de evaporação.

O problema pode ser representado através da simulação de um número de formigas artificiais que se deslocam em um grafo. Elas percorrem o grafo construindo a solução, passando pelos vértices e depositando feromônio nas arestas. (Figura 1).

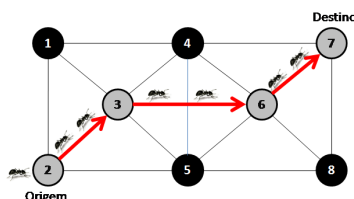


Figura 1. ACO - Exemplo de construção de uma solução por uma formiga.

Cada aresta (i, j) do grafo possui uma variável τ_{ij} que representa uma trilha de feromônio artificial. Essa variável pode ser modificada pelas formigas, sendo incrementada de acordo com o percurso realizado por elas. Assim, quanto maior o nível de feromônio, maior a chance da aresta ser visitada por outras formigas.

Conforme o exemplo apresentado pela figura 1, dado um grafo $G = (i, j)$, o caminho mais curto entre um determinado par de vértices pode ser encontrado. Nesse exemplo, a formiga percorre o grafo partindo inicialmente do vértice 2 e chegando até o vértice 7, obtendo como uma das possíveis soluções o caminho 2-3-6-7.

Em cada passo da construção da solução, a formiga seleciona o próximo vértice a ser visitado de acordo com um mecanismo estocástico que é influenciado pelo feromônio: quando no vértice i , o seguinte vértice é selecionado estocasticamente entre os anteriormente visitados (Figura: 2).

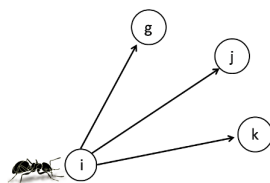


Figura 2. ACO - Seleção do próximo caminho - processo estocástico.

Fonte: [Dorigo et al. 2006]

Se j ainda não foi visitado, ele pode ser selecionado, com uma probabilidade que é proporcional ao índice de feromônio associado à aresta (i, j) .

Através de uma fórmula probabilística, cada formiga k constrói uma solução movendo-se por uma sequência de locais vizinhos. Ao final de cada iteração o feromônio é atualizado tomando-se por base a qualidade das soluções produzidas pelas formigas.

A equação é dada por:

$$p_{ij}^k = \frac{[\tau_{ij}]^\alpha \cdot [\eta_{ij}]^\beta}{\sum_{l \in J_i^k} [\tau_{il}]^\alpha \cdot [\eta_{il}]^\beta} \quad (1)$$

onde:

p_{ij}^k : é a probabilidade da formiga k , que se encontra no vértice i , escolher o vértice j como o próximo a ser visitado;

τ_{ij} : quantidade de feromônio existente na aresta (i, j) . Inicialmente, adota-se um mesmo valor τ_0 para todas as arestas do grafo;

η_{ij} : heurística que representa a atratividade da aresta (i, j) ;

$l \in J_i^k$: conjunto de vértices ainda não visitados pela formiga k , que se encontra no vértice i ;

α : parâmetro utilizado para determinar a importância da trilha de feromônio τ_{ij} ;

β : parâmetro que pondera a influência relativa da variável η_{ij} para os vértices i e j no processo de decisão;

Para que não haja uma intensificação demasiadamente elevada do nível de feromônio, a cada iteração da fase de construção as formigas atualizam o feromônio da aresta pela qual acabaram de atravessar, de acordo com a expressão:

$$\tau_{ij} = (1 - \rho) \cdot \tau_{ij} + \rho \cdot \tau_0 \quad (2)$$

O parâmetro ρ é utilizado para que os caminhos menos frequentados sejam esquecidos com o passar do tempo. ρ é um parâmetro definido entre 0 e 1 e τ_0 é o valor inicial do feromônio das arestas.

4. Abordagem Proposta

A proposta do presente trabalho fundamenta-se na recomendação de Objetos de Aprendizagem de Línguas, classificados segundo sua relevância para o usuário. Dessa forma, a saída desejada consiste em uma lista de "n" OAL, ordenados de acordo a predição realizada pelo sistema, ficando a critério do usuário selecionar ou não aquele que melhor atende ao objetivo da sua consulta.

