



Uma Abordagem Multiobjetivo para Recomendação de Caminhos de Aprendizagem para Grupo de Usuários

Title: A Multiobjective Approach to Recommend Learning Paths for Group of Users

Pérciles B. C. de Miranda
Universidade Federal
Rural de Pernambuco
pericles.miranda@ufrpe.br

Samuel J. Souza
Universidade Federal
de Pernambuco
samuel.souza@hotmail.com

Rafael Ferreira
Universidade Federal
Rural de Pernambuco
rafael.mello@ufrpe.br

Lucas A. Santos
Universidade Federal
Rural de Pernambuco
samuel.souza@hotmail.com

Mayara S. O. Castro
Universidade Federal
Rural de Pernambuco

Lizandra L. S. B. Silva
Universidade Federal
Rural de Pernambuco

Giuseppe F. Neto
Universidade Federal
Rural de Pernambuco

Resumo

A definição de caminhos de aprendizagem é um ponto fundamental para a aprendizagem eficiente. Este artigo propõe a utilização de um algoritmo evolucionário (NSGA-II) para otimizar o processo de recomendação de caminhos de aprendizagem para um grupo de indivíduos, levando em conta múltiplos critérios: satisfação dos membros da equipe e tempo gasto na realização da atividade. O método proposto foi aplicado em dois cenários: com uma turma online e com uma turma presencial. O método foi avaliado quanto ao desempenho, sendo comparado às abordagens exaustiva e aleatória; e quanto ao aspecto pedagógico, sendo comparado com os métodos aleatório e autosselcionado. Os resultados alcançados revelaram o potencial do método proposto tanto do ponto de vista computacional quanto do pedagógico.

Palavras-chave: Recomendação de caminhos de aprendizagem, Otimização multiobjetivo, Otimização combinatória

Abstract

Defining learning paths is a crucial issue for effective learning. This article proposes the application of an evolutionary algorithm (NSGA-II) to optimize the recommendation process of learning paths for a group of individuals taking into account multiple criteria: satisfaction of the team members and time spent in the accomplishment of the activity. The proposed method was applied in two scenarios: an online course and presencial course. The method was evaluated regarding the performance, being compared to the exhaustive and random approaches; And regarding the pedagogical aspect, being compared with random and self-selected methods. The results showed the potential of the proposed method from the computational point of view as well as the pedagogical point of view.

Keywords: Learning path recommendation, Multi-objective optimization, Combinatorial optimization



1 Introdução

Os avanços recentes no campo da *e-learning* tornaram possível o uso de ambientes virtuais para alcançar estudantes sem limites espaciais e temporais (Voss, Nunes, Herpich, & Medina, 2013). Porém, ainda existem dificuldades a serem enfrentadas nesses ambientes como: falta de motivação dos alunos, acompanhamento individual do professor para a grande quantidade de estudantes e utilização de todos os recursos disponíveis. Uma das razões para esses problemas ocorrerem é que a aprendizagem não se dá apenas por meio da disponibilização de recursos de aprendizagem, mas também da identificação de caminhos de aprendizagem interessantes para cada estudante (dos Santos, Cechinel, Araújo, & Brauner, 2015).

Um Caminho de aprendizagem é um conjunto de unidades de aprendizagem (por exemplo, materiais de estudo) que determina a rota pela qual o aluno deve seguir, em termos de conhecimento a ser adquirido ou aperfeiçoado, para que o mesmo realize tarefas relacionadas a essas competências (Li et al., 2009). Em muitas situações, alunos precisam participar de atividades em grupo, como desenvolver um site ou escrever um artigo. Assim, para finalizar a atividade, cada componente do grupo precisa desenvolver habilidades e aprender novos conteúdos para realizar suas tarefas. Portanto, é fundamental conhecer o perfil de cada membro do grupo e determinar um caminho de aprendizagem eficiente e efetivo para cada um, de modo que ajude a equipe terminar a atividade (Xie et al., 2017).

Em um trabalho recentemente publicado no *Simpósio Brasileiro de Informática na Educação 2017* (SBIE 2017), foi proposto um genético multiobjetivo (MOGA) para a sugestão de caminhos de aprendizagem para membros de um grupo, levando em conta os seus perfis (Miranda et al., 2017). Para isso, o problema em questão foi tratado como sendo uma tarefa de otimização com múltiplos critérios em um contexto combinatório. Foram considerados nesse artigo dois objetivos: maximização da satisfação dos membros do grupo, fazendo com que cada membro alcance seus resultados de aprendizagem almejados; e minimizar o tempo necessário para a conclusão da atividade em equipe. Foram realizadas avaliações quantitativa e pedagógica no método proposto. Para a realização da análise quantitativa, o método foi comparado com as abordagens exaustiva e aleatória, usando um conjunto de dados de valores simulados. Para a avaliação pedagógica, foi realizado um experimento real com estudantes de um *Massive Open Online Course* (MOOC): *Desenvolvendo um e-Commerce com Python 3 e Django*. O método foi comparado com duas abordagens tradicionais para definição de caminhos de aprendizagem: aleatória e auto-organizada. Os resultados constataram que, tanto do ponto de vista de desempenho quanto do pedagógico, o método proposto é eficiente e capaz de formar caminhos de aprendizagem adequados para os grupos, considerando os critérios de satisfação e tempo.

Este artigo é uma versão expandida do trabalho preliminar de Miranda et al. (2017), mencionado anteriormente. O objetivo principal desta expansão é avaliar a abordagem desenvolvida em (Miranda et al., 2017) em uma turma presencial. Além disso, encorpar mais as seções sobre o problema de recomendação de caminhos de aprendizagem, trabalhos relacionados e discussão dos resultados. Neste trabalho foram adicionadas as seguintes contribuições:

- Explicação mais detalhada à respeito do problema de recomendação de caminhos de aprendizagem;



- Discussão maior sobre os trabalhos relacionados;
- Explicação mais aprofundada sobre otimização multiobjetivo;
- Inclusão de um experimento real com estudantes de uma turma presencial do último ano do curso de programação de computadores;
- Apresentação e análise dos resultados obtidos com o novo experimento.

