

Um GRASP Híbrido com Reconexão por Caminhos e Mineração de Dados

Hugo Barbalho¹, Isabel Rosseti¹, Simone L. Martins¹, Alexandre Plastino¹

¹ Departamento de Ciência da Computação – Universidade Federal Fluminense
Niterói – RJ – Brazil.

{hbarbalho, rosseti, simone, plastino}@ic.uff.br

Abstract. *The exploration of hybrid metaheuristics – combination of metaheuristics with concepts and processes from other research areas – has been an important trend in combinatorial optimization research. In this work, we developed a hybrid version of the GRASP metaheuristic which incorporates the path-relinking procedure – a memory-based intensification strategy – and a data mining module. Computational experiments showed that employing the combination of path-relinking and data mining allowed GRASP to find better results in less computational time. Another contribution of this work is the application of the path-relinking hybrid proposal for the 2-path network design problem, which improved the state-of-the-art solutions for this problem.*

Resumo. *A exploração de metaheurísticas híbridas – combinação de metaheurísticas com conceitos e processos de outras áreas – vem sendo uma importante linha de pesquisa em otimização combinatória. Neste trabalho, propõe-se uma versão híbrida da metaheurística GRASP que incorpora a técnica de reconexão por caminhos e um módulo de mineração de dados. Experimentos computacionais mostraram que a combinação da técnica de reconexão por caminhos com mineração de dados contribuiu para que o GRASP encontrasse soluções melhores em um menor tempo computacional. Outra contribuição deste trabalho é a aplicação dessa proposta híbrida ao problema de síntese de redes a 2-caminhos, que proporcionou encontrar melhores soluções para esse problema.*

1. Introdução

Metaheurísticas representam uma importante classe de técnicas aproximativas utilizadas para resolver problemas de otimização combinatória, para os quais o uso de métodos exatos tornaria sua resolução impraticável. São procedimentos de propósito geral que podem explorar de forma eficiente o espaço de soluções de um problema específico de otimização.

A tendência das pesquisas em metaheurísticas é explorar a hibridização desses métodos. Uma forma de hibridização consiste na combinação de conceitos e estratégias de duas ou mais metaheurísticas clássicas. Outra abordagem seria a combinação de metaheurísticas com conceitos e processos de outras áreas de pesquisas que podem executar procedimentos específicos para melhorar o desempenho da metaheurística original. Um exemplo da segunda abordagem é a versão híbrida da metaheurística GRASP que incorpora um processo de mineração de dados, chamada DM-GRASP (Data Mining GRASP) [Santos et al. (2008)].

Desde que foi proposta, a metaheurística GRASP (*Greedy Randomized Adaptive Search Procedures*) [Feo e Resende (1989), Feo e Resende (1995)] tem sido aplicada com sucesso para resolver muitos problemas de otimização em áreas como escalonamento, roteamento, particionamento, localização e atribuição [Festa e Resende (2009)]. GRASP é de fácil implementação e é capaz de obter soluções muito boas em um tempo computacional aceitável. O processo de busca por soluções utilizado no GRASP é executado de forma iterativa e cada iteração é composta de duas fases: construção e busca local. Uma solução viável é gerada na fase de construção e, então, sua vizinhança é explorada pela busca local com a finalidade de encontrar soluções melhores. O resultado desse procedimento é a melhor solução encontrada entre todas as iterações.

Mineração de dados consiste na extração de conhecimento de bases de dados de forma automática [Han e Kamber (2006)]. As informações extraídas, na forma de padrões ou regras, representam características importantes da base de dados. Inicialmente, o processo de hibridização do GRASP com mineração de dados foi aplicado ao problema de empacotamento [Ribeiro et al. (2004)]. A hipótese básica utilizada foi que padrões extraídos de soluções de boa qualidade poderiam ser usados para guiar a busca no espaço de soluções, proporcionando uma exploração mais efetiva desse espaço. O método resultante, denominado DM-GRASP, obteve resultados promissores em termos de qualidade de solução e em termos do tempo necessário para obter soluções com boa qualidade. Posteriormente, o método foi avaliado em três outras aplicações: o problema de diversidade máxima [Santos et al. (2005)], o problema de replicação de servidores para *multicast* confiável [Santos et al. (2006)] e o problema das *p*-medianas [Plastino et al. (2009)], e os resultados foram igualmente bem sucedidos.