O modelo proposto incorpora tanto características baseadas em conteúdo e colaborativas, combinando ambas as técnicas em único recomendador baseado em regras, constituindo dessa forma um SR híbrido. A recomendação propriamente dita considera as ações do usuário, o valor e o número de *downloads* associado a cada objeto, a intensidade

das conexões existentes entre os possíveis objetos a serem recomendados e ocorre através da aplicação do algoritmo de Otimização por Colônia de Formigas.

Através da utilização da medida TF-IDF (*Term Frequency-Inverse Document Frequency*), uma técnica bastante utilizada para indexação de frequência de termos, é possível fazer uma recomendação inicial a partir da consulta de um usuário por uma determinada palavra-chave. Além disso, tendo em vista que a solução proposta não prevê a atribuição de avaliações explícitas para os objetos recomendados, o valor da medida TF-IDF é associado a cada objeto e utilizado para o cálculo do índice de feromônio, podendo aumentar ou diminuir de acordo com as ações realizadas pelo usuário.

A descrição de interesses do usuário ocorre por meio de ações como busca, visualização e seleção de OAL. Essas ações são utilizadas no contexto de aplicação da filtragem colaborativa, contribuindo, de forma implícita, para o estabelecimento de conexões entre os objetos.

A Figura 3 apresenta o exemplo da construção de uma solução representada através de um grafo. Nesse exemplo cada vértice do grafo representa um OAL e cada aresta representa uma conexão entre dois OAL.

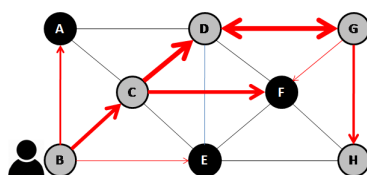


Figura 3. ACO - Exemplo de construção de uma solução por um usuário.

Cada aresta (i, j) do grafo possui uma variável τ_{ij} que representa uma trilha de feromônio artificial. Essa variável pode ser modificada pelo usuário, sendo incrementada de acordo com o percurso realizado por ele.

Através de uma fórmula probabilística, possíveis soluções são apresentadas para o usuário a cada nova interação. Em cada passo da construção da solução, o usuário seleciona o próximo vértice a ser visitado, através de uma lista, ordenada de acordo com um mecanismo estocástico que é influenciado pelo valor da variável τ_{ij} .

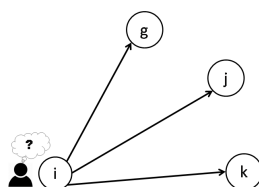


Figura 4. ACO - Seleção do próximo caminho pelo usuário.

De forma análoga às soluções construídas pelas formigas, o usuário percorre o grafo construindo a solução, passando pelos vértices e pelas arestas, modificando o valor da variável τ_{ij} associada a cada aresta. Ao final de cada iteração o feromônio é atualizado, podendo ser incrementado ou decrementado, conforme a ação realizada pelo usuário. Quando o nível de feromônio é decrementado a conexão entre os objetos é enfraquecida,

podendo até mesmo ser desfeita ao longo do tempo. Dessa forma, quanto maior o nível de feromônio, maior a chance da aresta ser visitada por outros usuários. Essa característica é evidenciada pela figura 5, onde pode-se perceber que a conexão entre os vértices $B \rightarrow C$ é mais forte que $B \rightarrow A$ e $B \rightarrow E$, sendo denotada pela espessura da aresta.

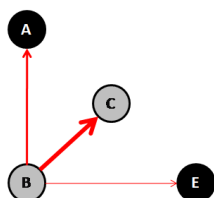


Figura 5. ACO - Representação do relacionamento entre OALs.

A figura 6 demonstra as etapas percorridas pelo usuário para construção de uma solução. Tomando-se como exemplo o grafo representado pela figura 3, ao final da interação do usuário tem-se a solução dada pelo caminho B-C-D-G-H (figura 6e).

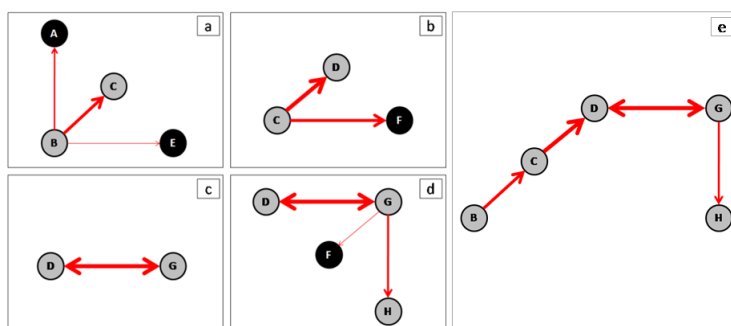


Figura 6. ACO - Construção de uma solução.