Este artigo está organizado da seguinte forma: a Seção 2 introduz o problema de definição de caminhos de aprendizagem e os principais trabalhos relacionados. A Seção 3 apresenta a modelagem computacional do problema para ser resolvido por MOGAs. A Seção 4 apresenta a metodologia experimental utilizada para avaliar o método proposto. A Seção 5 apresenta os resultados alcançados. Finalmente, a Seção 6 destaca as conclusões e trabalhos futuros.

2 Trabalhos Relacionados

Numerosos estudos sobre sistemas de *e-learning* e o desenvolvimento desses sistemas foram realizados nos últimos anos (Xie et al., 2017). A pesquisa sobre a descoberta de caminhos de aprendizagem investigou principalmente dois tipos de questões: questões explícitas, como a provisão direta de métodos de aprendizagem (por exemplo, caminhos de aprendizagem, planos de aprendizagem, etc.); questões implícitas, como sugestões indiretas de métodos de aprendizagem, por meio da apresentação de materiais de aprendizagem (por exemplo, recomendações de cursos ou materiais didáticos).

A tarefa de recomendar caminhos de aprendizagem torna-se difícil, devido à complexidade na qual as características e condições iniciais do indivíduo ou grupo de indivíduos interagem entre si (Xie et al., 2017). Esta relação entre essas características e a explosão combinatória relativa ao número de estudantes, que é fatorial, faz da tarefa em questão um problema combinatório *NP-hard*, ou seja, um problema cuja solução ótima não pode ser encontrada por um algoritmo em tempo polinomial (Xie et al., 2017).

Além de ser um problema complexo, a atribuição de um caminho de aprendizagem inadequado à um indivíduo pode levar a consequências negativas (Xie et al., 2017), tais como: falta de interesse, atraso na entrega da atividade e má formação de competências (Cardoso, Rissoli, Moreira, & Borges, 2016; Belacel, Durand, & Laplante, 2014). Diante da complexidade, importância e impacto desta tarefa no processo de aprendizagem de um aluno ou grupo de alunos, abordagens computacionais têm sido desenvolvidas e aplicadas neste contexto.

Estudos na área de descoberta de caminhos de aprendizagem focam, em sua maioria, na tentativa de facilitar o processo de aprendizagem. Ao longo dos anos, mineração de dados educacionais e sistemas de recomendação propuseram contribuições significativas para fornecer, aos aprendizes, material de aprendizagem adequado recomendando documentos educacionais (Tang & McCalla, 2010) ou links de internet (Godoy & Amandi, 2010), usando filtragem baseada em conteúdo ou colaborativa. Essas abordagens geralmente visam a recomendação de material didático que satisfaça um interesse imediato em vez de se adequar ao processo de aprendizagem sequencial do aprendiz.



As abordagens a seguir não se restringem à recomendação de objetos de aprendizagem, mas seqüências de caminhos de aprendizagem. O trabalho de Steiner et al. (2007) propõe, a partir de redes conceituais dinâmicas, um mecanismo responsável por entregar planos de estudo personalizados e individuais. Feng et al. (2010) definiram métodos para encontrar os caminhos de aprendizagem coletivos, partindo da análise de cada indivíduo do grupo. Diferentemente dos trabalhos anteriores, (Hwang, Kuo, Yin, & Chuang, 2010) considerara a descoberta de caminhos de aprendizagem como sendo um problema de otimização, na qual a solução ideal seria a maximização da aprendizagem dos alunos. Já Madhour et al. (2008) utilizaram um algoritmo inspirado em colônia de formigas (ACO) baseado no modelo do usuário e na experiência de outros alunos, proporcionando caminhos de aprendizagem personalizados. O trabalho desenvolvido por Durand et al. (2013) propôs um modelo baseado em grafos dinâmicos e uma heurística para encontrar o caminho de aprendizagem em um grafo contendo objetos de aprendizagem. Também vale mencionar o trabalho de Belacel et al. (2014), que aplicou um modelo de programação inteira binária para a otimização de caminhos de aprendizagem. Por fim, destaca-se o trabalho desenvolvido por Xie et al. (2017) que utilizou um método exaustivo para encontrar caminhos ótimos para grupos de indivíduos levando em conta o perfil de cada membro. Este último trabalho foi o primeiro a definir caminhos de aprendizagem em um cenário coletivo.

Com o intuito de facilitar a comparação entre os trabalhos relacionados, a Tabela 1 sumariza as características de cada trabalho mencionado anteriormente. É importante salientar que como o trabalho desenvolvido por (Steiner & Albert, 2007) não considerou o problema de recomendação como uma tarefa de otimização, portanto, este não apresenta número de objetivos.

Table 1: Análise comparativa dos trabalhos relacionados..

Artigos	Abordagem	Características	Número de objetivos	Experimento
(Steiner & Albert, 2007)	Redes conceituais	Interesse Habilidades Personalidade	-	Qualitativo
(Madhour & Wentland Forte, 2008)	ACO	Interação social Habilidades	1	Qualitativo
(Feng et al., 2010)	Heurística - GLDP	Habilidades	1	Qualitativo
(Hwang et al., 2010)	Heurística - TSP	Interesse Habilidades	1	Qualitativo
(Durand et al., 2013)	Heurística - Grafo	Habilidades Notas	1	Qualitativo
(Belacel et al., 2014)	<i>Branch and Bound</i>	Habilidades Notas	1	Qualitativo
(Xie et al., 2017)	Exaustiva	Habilidades Notas Personalidade	1	Qualitativo

Embora os trabalhos anteriores sejam relevantes, os mesmos apresentam três limitações. A primeira delas é quanto ao número de características consideradas, no máximo três características para descrever os indivíduos. A segunda é quanto ao número de objetivos levados em consideração na otimização dos grupos. Apenas um único critério é considerado. No entanto, a recomendação



de caminhos de aprendizagem para é um problema com múltiplos objetivos que necessitam ser satisfeitos em conjunto. Finalmente, a terceira limitação é quanto ao experimento. Dentre os trabalhos, apenas o experimento qualitativo foi realizado. Embora seja importante saber se a solução encontrada é útil na prática, é fundamental saber o custo computacional para encontrá-la (experimento quantitativo).