A primeira contribuição deste trabalho é mostrar que tanto a heurística GRASP tradicional quanto a heurística GRASP integrada com a técnica de reconexão por caminhos (um mecanismo que utiliza as melhores soluções das iterações anteriores para aprimorar as subsequentes) podem se beneficiar da incorporação de um módulo de mineração de dados.

A técnica de reconexão por caminhos foi originalmente proposta por Glover [Glover et al. (2000)] como uma estratégia de intensificação, explorando trajetórias que conectavam soluções elites obtidas por busca tabu ou *scatter search*. Nessa técnica, são percorridos caminhos a partir de soluções elite para outras soluções elite e, nesse percurso, soluções intermediárias são geradas possibilitando uma exploração eficiente do espaço de soluções.

Neste trabalho, apresentam-se duas estratégias híbridas com reconexão por caminhos (DM-GRASP-PR e MDM-GRASP-PR), que combinam, de forma diferente, o procedimento de mineração de dados com o GRASP integrado com a técnica de reconexão por caminhos. Mostra-se que essas estratégias podem melhorar a qualidade da solução e o tempo computacional do GRASP original com reconexão por caminhos.

A segunda contribuição deste trabalho é a aplicação da estratégia híbrida com reconexão por caminhos proposta para resolver o problema da síntese de redes a 2-caminhos (*2-path network design problem* – 2PNDP). Trata-se de um problema NP-difícil e muitas aplicações do mesmo podem ser encontradas em projetos de comunicação de redes, onde caminhos com poucos nós são mais confiáveis e produzem atrasos menores

[Ribeiro e Rosseti (2007), Ribeiro et al. (2007)]. O GRASP com reconexão por caminhos desenvolvido em [Ribeiro e Rosseti (2007)] obteve excelentes resultados para esse problema. Os experimentos computacionais realizados nesse trabalho mostraram que a implementação híbrida do GRASP com reconexão por caminhos conseguiu melhorar soluções obtidas pela heurística considerada estado da arte para o 2PNDP.

O restante deste artigo está organizado da seguinte maneira. Na Seção 2, são apresentados os principais conceitos da metaheurística GRASP e da técnica de reconexão por caminhos. A Seção 3 apresenta a estratégia híbrida DM-GRASP-PR desenvolvida para o 2PNDP e os resultados computacionais obtidos com essa estratégia são comparados ao obtidos com o GRASP tradicional com reconexão por caminhos. Na Seção 4, a estratégia MDM-GRASP-PR é descrita e os resultados computacionais são apresentados e comparados com os resultados da estratégia DM-GRASP-PR. Finalmente, na Seção 5, são apresentadas considerações finais e apontados alguns trabalhos futuros.

2. GRASP com Reconexão por Caminhos

GRASP [Resende e Ribeiro (2003)] é uma metaheurística que já foi aplicada com sucesso em muitos problemas de otimização [Festa e Resende (2009)].

Na primeira fase de uma iteração do GRASP, uma solução inicial é construída. Como essa solução não é garantidamente um ótimo local, um processo de busca local é executado na segunda fase e uma solução é gerada para o problema. Esse processo iterativo é repetido até que um critério de parada seja satisfeito e o resultado do algoritmo é a melhor solução encontrada entre as soluções geradas em todas iterações.

Na fase de construção, a solução se inicia com o conjunto vazio. Os elementos que não fazem parte da solução são ordenados de acordo com uma função gulosa e uma fração dos melhores elementos compõe uma lista restrita de candidatos. Um elemento é selecionado aleatoriamente dessa lista e é incorporado à solução. Esse processo é repetido até que seja construída uma solução viável para o problema.

A solução obtida na fase de construção é o ponto de partida da busca local na qual a vizinhança da solução é explorada. A vizinhança de uma solução é definida por uma função que relaciona essa solução a um conjunto de outras soluções. Se uma solução melhor for encontrada na vizinhança, a busca local começa novamente partindo do melhor vizinho encontrado. Caso contrário, a busca local termina.