O exemplo apresentado considera que o usuário selecionou o primeiro objeto recomendado pelo SR em todas as interações. Cabe ressaltar que outros objetos do grafo poderiam ser selecionados, dando origem à novas soluções, através da criação de novas conexões ou reforçando conexões já existentes. Os passos para construção da solução representada pela figura 6e são descritos a seguir.

1. O usuário faz uma busca por palavra-chave;
2. O SR retorna um conjunto de objetos, ordenados em uma lista, de acordo com a importância do objeto para a consulta realizada, tomando-se por base a medida TF-IDF associada a cada objeto;
3. Ao clicar no objeto B o SR retorna uma nova lista, ordenada de acordo com o valor da variável p_{ij}^k ; (objetos C , A e E , respectivamente (figura 6a));
4. Ao clicar no objeto C são apresentados os objetos D e F (figura 6b);
5. Ao clicar no objeto D é apresentado apenas o objeto G , pois não existem outros objetos relacionados. A figura 6c denota a existência de uma forte conexão entre dois objetos, podendo um referenciar o outro no processo de recomendação.
6. Por fim, ao clicar no objeto G , são apresentados os objetos H e F , respectivamente, resultando na solução representada pela figura 6e.

O usuário constrói possíveis soluções a cada interação com o sistema, reforçando ou enfraquecendo a conexão entre os objetos de acordo com suas preferências.

A equação 1 é reescrita logo a seguir para melhor compreensão em relação aos valores associados às variáveis utilizadas por esta solução.

$$p_{ij}^k = \frac{[\tau_{ij}]^\alpha \cdot [\eta_{ij}]^\beta}{\sum_{l \in J_i^k} [\tau_{il}]^\alpha \cdot [\eta_{il}]^\beta}$$

Onde:

p_{ij}^k : é o valor utilizado para ordenar os objetos recomendados, segundo a predição calculada pelo sistema. Se j ainda não foi visitado, ele pode ser selecionado, com uma probabilidade que é proporcional ao índice de feromônio associado à aresta (i, j) .

τ_{ij} : variável que controla o feromônio depositado nas arestas. Na solução proposta o valor utilizado para representar o índice de feromônio associado à variável τ_{ij} é a medida TF-IDF. A aplicação da atualização de feromônio é executada a cada interação do usuário com um OAL. Se o usuário visualizou o objeto e não efetuou o download o feromônio é decrementado de acordo com a equação 2. Caso contrário, o feromônio é incrementado de acordo com a medida TF-IDF correspondente ao OAL selecionado.

η_{ij} : é um valor heurístico associado ao contexto do problema em questão, sendo representado neste trabalho pelo número de visitas ou interações, realizadas por usuários distintos, à uma conexão existente entre determinados objetos.

Três parâmetros α , β e ρ são utilizados para controlar a intensidade do feromônio (τ_{ij}), e a qualidade da aresta (η_{ij}):

- α - Utilizado para determinar a importância do feromônio. Um valor muito alto para α pode causar um efeito de intensificação do feromônio.
- β - Trata-se de um parâmetro importante para uma boa diversidade de soluções. Um valor baixo demais para β pode levar à estagnação precoce do algoritmo, assim como um valor alto demais pode aproximá-lo de uma construção gulosa.
- ρ - Utilizado na equação 2, responsável pela evaporação do feromônio. Neste trabalho a evaporação é aplicada sempre que o usuário visualiza um objeto recomendado e não efetua o download. Dessa forma, conexões que não são consideradas relevantes por muito usuários, vão perdendo a intensidade ao longo do tempo, podendo até mesmo serem desfeitas.

Nos experimentos realizados para o TSP, verificou-se que diferentes combinações de parâmetros (ou seja, $(\alpha = 1, \beta = 1)$, $(\alpha = 1, \beta = 2)$, $(\alpha = 1, \beta = 5)$, $(\alpha = 0.5, \beta = 5)$) resultaram no mesmo nível de desempenho [Dorigo et al. 1996]. Assim, os valores utilizados na abordagem proposta foram aqueles sugeridos pelo autor para o algoritmo Ant cycle, que obteve melhor desempenho em comparação aos outros dois: Ant density e Ant quantity. Sendo $\alpha = 1$, $\beta = 2$ e $\rho = 0.5$.