Diferentemente dos estudos apresentados, propomos um método para a recomendação automática de caminhos de aprendizagem considerando múltiplos critérios: maximização da satisfação dos membros do grupo e minimização do tempo necessário para a conclusão da atividade em equipe; uma gama diversa de características para a descrição dos membros; e, além disso, foram realizados experimentos quantitativo e qualitativo. Vale salientar que este trabalho é pioneiro, ao considerar um cenário multiobjetivo.

3 MOGA para Recomendação de Caminhos de Aprendizagem

O problema de recomendação de caminhos de aprendizagem para um grupo alunos trata-se de um problema combinatório que apresenta um espaço de soluções de dimensão elevada, onde múltiplos critérios podem ser envolvidos. Por este motivo, decidimos utilizar um algoritmo genético multiobjetivo (MOGA). Para isto, faz-se necessário adaptar o algoritmo para o problema em questão, ou seja, definir como as soluções serão representadas e como serão avaliadas.

3.1 Otimização MultiObjetivo

A otimização multiobjetivo (MO) visa encontrar soluções que satisfaçam mais de um critério ao mesmo tempo. Considere um vetor $\vec{C} = (C_1, C_2, \dots, C_n)$ de objetivos, onde n é o número de objetivos a serem considerados na otimização. Deste modo, cada solução possui um vetor $\vec{f} = (f_1, f_2, \dots, f_n)$, onde cada valor (*fitness*) de \vec{f} representa a avaliação da solução em cada critério pertencente a \vec{C} . Como neste trabalho são adotados dois critérios na otimização $\vec{C} = (C_1, C_2)$, cada solução possui um vetor $\vec{f} = (f_1, f_2)$. Como cada solução possui um vetor de *fitness*, a comparação entre diferentes soluções se dá da seguinte forma: dadas duas soluções distintas S_1 e S_2 , onde $f_{(S_1)}$ representa o vetor de *fitness* de S_1 e $f_{(S_2)}$ representa o vetor de *fitness* de S_2 ; S_1 domina S_2 se $f_{(S_1)}$ é melhor que $f_{(S_2)}$ em ao menos um *fitness*, e não é pior em nenhum outro. Deste modo, existem três casos possíveis: 1) S_1 domina S_2 , 2) S_1 é dominada por S_2 , ou 3) S_1 é incomparável em relação à S_2 . Este último caso ocorre quando S_1 e S_2 vencem em pelo menos um critério. O conjunto de soluções incomparáveis (ou não-dominadas) é conhecido como *Pareto Front*. A Figura 1 ilustra o conceito de dominância em um problema de minimização de dois objetivos. Como se pode ver, as soluções sobre a linha pontilhada são consideradas não-dominadas (entre elas as soluções A e B), pois não existe nenhuma outra solução que possua melhores valores de *fitness* em todos os objetivos. Estas soluções compõem o *Pareto front*. As demais soluções (C, D, E, F e G) são consideradas dominadas, ou seja, existe ao menos uma solução que as supere em todos os objetivos.

Deste modo, algoritmos de otimização multiobjetivo tentam encontrar o *Pareto Front* ótimo, ou seja, aquele conjunto de soluções não-dominadas que melhor satisfazem os critérios considerados. No MOGA, uma população inicial de indivíduos (soluções) é definida, e esta passa,

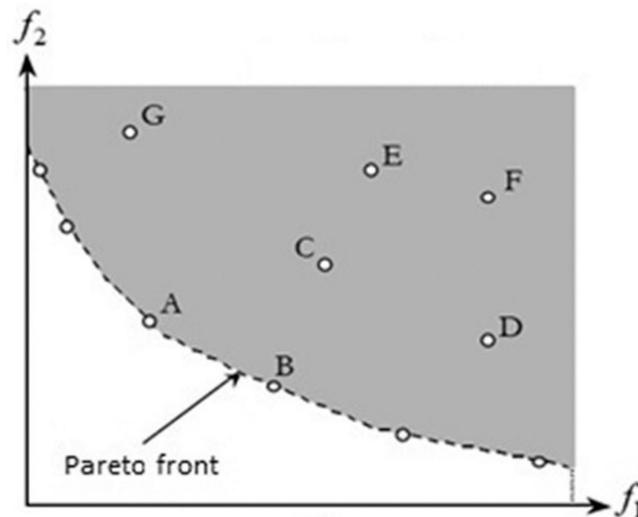


Figure 1: Exemplo de dominância em um *Pareto front*..

iterativamente, por processos evolucionários, tais como avaliação, seleção, cruzamento e mutação de indivíduos. Após um determinado número de gerações, o *Pareto front* resultante, que reflete as melhores soluções, é gerado como saída do algoritmo.

3.2 Representação do Indivíduo

No GA, cada possível solução, do espaço de soluções do problema de otimização, é chamada de indivíduo ou cromossomo. No problema em questão, deseja-se alocar um conjunto de unidades de aprendizagem $\vec{U} = (U_1, U_2, \dots, U_t)$, onde t é o número de unidades, entre os membros do grupo. Diante deste cenário, cada cromossomo representa uma possível alocação de unidades de aprendizagem para um grupo de membros, onde cada membro do grupo é responsável pelas suas unidades. Definindo-se formalmente, um cromossomo seria um vetor de membros do grupo $\vec{I} = (M_1 : \vec{u}_1, M_2 : \vec{u}_2, \dots, M_l : \vec{u}_l)$, onde l é o número de membros do grupo; e cada membro M_i possui um vetor de unidades de aprendizagem alocadas a ele \vec{u}_i . Vale mencionar que diferentes membros não podem possuir a mesma unidade de aprendizagem, e que não há membros sem unidades de aprendizagem alocadas.