Reconexão por caminhos [Glover et al. (2000)] é uma técnica proposta por Glover para explorar trajetórias compostas por soluções que conectam duas boas soluções encontradas por heurísticas.

A metaheurística GRASP é um processo que não utiliza memória, pois todas as iterações são independentes e nenhuma informação sobre as soluções são transmitidas de uma iteração para outra. O objetivo de usar a técnica de reconexão por caminhos durante o processo do GRASP é armazenar boas soluções encontradas e usá-las para guiar a busca por outras boas soluções. Laguna e Martí [Laguna e Martí (1999)] foram os primeiros a usar a reconexão por caminhos com GRASP. Algumas melhorias e aplicações dessa técnica podem ser encontradas na literatura [Resende e Ribeiro (2005)].

Basicamente, a reconexão por caminhos é aplicada a um par de soluções $\{s_i, s_g\}$ e, começando da solução inicial s_i , gradualmente incorpora os atributos da solução

guia s_g a s_i , até que s_i seja igual s_g , gerando soluções intermediárias nesse procedimento. Existem diferentes maneiras de se explorar os caminhos entre as duas soluções [Resende e Ribeiro (2005)]: reconexão *backward*, reconexão *forward*, reconexão bidirecional, reconexão periódica, reconexão aleatória e reconexão truncada.

Quando se utiliza o GRASP com reconexão de caminhos, deve-se manter um conjunto elite, onde as melhores soluções encontradas pelas iterações anteriores do GRASP são armazenadas. Neste trabalho, a reconexão por caminhos é executada em cada iteração do GRASP usando uma solução do conjunto elite e a solução encontrada após a execução da busca local na iteração corrente. A partir dessas duas soluções, definem-se as soluções inicial (s_i) e guia (s_g). O conjunto Δ , composto das posições nas quais s_i e s_g diferem, é calculado. A melhor solução e seu custo são determinados. A reconexão por caminhos é executada até que o caminho entre s_i e s_g seja totalmente percorrido. Para cada posição $m \in \Delta$, $s_i \oplus m$ é a solução obtida a partir de s_i pela troca da m -ésima posição por sua correspondente em s_g . Em seguida, a componente m^* de Δ em que a operação $s_i \oplus m$ resultou na solução de menor custo é obtida, m^* é removido de Δ e a solução atual é atualizada trocando o valor da m^* -ésima posição. Essa solução é mais parecida com a solução guia pois um elemento da solução inicial foi substituído por um da solução guia. Se essa nova solução tiver um custo melhor do que a melhor solução conhecida até então, a nova solução passa a ser a melhor solução. Essa nova solução será usada como solução inicial no próximo passo da reconexão por caminhos.

3. Proposta Híbrida DM-GRASP-PR

Nesta seção, será descrito o problema de projeto de redes a 2-caminhos e o GRASP com reconexão por caminhos desenvolvido por [Ribeiro e Rosseti (2007)] para resolver esse problema. Em seguida, será apresentada a heurística DM-GRASP-PR, que é uma versão híbrida da metaheurística GRASP com reconexão por caminhos apresentada em [Ribeiro e Rosseti (2007)] acrescida de um processo de mineração de dados.

Seja $G = (V, E)$ um grafo conexo não direcionado, onde V é o conjunto de vértices e E o conjunto de arestas. Um k -caminho entre dois vértices é uma sequência de no máximo k arestas que os conecta. Dada uma função de pesos não negativos $w : E \rightarrow R_+$, associada às arestas de G , e um conjunto D de pares de vértices origem-destino, o problema de projeto de redes a 2-caminhos consiste em encontrar um subconjunto de arestas $E' \subseteq E$ de menor peso que forneça um 2-caminho entre cada par origem-destino pertencente a D . O 2PNDP tem aplicações em projetos de redes de comunicações, onde caminhos com poucas arestas aumentam a confiabilidade e diminuem os atrasos. Dahl e Johannessen provaram que a versão de decisão do 2PNDP é NP-completo [Dahl e Johannessen (2004)]. Em [Ribeiro e Rosseti (2007)], os autores usaram GRASP com reconexão por caminhos para resolver o problema de forma aproximada.

3.1. GRASP-PR para o 2PNDP

Nesta seção, é apresentada a heurística GRASP com reconexão por caminhos (GRASP-PR) para o 2PNDP apresentado em [Ribeiro e Rosseti (2007)].