5. Experimentos e Resultados

Para validação do sistema proposto foram cadastrados e disponibilizados por meio de um protótipo 665 objetos de aprendizagem (OAs). Os experimentos foram realizados através da simulação de 10 usuários pré-cadastrados no sistema e contaram com a participação de um professor especialista da área de Letras, e três professores de inglês em formação vinculados ao projeto de um Programa de Pós-Graduação em Letras.

As três etapas de simulação resultaram na criação de 72 conexões entre objetos. Essas conexões, quando representadas através de um grafo, assemelham-se aos caminhos construídos pelas formigas durante o forrageamento. A cada interação podem surgir novas conexões, ocorrer o reforço ou serem desfeitas conexões já existentes. A intensidade das conexões, resultantes dos experimentos, evidencia essa característica.

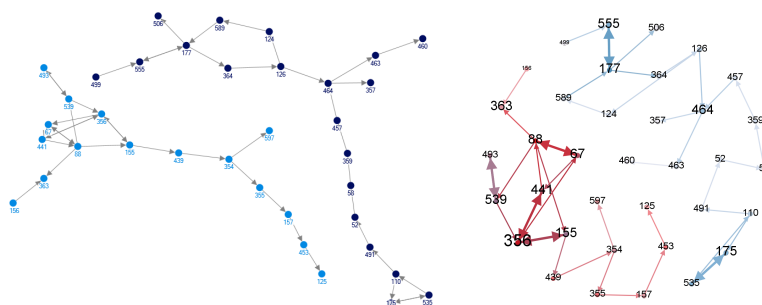


Figura 7. Exemplo de conexões resultantes dos experimentos

Como forma de validação foram implementados, além do algoritmo proposto, dois dos algoritmos de recomendação do Apache Mahout, amplamente adotados em Sistemas de Recomendação. Os resultados obtidos podem ser avaliados conforme descrito a seguir:

Tabela 1. Classificação do resultado da recomendação de um item para um usuário

	Recomendado	Não recomendado
Selecionado	Verdadeiro-Positivo (vp)	Falso-Negativo (fn)
Não selecionado	Falso-Positivo (fp)	Verdadeiro-Negativo (vn)

A tabela 2 apresenta os valores resultantes da análise das recomendações realizadas pelos três algoritmos implementados.

Tabela 2. Valores obtidos a partir da análise das recomendações

	vp	fp	vn	fn
Swarm	14	6	637	8
Mahout (baseada em item)	2	6	637	20
Mahout (baseada em usuário)	12	8	635	10

Com base nos valores da tabela 2 podemos calcular as seguintes medidas:

Precision: Quantidade de recomendações adequadas para o usuário alvo.

Recall: Proporção de recomendações corretamente realizadas.

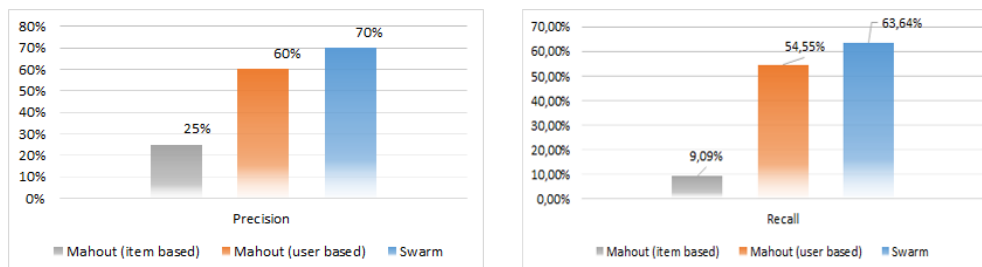


Figura 8. Percentuais Precison e Recall

False Positive Rate: A taxa de falsos positivos descreve a proporção de recomendações realizadas pelo sistema que foram inadequadas para usuário alvo.

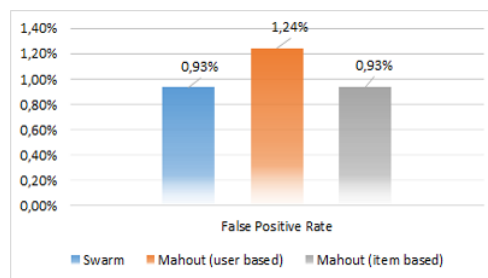


Figura 9. Percentual de recomendações inadequadas para os usuários I, J e K.