Neste trabalho, o processo de otimização dos caminhos de aprendizagem para os membros de um grupo baseia-se no perfil de cada um. Deste modo, é fundamental definir quais características são relevantes para a descrição dos membros e das unidades. Vamos considerar que existam k características importantes para a realização das unidades contidas em \vec{U} . Deste modo, cada membro do grupo M_i , $i = 1 \dots l$, possui um vetor de preferências \vec{P} e um vetor de habilidades \vec{H} , onde cada valor desses vetores estão relacionados a cada uma das k características. \vec{P} armazena valores reais no intervalo $[0,1]$ equivalente ao interesse do membro em relação a certos temas; E \vec{H} armazena valores reais no intervalo $[0,1]$ equivalente às habilidades do membro em relação a certos temas. Esses dois vetores são usados para definir o perfil de cada membro do grupo em relação às unidades de aprendizagem consideradas.



Assim como cada membro do grupo precisou ser descrito por meio das k características, cada unidade de aprendizagem contida em \vec{U} também o foi. Cada unidade possui: 1) a duração de tempo (T) em que um aluno com valor 0 de habilidade naquela unidade em questão gastará para realizá-la; e 2) vetor de habilidades requeridas \vec{HR} , atribuindo-se a cada habilidade um valor no intervalo $[0,1]$, equivalente ao quanto a atividade a necessita para ser concluída.

3.3 Avaliação do Indivíduo

Uma das etapas mais importantes do MOGA é a avaliação das soluções (cromossomo) da população. Nesta etapa, cada solução é avaliada, sendo atribuído a cada solução um valor relativo a sua qualidade. Este valor, chamado de valor de *fitness*, é utilizado como base para a seleção das soluções que permanecerão para as próximas etapas na evolução do algoritmo. Neste trabalho, dois objetivos são considerados na avaliação das soluções, sendo: o tempo, que deve ser minimizado, visto que as atividades devem ser concluídas de maneira a reduzir o tempo total gasto; E a preferência ou satisfação, que deve ser maximizada, garantindo que cada estudante fique responsável por uma atividade que tenha maior afinidade. Os cálculos necessários para a avaliação das soluções serão apresentados a seguir.

Cálculo da Satisfação (ST): Com relação à avaliação da satisfação de cada membro do grupo, foi criada uma matriz $MA_{membros,unidades}$ que armazena o somatório do módulo da diferença, entre os valores de mesmo critério, de \vec{P} (vetor de preferências do membro) e \vec{HR} (vetor de habilidades requeridas pela unidade de aprendizagem). Cada valor de MA é calculado da seguinte forma:

$$MA_{a,b} = \sum_{k=1}^{n^{\circ}\text{-habilidades}} |M_a.P_k - U_b.HR_k|, \quad (1)$$

onde $M_a.P_k$ é o k -ésimo valor do vetor de preferências \vec{P} do membro M_a , e $U_b.HR_k$ é o k -ésimo valor do vetor de habilidades requeridas \vec{HR} da unidade de aprendizagem U_b . Dessa forma, um valor de preferência de cada estudante em relação à cada unidade de aprendizagem é gerado. Quanto mais próximo de zero o valor obtido for, maior será o grau de satisfação em trabalhar com aquela unidade de aprendizagem. Uma vez que MA está preenchida, esta é utilizada para calcular o valor (*fitness*) de satisfação total do grupo (ST). ST nada mais é que a soma da satisfação individual (Si) de cada componente em relação às unidades de aprendizagem alocadas a cada um deles. Deste modo, o valor de ST é obtido através da Equação 2. Diante disto, quanto menor o valor de ST , maior a satisfação total do grupo. Vale salientar que o valor de ST é normalizado entre $[0,1]$.

$$ST = \sum_{j=1}^{n^{\circ}\text{-membros}} Si_j. \quad (2)$$

Cálculo do Tempo (TT): O método proposto otimiza as soluções em relação ao tempo considerando duas situações: minimização do tempo geral de finalização das unidades e aproximando a duração total gasta por cada membro nas unidades alocadas a cada um. O objetivo de levar em conta estas duas situações, é fazer com que haja o estudo das unidades de aprendizagem em menor



tempo, garantindo uma menor probabilidade de ocorrência de sobrecargas entre os membros do grupo.

Na avaliação do tempo, as habilidades dos membros foram consideradas. Como já dito anteriormente, cada valor em \vec{H} (vetor de habilidades de cada membro) é definido entre o intervalo $[0,1]$. Deste modo, quanto maior for o valor de uma habilidade, as unidades de aprendizagem que requererem aquele tipo habilidade, tendem a serem finalizadas em menor tempo. Assim, sendo possível a geração de uma matriz, $MT_{membros,unidades}$, encarregada de definir a previsão do tempo restante em que cada membro levaria para finalizar cada unidade. O cálculo de cada valor de MT é realizado da seguinte forma:

$$MT_{a,b} = (1 - M_a \cdot H_b \times U_b \cdot T), \quad (3)$$

onde $M_a \cdot H_b$ é o b -ésimo valor do vetor de habilidades \vec{H} do membro M_a , e $U_b \cdot T$ é o tempo de um aluno com valor 0 de habilidade na unidade de aprendizagem U_b . Uma vez preenchida, MT é utilizada para calcular o valor (*fitness*) de tempo total do grupo (TT). TT nada mais é do que a soma do tempo gasto de cada membro (Ti) na finalização de suas unidades de aprendizagem. Deste modo, o valor de TT é obtido assim:

$$TT = \sum_{j=1}^{n^o_membros} Ti_j. \quad (4)$$

Como havíamos dito previamente, todos os membros do grupo precisam trabalhar de forma igualitária para evitar a sobrecarga dos demais alunos. Como não há garantia de que um valor de TT baixo signifique que todos os membros do grupo trabalharam de forma igualitária, decidimos criar uma penalidade. Para isto, foi realizado um somatório das diferenças entre o tempo individual de cada membro do grupo:

$$Penalidade = \sum_{a=1}^{n^o_membros} \sum_{b=a}^{n^o_membros} |Ti_a - Ti_b|. \quad (5)$$

Deste modo, quanto menor o valor da *Penalidade*, maior é a aproximação do tempo gasto, entre os estudantes, para a finalização de suas unidades. Diante da importância desta penalidade, incorporamos ao cálculo de TT , o valor da *Penalidade*², resultando em $TT = TT + Penalidade^2$. Em situações em que o valor de TT é baixo e há sobrecarga de trabalho, a *Penalidade* aumenta o valor do *fitness* fazendo com que a solução tenha menos chances de ser escolhida para a próxima geração. Vale salientar que o valor de TT é normalizado entre $[0,1]$.