Cada iteração desse GRASP possui três fases: construção, busca local e reconexão por caminhos. A fase de construção gera uma solução x considerando uma função de pesos das arestas w' , que inicialmente é igual à função de pesos original w . O procedimento

é executado até que um 2-caminho seja encontrado para todo par origem-destino de D . Cada iteração começa selecionando aleatoriamente um par (a, b) para o qual será gerado um 2-caminho. Um menor caminho Pt de a para b é determinado usando a função de pesos w' . Em seguida, o par (a, b) é removido do conjunto de pares origem-destino e as arestas em Pt são inseridas na solução em construção. Para as iterações seguintes da fase de construção, a função w' é modificada de forma a atribuir o valor 0 às arestas em Pt , pois essas já fazem parte da solução.

Cada solução x é um conjunto $|D|$ de 2-caminhos. Na fase de busca local, dada uma solução x , suas soluções vizinhas são obtidas trocando algum 2-caminho em x por algum outro 2-caminho que exista entre a origem e o destino do mesmo par em x .

A fase de reconexão de caminhos é aplicada duas vezes. Na primeira vez, percorre-se o caminho partindo da solução obtida após a busca local e chegando em uma solução selecionada aleatoriamente do conjunto elite e, na segunda vez, percorre-se o caminho no sentido contrário. A solução encontrada após a busca local e a melhor solução encontrada na reconexão por caminhos são candidatas a entrarem no conjunto elite. Uma solução é inserida no conjunto elite se for diferente de todas as soluções no conjunto elite e seu custo for melhor do que a pior solução do conjunto elite.

3.2. Heurística DM-GRASP-PR

A hibridização da metaheurística GRASP com mineração de dados, denominada DM-GRASP, já foi aplicada com sucesso para resolver alguns problemas de otimização [Plastino et al. (2009), Santos et al. (2005), Santos et al. (2006), Santos et al. (2008)]. O DM-GRASP é dividido em duas fases. A primeira fase é chamada de geração do conjunto elite, que consiste em executar n iterações do GRASP puro. As d melhores soluções encontradas formam o conjunto elite.

Após essa primeira fase, é executado um procedimento de mineração de dados para extrair t padrões do conjunto elite. Os padrões a serem extraídos são conjuntos de elementos que aparecem no conjunto elite de forma frequente. Essa extração de padrões é realizada pela mineração de conjuntos frequentes [Han e Kamber (2006)]. Um conjunto frequente minerado com suporte s representa um conjunto de elementos que aparece em $s\%$ das soluções do conjunto elite.

Em seguida, é executada a segunda fase, chamada fase híbrida. Outras n iterações do GRASP adaptado são executadas. Nessas n iterações, a fase de construção adaptada constrói uma solução guiada por um padrão selecionado do conjunto de padrões minerados. Inicialmente, todos os elementos do padrão selecionado são inseridos na solução parcial e uma solução completa é construída a partir dessa solução parcial executando-se o procedimento de construção original.

Neste trabalho, foi desenvolvida a heurística híbrida DM-GRASP-PR, onde a mineração de dados é incorporada ao GRASP com reconexão por caminhos (GRASP-PR), com a intenção de mostrar que o GRASP com reconexão por caminhos – mecanismo de intensificação baseado em memória – pode se beneficiar da utilização de mineração de dados assim como o GRASP tradicional.

Os padrões a serem extraídos são conjuntos de arestas que comumente aparecem em soluções sub-ótimas do 2PNDP. Um conjunto frequente extraído do conjunto elite

com suporte s representa um conjunto de arestas que aparecem em $s\%$ das soluções elite. Um conjunto frequente é chamado maximal se não é subconjunto de outro conjunto frequente. Para evitar a extração de conjuntos frequentes que sejam subconjuntos de outros, na proposta DM-GRASP-PR para o 2PNDP, decidiu-se extrair apenas conjuntos frequentes maximais.