6. Conclusões

Este trabalho apresentou uma proposta para aplicação de técnicas de inteligência de enxames para a recomendação de Objetos de Aprendizagem de Línguas. Os resultados obtidos a partir dos experimentos demonstraram que a abordagem utilizada mostra-se eficiente para o contexto de aplicação a que se propõe, apresentando resultados melhores que outros dois algoritmos, amplamente utilizados em Sistemas de Recomendação. Além disso, dois aspectos fundamentais a serem considerados durante o processo de recomendação foram contemplados, quais sejam: a capacidade de combinar um objeto com outro e a reusabilidade, isto é, a possibilidade de sua aplicação em diferentes contextos de aprendizagem. Para tanto, a tarefa de recomendação prevê a existência de múltiplas conexões conceituais entre os objetos, que são construídas à medida que o sistema é utilizado, e ocorre através da aplicação do algoritmo de Otimização por Colônia de Formigas.

Referências

- Bedi, P. and Sharma, R. (2012). Trust based recommender system using ant colony for trust computation. *Expert Systems with Applications*, 39(1):1183–1190.
- Bellaachia, A. and Alathel, D. (2016). Improving the recommendation accuracy for cold start users in trust-based recommender systems. *International Journal of Computer and Communication Engineering*, 5(3):206.

- Cazella, S. C., Behar, P., Schneider, D., da Silva, K. K., and Freitas, R. (2012). Desenvolvendo um sistema de recomendação de objetos de aprendizagem baseado em competências para a educação: relato de experiências. In *Anais do Simpósio Brasileiro de Informática na Educação*, volume 23.
- Cazella, S. C., Reategui, E., Machado, M., and Barbosa, J. (2009). Recomendação de objetos de aprendizagem empregando filtragem colaborativa e competências. *Simpósio Brasileiro de Informática na Educação (SBIE)*.
- Dorigo, M., Birattari, M., Blum, C., Clerc, M., Stützle, T., and Winfield, A. (2008). *Ant Colony Optimization and Swarm Intelligence: 6th International Conference, ANTS 2008, Brussels, Belgium, September 22-24, 2008, Proceedings*, volume 5217. Springer.
- Dorigo, M., Birattari, M., and Stutzle, T. (2006). Ant colony optimization: artificial ant as a computational intelligence technique. university libre de bruxelles. Technical report, IRIDIA Technical report Series, Belgium.
- Dorigo, M., Maniezzo, V., and Colorni, A. (1996). Ant system: optimization by a colony of cooperating agents. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 26(1):29–41.
- Prabha, V. D. and Rathipriya, R. (2013). A study on swarm intelligence techniques in recommender system. In *IJCA Proceedings on International Conference on Research Trends in Computer Technologies 2013*, number 4, pages 32–34. Foundation of Computer Science (FCS).
- Serapião, A. B. d. S. (2009). Fundamentos de otimização por inteligência de enxames: uma visão geral. *Sba: Controle & Automação Sociedade Brasileira de Automatica*, 20(3):271–304.
- Soares, D. d. A. (2009). Objetos de aprendizagem eo ensino de língua inglesa para fins específicos a distância—. *Anais do 7º Encontro de Educação e Tecnologias de Informação e Comunicação-Universidade Estácio de Sá, Setembro*, pages 1–16.
- Sobecki, J. and Tomczak, J. M. (2010). Student courses recommendation using ant colony optimization. In *Asian Conference on Intelligent Information and Database Systems*, pages 124–133. Springer.
- Vetromille-Castro, R., Moor, A. M., Duarte, G. B., and Sedrez, N. H. (2012). Objetos de aprendizagem de línguas: uma proposta. *Aprendizagem de Línguas-a Presença na Ausência: CALL, Atividade e Complexidade. Uma Homenagem aos*, 70.
- Vetromille-Castro, R., Moor, A. M., Duarte, G. B., and Sedrez, N. H. (2013). From learning objects to language learning objects: Communicative language teaching principles in call material. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT)*, 3(2):82–96.
- Wang, Y., Liao, X., Wu, H., and Wu, J. (2012). Incremental collaborative filtering considering temporal effects. *arXiv preprint arXiv:1203.5415*.
- Zaina, L., Bressan, G., Cardieri, M., and Júnior, J. F. R. (2012). e-lors: Uma abordagem para recomendacao de objetos de aprendizagem. *Revista Brasileira de Informática na Educação*, 20(1):04.