4 Metodologia Experimental

Nesta seção serão apresentadas as análises realizadas para avaliação do método proposto, as unidades de aprendizagem e características consideradas, e a configuração adotada para o MOGA.



4.1 Análises

Neste trabalho, duas análises foram realizadas: uma que avalia o *desempenho* do algoritmo proposto e outra que avalia o ponto de vista *pedagógico*.

4.1.1 Análise de Desempenho

Nesta análise, objetivamos investigar se o MOGA consegue obter resultados competitivos no problema em questão. Com isto, o desempenho do método proposto foi comparado com um método exaustivo e um aleatório (ambos em suas versões multiobjetivo, levando em consideração a satisfação e o tempo). Todos os métodos usaram um conjunto de dados (características e unidades de aprendizagem) de valores simulados (uniformemente distribuídos). Vale salientar que as características consideradas são as mesmas descritas na Seção 4.3. Por questões práticas de comparação dos resultados, decidimos, ao invés de comparar os *Pareto* obtidos por cada método, selecionar apenas uma solução de cada *Pareto Front*. Esta prática é muito comum em análises de métodos multiobjetivo, e geralmente utiliza-se o método de *Borda Count* (Orouskhani, Teshnehlab, & Nekoui, 2017). Este método utiliza uma estratégia de ranqueamento por objetivo, de modo que a solução do *Pareto Front* que estiver melhor ranqueada é selecionada. O objetivo da análise de desempenho é verificar se o método proposto é capaz de encontrar um bom caminho de aprendizagem em tempo hábil.

4.1.2 Análise Pedagógica

A análise pedagógica objetiva investigar se as recomendações da abordagem proposta trás consequências positivas para o desempenho do grupo. Para a realização da análise pedagógica, foram realizados dois experimentos: o primeiro experimento foi feito com estudantes do MOOC *Desenvolvendo um e-Commerce com Python 3 e Django*, em fase de conclusão do curso e elaboração do projeto final; o segundo experimento foi feito com estudantes do último ano do curso de programação de computadores.

No primeiro experimento, foi considerada uma turma com 120 alunos: graduandos, profissionais e pesquisadores com experiência e conhecimentos diversos, fatores importantes nesta pesquisa. Os 120 alunos foram divididos em 24 grupos de 5 membros, onde os próprios alunos se auto-organizaram. A cada grupo foi atribuída a tarefa de desenvolver um *e-commerce em Python 3 e Django*, com o auxílio de 25 unidades de aprendizagem (5 unidades por aluno), sendo cada uma um material e estudo relativo a um determinado assunto. Além disso, embora os projetos de cada grupo apresentem o mesmo nível de complexidade e requisitos similares, cada equipe é responsável por um projeto diferente para evitar cópia.

No segundo experimento foi selecionada uma amostra de 60 alunos de três turmas diferentes (20 alunos por turma). Cada turma foi dividida em 5 grupos de 4 alunos, onde os próprios alunos se auto-organizaram. Vale salientar que o número de 4 alunos por grupo foi definido pelos professores de cada turma, uma vez que a complexidade do projeto é proporcional a este número de alunos. A atividade de aprendizagem sugerida consistiu em cada grupo de cada turma desenvolver um *e-commerce em Python 3 e Django*, fazendo uso do conhecimento adquirido durante o curso. Para evitar cópia, cada equipe ficou responsável por um projeto diferente.



Em cada experimento (online e presencial), o método proposto é comparado com duas abordagens tradicionais para definição de caminhos de aprendizagem: aleatória e autosselecionada. No caso do experimento online, como existem ao todo 24 grupos de 5 alunos, cada método foi aplicado sobre 8 grupos diferentes. No caso do experimento presencial, como existem ao todo 15 grupos de 5 alunos, cada método foi aplicado sobre 4 grupos diferentes.

O método aleatório atribui, a cada aluno, unidades de aprendizagem aleatórias. Já no método autosselecionado, os alunos escolhem as suas próprias unidades. Os grupos atrelados a cada abordagem serão avaliados pedagogicamente, ou seja, em relação aos seus desempenhos (notas) obtidos em seus projetos. O objetivo é avaliar se, na prática, o método proposto é capaz de melhorar o desempenho dos seus grupos, quando comparados aos das demais abordagens. O cálculo da nota final do grupo é a média entre a nota do projeto, e a média das notas de cada aluno do grupo. Deste modo, projetos que não foram entregues por completo e/ou alunos que não tenham contribuído com a realização do projeto, sejam motivo para penalização da nota do grupo.

4.2 Unidades de Aprendizagem

Como se pode ver na Tabela 2, existem 25 unidades que abrangem diversos tipos de conhecimento: *design*, *configuração*, *codificação*, *modelagem*, *teste* e *deploy*. Cada unidade está associada a um conjunto de requisitos. Deste modo, os alunos devem resolver os requisitos do sistema que estiverem associados às suas 5 unidades.

Table 2: Unidades de Aprendizagem relativas ao projeto de um e-commerce.