O algoritmo de construção adaptada é bem similar ao procedimento da fase de construção do GRASP, com a seguinte diferença: tenta-se construir um 2-caminho entre um par (a, b) utilizando apenas arestas do padrão ou arestas que já pertencem à solução parcial, que para isso têm seu peso modificado para 0. Se um 2-caminho não for encontrado utilizando apenas essas arestas, gera-se um 2-caminho, utilizando todas as arestas de E , partindo-se da solução parcial encontrada até o momento.

3.3. Resultados Computacionais do DM-GRASP-PR

Nesta seção, são comparados os resultados obtidos pelo GRASP-PR e DM-GRASP-PR. Foram geradas 25 instâncias similares às geradas em [Ribeiro e Rosseti (2007)]. Essas instâncias são grafos completos com $|V| \in \{100, 200, 300, 400, 500\}$. Os custos das arestas são gerados aleatoriamente com distribuição uniforme no intervalo $(0, 10]$ e $10 \times |V|$ pares origem-destino são escolhidos aleatoriamente. Os algoritmos foram implementados em C e compilados com gcc 4.4.1. Os testes foram executados em um processador 2.4 GHz Intel Core 2 Quad CPU Q6600 com 3 Gbytes de RAM, com Linux Kernel 2.6.24. Os procedimentos GRASP-PR e DM-GRASP-PR foram executados 10 vezes com diferentes sementes aleatórias em cada execução. Em cada uma dessas estratégias, foram executadas 1000 iterações. Após realizar alguns experimentos para calibragem dos parâmetros, os valores de d e t foram definidos ambos como 10, e o valor de s foi definido como 2.

Na Tabela 1, apresentam-se os resultados referentes à qualidade das soluções e ao tempo computacional. A primeira coluna representa o identificador $ax-y$ das instâncias, onde $x = |V|$ e y é a semente usada para gerar a instância aleatória. As colunas *Melhor*, *Média* e *Desv* representam, respectivamente, o custo da melhor solução, a média dos custos e o desvio padrão da média dos custos obtidos pelas estratégias. Os melhores resultados estão em negrito. Esses resultados mostram que a estratégia DM-GRASP-PR proposta conseguiu melhorar todos os resultados obtidos pelo GRASP-PR.

A coluna *Tempo* exibe a média dos tempos de execução (em segundos) das estratégias, obtidos em 10 execuções, e a coluna *TDesv* mostra seu desvio padrão. A última coluna mostra a diferença percentual entre os tempos médios obtidos pelas estratégias. Em todas as instâncias, o tempo de execução do DM-GRASP-PR foi menor. A última linha da tabela apresenta a média das diferenças percentuais. Pode-se observar que, na média, DM-GRASP-PR foi 20,23% mais rápido que o GRASP-PR.

Existem duas razões para o DM-GRASP-PR ter sido mais rápido. A primeira é o fato de que o esforço computacional dispendido na fase de construção adaptada é menor que na fase de construção original, dado que o conjunto de arestas avaliado para se encontrar um 2-caminho é menor. A segunda razão é que o uso de padrões proporcionou a construção de soluções de melhor qualidade. Essas soluções são utilizadas como soluções iniciais na busca local e, como possuem melhor qualidade, fazem com que a busca local convirja mais rapidamente para um ótimo local.