Unidades de Aprendizagem		
Orientação a Objetos em Python	django.contrib.messages	Teste de Performance
Configuração do Django	Custom Template Tags	Teste de Views
Configuração do Banco de Dados	Custom User e Custom User Model	Teste de Forms e E-mail
Desenvolvimento de Templates	Class-based views para o Fórum	Teste de Models com model-mommy
App Courses	Signal para envio de emails	Deploy no Heroku
Model Objects	Ajustando Design do Dashboard	Prototipação do design da aplicação
URL's	Modelagem do Banco de Dados	Html e CSS
Views: Página, Listagem e Login	Mapeamento dos Dados	
Formulários	Testes Unitários	

4.3 Características Consideradas

Neste trabalho foram adotados dois aspectos para definir o perfil dos alunos e o perfil das unidades de aprendizagem: preferência e habilidades em certos assuntos. Estes aspectos foram selecionados com base em uma revisão da literatura, nas características particulares da população estudada e nas unidades de aprendizagem consideradas (ver Tabela 2). Diante disto, 7 características foram selecionadas ($k = 7$) relativas à: 1) *design*, 2) configuração, 3) desenvolvimento de programas, 4) modelagem, 5) teste de software, 6) *deploy*, e 7) liderança. A escolha destas características é baseada no fato de que é interessante que cada grupo tenha membros responsáveis por atividades que sabem e gostam de fazer.

A extração de características dos alunos envolvidos nos experimentos se deu por meio da aplicação de um questionário, e da análise do desempenho de cada aluno no curso. O questionário consistiu em um conjunto de perguntas de resposta única e de múltipla escolha sobre tópicos relacionados às características.



4.4 Configurações

O método proposto adotou o MOGA chamado *Non-Dominated Sorting Genetic Algorithm* (NSGA-II) adaptado para o problema em questão. O método utilizou a seleção baseada em *roleta*, onde 40% dos indivíduos da população são clonados e mantidos na geração seguinte. Os outros 60% dos indivíduos são utilizados pelo operador de cruzamento. O operador de mutação foi aplicado na população em cada geração com uma probabilidade de mutação para cada indivíduo $\theta = 0.2$ e uma probabilidade de mutação para cada cromossomo $\theta = 0.15$. O tamanho populacional, $TAM = 50$, bem como o número máximo de gerações, $n_{gen} = 500$, foram definidos empiricamente. O método aleatório adotou o mesmo número de soluções candidatas que as utilizadas para o método proposto. Os métodos aleatório e proposto foram executados 30 vezes para gerar um desempenho médio. Todos os métodos foram implementados em Python usando o ambiente Microsoft Visual Studio 2010 e todas as execuções foram realizadas em um laptop com Intel Core i7-5600U com 4M Cache, 8GB RAM e 3.20 GHz.

5 Resultados

Esta seção apresenta resultados provenientes de duas análises diferentes. A primeira leva em conta o aspecto desempenho (quantitativo), enquanto a segunda, o aspecto pedagógico (qualitativo).

5.1 Análise de Desempenho

Nesta análise, o desempenho do método proposto foi comparado com o da abordagem exaustiva e aleatória. A Tabela 3 mostra o tempo médio de execução de cada método (medido em segundos) e o valor de *fitness* médio alcançado em cada objetivo (satisfação e tempo de realização da atividade). Por uma questão de simplicidade, o número de membros por grupo foi de 5 em todos os casos. Além disso, o desempenho de cada método foi avaliado variando-se o número de unidades de aprendizagem.

Table 3: Análise comparativa do desempenho entre as abordagens exaustiva, aleatória e proposta..

Unidades	Exaustivo			Aleatório			Proposta		
	Tempo	ST	TT	Tempo	ST	TT	Tempo	ST	TT
10	0.49	0.21	0.23	0.59	0.25	0.23	0.62	0.21	0.23
15	90.45	0.16	0.19	7.89	0.28 (↓)	0.27 (↓)	9.87	0.19	0.22
20	878.43	0.13	0.17	12.34	0.35 (↓)	0.34 (↓)	15.76	0.16	0.21
25	7349.2	0.09	0.13	32.12	0.37 (↓)	0.37 (↓)	40.98	0.14	0.18
30	18216.1	0.03	0.06	59.76	0.41 (↓)	0.42 (↓)	72.34	0.11 (↓)	0.16 (↓)
50	N/A	N/A	N/A	187.36	0.52	0.47	210.39	0.07	0.13
100	N/A	N/A	N/A	420.29	0.63	0.51	480.74	0.07	0.10

Como se pode ver, o tempo de execução necessário para o método exaustivo avaliar todas as soluções possíveis aumenta drasticamente a medida que o número de unidades de aprendizagem também aumenta. Por outro lado, o tempo de execução do método proposto e aleatório aumenta quase que linearmente. É importante mencionar que o método exaustivo alcança melhores valores de *fitness* à medida que aumenta a quantidade de unidades de aprendizagem. Isso se deve ao fato de que mais unidades de aprendizagem, mais combinações de caminhos podem ser formados. O



mesmo não acontece com o método aleatório pois, quanto maior o número de soluções possíveis, menores são as chances de encontrar boas combinações de caminhos por sorte.

Quanto ao desempenho do método proposto, vale destacar que mesmo não alcançando valores ótimos (iguais aos do exaustivo), o método foi capaz de encontrar boas soluções em tempo de execução menor, até quando a quantidade de unidades de aprendizagem aumenta. Para realizar uma comparação justa entre os métodos, aplicamos o teste de hipótese *t-student*, considerando 95% de significância. Aqueles valores de *fitness*, apresentados na Tabela 3, que estiverem em negrito significam que são estatisticamente iguais aos do exaustivo. Em contrapartida, os valores com o símbolo ↓ são estatisticamente inferiores aos do exaustivo. A comparação estatística realizada não levou em consideração os valores para 50 e 100 alunos, uma vez que os resultados do método exaustivo não foram obtidos por questões de custo computacional.

Como se pode ver, a maioria dos resultados alcançados pelo método aleatório, para cada objetivo, foram superados estatisticamente pelos resultados da abordagem exaustiva. Apenas no cenário mais simples que a abordagem aleatória teve bons resultados. Por outro lado, os valores de *fitness* obtidos pelo método proposto foram estatisticamente iguais aos do método exaustivo nos cenários de 10, 15, 20 e 25 alunos. Embora os valores para 30 unidades tenham sido estatisticamente piores que o do exaustivo, os resultados foram bons e alcançados em bem menos tempo. Talvez, com mais gerações, melhores resultados seriam alcançados. Com relação aos resultados para 50 e 100 alunos, o método proposto foi capaz de melhorar ainda mais os valores de *fitness* dos dois objetivos. Quando comparado à abordagem aleatória, o método proposto a superou estatisticamente em todos os objetivos, na maioria dos cenários.

5.2 Análise Pedagógica

Os grupos, atrelados a cada método, tiveram seu desempenho pedagógico (nota do grupo) avaliado. A nota do grupo é a composição da nota do projeto e da nota de cada membro da equipe (ver Seção 4).

5.2.1 Turma à Distância

A Tabela 4 apresenta os resultados pedagógicos colhidos por meio da aplicação do método proposto, aleatório e autosselecionado, em 24 turmas de 5 estudantes (8 turmas por método). Estão explicitados na Tabela 4, por método, a média e desvio padrão das notas dos projetos dos grupos; a média e o desvio padrão das notas dos indivíduos dos grupos; e a média e o desvio padrão das notas dos grupos. De acordo com os professores, os projetos feitos por equipes com caminhos de aprendizagem definidos a partir do método entregaram o projeto com todos os requisitos. A média das notas dos projetos entregues é 8.8 com um desvio padrão de 1.1, ou seja, o projeto com menor nota ainda assim atingiu nota acima de 7.0 (nota necessária para aprovação). Quanto às notas individuais, nota-se que a média das notas dos membros também é alta. Isso comprova a efetividade da contribuição de cada membro das equipes em seus projetos, e a entrega das funcionalidades requisitadas. Como consequência, a média final dos grupos atendidos pelo método proposto foi bastante satisfatória.

Os grupos que adotaram a abordagem autosselecionada obtiveram 7.5 como média das no-



Table 4: Média das notas dos grupos e individuais de cada abordagem..

	N. Alunos	Notas dos Projetos	Notas Individuais	Notas dos Grupos
Proposta	40	8.8 (± 1.1)	8.4 (± 0.6)	8.6 (± 0.85)
Aleatória	40	5.1 (± 0.98)	6.4 (± 0.47)	5.75 (± 0.72)
Autosselecionada	40	7.5 (± 3.5)	6.2 (± 3.1)	6.85 (± 3.3)

tas nos projetos, mas com 3.5 de desvio padrão, ou seja, as notas dos projetos das equipes tiveram grande variação. O mesmo aconteceu com as notas individuais, onde a média foi de 6.2 e o desvio padrão de 3.1. Este cenário, em que as notas apresentam grande variação, mostra o desbalanceamento da contribuição dos alunos. De acordo com os professores, alguns grupos, que adotaram a abordagem autosselecionada, não entregaram o projeto com todos os requisitos. Como consequência, a média do grupo foi penalizada, resultando em 6.85. Por fim, os resultados mostram que a abordagem aleatória não foi capaz de atingir resultados competitivos frente aos métodos proposto e autosselecionado. A média das notas dos projetos daqueles grupos que adotaram a abordagem aleatória é 5.1 e a média das notas individuais é 6.4, ambas com baixo desvio padrão. De acordo com o relato dos professores, algumas equipes não chegaram a entregar todos os requisitos do projeto. Diante de tudo isto, a nota média dos grupos, que adotaram esta abordagem, é 5.75.

5.2.2 Turmas presenciais

A Tabela 5 apresenta os resultados pedagógicos do segundo experimento (turma presencial do último ano do curso de programação de computadores), colhidos por meio da aplicação das três abordagens de recomendação: proposta, aleatória e autosselecionada, em turmas de 20 estudantes (5 grupos de 4 alunos). São apresentados na Tabela 5, por abordagem de recomendação, a média e desvio padrão das notas dos projetos de cada turma; a média e o desvio padrão das notas dos indivíduos de cada turma; e a média e o desvio padrão das notas dos grupos de cada turma. Os professores relataram que os projetos desenvolvidos por equipes que adotaram caminhos de aprendizagem fornecidos pelo método proposto conseguiram entregar o projeto completo. A média das notas dos projetos entregues é 8.5 com um desvio padrão de 1.2. Em relação às notas individuais, percebeu-se também a contribuição efetiva de cada componente de cada equipe em seus projetos. Como consequência, a nota média dos grupos que utilizaram a proposta foi bastante satisfatória.

Table 5: Média das notas dos grupos e individuais de cada abordagem no primeiro experimento..

	N. Alunos	Notas dos Projetos	Notas Individuais	Notas dos Grupos
Proposta	20	8.5 (± 1.2)	8.4 (± 0.98)	8.45 (± 1.09)
Aleatória	20	5.4 (± 2.34)	6.2 (± 0.42)	5.8 (± 1.38)
Autosselecionada	20	6.0 (± 3.2)	6.5 (± 2.8)	6.25 (± 3.0)

Em relação à abordagem autosselecionada, se pode ver que a média das notas nos projetos foi 6.0, mas com 3.2 de desvio padrão, ou seja, as notas dos projetos das equipes variaram bastante. O resultado das notas individuais seguiram o mesmo comportamento, a média foi de 6.5 e o desvio padrão de 2.8. Esta informação confirma que a contribuição dos alunos foi desbalanceada, ou seja, alguns alunos não colaboraram como deveriam, enquanto outros ficaram sobrecarregados. Por este motivo, alguns projetos, das equipes que adotaram a abordagem autosselecionada, não foram entregues por completo. A média do grupo, portanto, foi penalizada, resultando em 6.25. Quanto



à abordagem aleatória, os resultados mostraram que esta não foi capaz prover bons caminhos de aprendizagem para os alunos do grupo. Como se pode ver, a média das notas individuais dos membros dos grupos formados aleatoriamente é 6.2, com baixo desvio padrão, ou seja, as notas dos alunos em grupos com caminhos de aprendizagem aleatórios não variam muito, ficando próximas da média. O que torna crítica a situação dos grupos que adotaram o método aleatório é a nota do projeto, 5.4. De acordo com o relato dos professores, algumas equipes com esta abordagem não chegaram a entregar todos os requisitos do projeto. Diante de tudo isto, a nota média das equipes com esta abordagem é 5.8.

6 Conclusão e Trabalhos Futuros

Este trabalho apresenta um método automático para a definição de caminhos de aprendizagem para membros de um grupo, considerando critérios como a satisfação dos membros da equipe e o tempo gasto pelo grupo na realização de uma atividade. Os resultados mostraram que o método proposto foi capaz de atingir desempenho estatisticamente similar ao do método exaustivo na maioria dos cenários, mas, com custo de tempo quase linear. Além disso, o método superou a abordagem aleatória na maioria dos critérios. Em relação à análise pedagógica, foi constatado que grupos que utilizaram a recomendação de caminhos de aprendizagem sugeridos pelo método proposto obtiveram melhores notas na atividade dada quando comparados com os demais grupos que adotaram um método autosselecionado e aleatório. Futuramente, pretendemos considerar outras características para ajudar na melhor definição do perfil do aluno. Além disso, objetivamos investigar e incluir novos critérios no processo de otimização, além de satisfação e tempo de entrega da atividade (considerados neste trabalho).

References

- Belacel, N., Durand, G., & Laplante, F. (2014). A binary integer programming model for global optimization of learning path discovery. In *Edm (workshops)*. [\[\[GS Search\]\]](#)
- Cardoso, I. C., Rissoli, V., Moreira, T., & Borges, V. (2016). Construção de ambientes interativos de apoio à reflexão docente aplicados à teoria da aprendizagem significativa. In *Anais dos workshops do congresso brasileiro de informática na educação* (Vol. 5, p. 637). [DOI: 10.5753/cbie.wcbie.2016.637 [\[GS Search\]\]](#)
- dos Santos, H., Cechinel, C., Araújo, R., & Brauner, D. (2015). Recomendação de objetos de aprendizagem utilizando filtragem colaborativa: Uma comparação entre abordagens de pré-processamento por meio de clusterização. In *Brazilian symposium on computers in education (simpósio brasileiro de informática na educação-sbie)* (Vol. 26, p. 1127). [DOI: 10.5753/cbie.sbie.2015.1127 [\[GS Search\]\]](#)
- Durand, G., Belacel, N., & LaPlante, F. (2013). Graph theory based model for learning path recommendation. *Information Sciences*, 251, 10-21. [DOI: 10.1016/j.ins.2013.04.017 [\[GS Search\]\]](#)
- Feng, X., Xie, H., Peng, Y., Chen, W., & Sun, H. (2010). Groupized learning path discovery based on member profile. In *International conference on web-based learning* (pp. 301–



- 310). [DOI: 10.1007/978-3-642-20539-2_32 [\[GS Search\]](#)]
- Godoy, D., & Amandi, A. (2010). Link recommendation in e-learning systems based on content-based student profiles. *Handbook of educational data mining*, 273–286. [[\[GS Search\]](#)]
- Hwang, G.-J., Kuo, F.-R., Yin, P.-Y., & Chuang, K.-H. (2010). A heuristic algorithm for planning personalized learning paths for context-aware ubiquitous learning. *Computers & Education*, 54(2), 404–415. [DOI: 10.1016/j.compedu.2009.08.024 [\[GS Search\]](#)]
- Li, Q., Lau, R. W., Wah, B. W., Ashman, H., Leung, E. W., Li, F., & Lee, V. (2009). Guest editors' introduction: Emerging internet technologies for e-learning. *IEEE Internet Computing*, 13(4), 11–17. [DOI: 10.1109/MIC.2009.83 [\[GS Search\]](#)]
- Madhour, H., & Wentland Forte, M. (2008). Personalized learning path delivery: Models and example of application. In *Intelligent tutoring systems* (pp. 725–727). [DOI: 10.1007/978-3-540-69132-7_90 [\[GS Search\]](#)]
- Miranda, P., Ferreira, R., Fiorentino, G., Ligia, L., Souza, S., Castro, M., & André, L. (2017). Seleção de caminho de aprendizagem para grupo de usuários: uma abordagem baseada em perfil. In *Brazilian symposium on computers in education (simpósio brasileiro de informática na educação-sbie)* (Vol. 28, p. 1167). [DOI: 10.5753/cbie.sbie.2017.1167 [\[GS Search\]](#)]
- Orouskhani, M., Teshnehlab, M., & Nekoui, M. A. (2017). Evolutionary dynamic multi-objective optimization algorithm based on borda count method. *International Journal of Machine Learning and Cybernetics*, 1–29. [DOI: 10.1007/s13042-017-0695-3 [\[GS Search\]](#)]
- Steiner, C. M., & Albert, D. (2007). Personalising learning through prerequisite structures derived from concept maps. In *International conference on web-based learning* (pp. 43–54). [DOI: 10.1007/978-3-540-78139-4_5 [\[GS Search\]](#)]
- Tang, T. Y., & McCalla, G. G. (2010). Data mining for contextual educational recommendation and evaluation strategies. *Handb. Educ. Data Min.*, 257. [[\[GS Search\]](#)]
- Voss, G. B., Nunes, F. B., Herpich, F., & Medina, R. D. (2013). Ambientes virtuais de aprendizagem e ambientes imersivos: um estudo de caso utilizando tecnologias de computação móvel. In *Anais do simpósio brasileiro de informática na educação* (Vol. 24, p. 12). [DOI: 10.5753/cbie.sbie.2013.12 [\[GS Search\]](#)]
- Xie, H., Zou, D., Wang, F. L., Wong, T.-L., Rao, Y., & Wang, S. H. (2017). Discover learning path for group users: A profile-based approach. *Neurocomputing*, 254(1), 59–70. [DOI: 10.1016/j.neucom.2016.08.133 [\[GS Search\]](#)]