Tabela 1. Resultados de Qualidade e Tempo do GRASP-PR e DM-GRASP-PR

Instância	GRASP-PR					DM-GRASP-PR					Tempo
	Melhor	Média	Desv	Tempo	TDesv	Melhor	Média	Desv	Tempo	TDesv	%
a100-1	679	687,50	4,06	44,22	0,76	676	682,00	3,55	37,39	0,65	15,44
a100-10	663	669,80	3,25	43,29	0,58	662	668,70	2,83	36,14	0,54	16,51
a100-100	670	674,60	2,65	46,66	0,30	666	670,30	2,10	38,89	0,32	16,66
a100-1000	644	649,90	3,33	42,98	0,55	641	647,00	4,31	36,11	0,79	15,99
a100-10000	664	669,20	3,40	43,57	0,50	661	666,50	3,58	36,87	0,58	15,37
a200-1	1386	1391,90	4,66	201,30	2,57	1379	1384,60	3,80	161,87	1,77	19,59
a200-10	1374	1386,00	8,26	206,32	1,95	1362	1376,10	8,19	166,02	1,85	19,53
a200-100	1361	1369,40	4,27	197,35	2,71	1354	1362,00	4,80	157,37	1,96	20,26
a200-1000	1363	1374,50	7,77	199,61	2,42	1358	1367,90	8,63	158,63	2,35	20,53
a200-10000	1375	1387,40	8,66	207,02	2,20	1369	1377,50	7,57	166,49	1,75	19,58
a300-1	2106	2117,00	7,94	516,63	3,53	2081	2102,40	9,36	401,89	3,01	22,21
a300-10	2134	2148,00	6,88	515,14	3,47	2122	2133,70	7,89	401,34	4,00	22,09
a300-100	2088	2096,20	7,08	517,84	3,54	2072	2082,30	7,04	412,27	29,71	20,39
a300-1000	2100	2105,70	6,69	516,14	4,42	2080	2094,50	8,69	398,99	4,41	22,70
a300-10000	2077	2092,80	9,13	515,48	3,63	2067	2078,20	7,70	399,88	5,29	22,43
a400-1	2807	2816,20	5,33	1000,79	6,59	2788	2797,50	4,76	769,70	10,15	23,09
a400-10	2848	2864,70	9,24	1003,74	4,89	2833	2847,80	7,17	780,44	17,78	22,25
a400-100	2818	2834,20	7,40	1026,18	4,57	2803	2818,90	13,36	854,99	97,01	16,68
a400-1000	2822	2833,40	7,47	1022,98	3,21	2800	2816,40	10,28	824,99	78,34	19,35
a400-10000	2856	2874,80	9,47	1028,98	5,62	2844	2857,20	8,94	808,78	56,69	21,40
a500-1	3598	3606,60	6,62	1727,36	13,52	3571	3579,60	6,04	1330,40	28,97	22,98
a500-10	3595	3607,70	6,66	1712,67	6,40	3573	3580,70	7,55	1302,53	38,48	23,95
a500-100	3598	3612,40	10,46	1747,14	5,43	3576	3584,70	7,99	1396,26	125,10	20,08
a500-1000	3573	3592,00	7,69	1721,36	9,88	3554	3564,20	5,84	1332,65	26,15	22,58
a500-10000	3605	3625,00	10,85	1760,41	4,69	3580	3597,90	11,26	1337,25	20,10	24,04
Média											20,23

4. Proposta Híbrida MDM-GRASP-PR

Na proposta híbrida DM-GRASP-PR, a mineração de dados é executada apenas uma vez e no meio de todo o processo. Apesar de o resultado obtido ter sido satisfatório, acreditamos que realizar a mineração de dados mais de uma vez, sempre que o conjunto elite estiver suficientemente estável com soluções de boa qualidade, pode ocasionar uma melhora no desempenho da estratégia DM-GRASP. Baseando-se nessa hipótese, propõe-se e avalia-se, neste trabalho, uma outra versão do DM-GRASP-PR para o 2PNDP, chamada MDM-GRASP-PR (*Multi Data Mining GRASP-PR*).

A ideia principal dessa proposta é executar a mineração de dados: (a) assim que o conjunto elite se tornar estável – isto é, quando nenhuma mudança ocorrer durante um certo número de iterações – e (b) sempre que o conjunto elite for alterado e se tornar novamente estável. A hipótese explorada é que minerar mais de uma vez irá explorar a evolução gradual do conjunto elite e isso permitirá a extração de padrões mais refinados.

4.1. Resultados Computacionais

Nesta seção, são apresentados os resultados computacionais obtidos pela estratégia MDM-GRASP-PR usando o mesmo ambiente de execução da seção anterior. Após experimentos preliminares, definiu-se o número de iterações avaliadas para considerar o conjunto elite estável como 1% do número total de iterações.

Como na análise anterior os resultados obtidos pelo DM-GRASP-PR foram superiores aos obtidos pelo GRASP-PR, optou-se por comparar os resultados obtidos pela

nova proposta MDM-GRASP-PR somente com os resultados obtidos pela estratégia DM-GRASP-PR. Na Tabela 2, os resultados relacionados à qualidade e ao tempo computacional são apresentados. O MDM-GRASP-PR encontrou 18 melhores resultados no que diz respeito aos melhores valores e o DM-GRASP-PR encontrou quatro. O MDM-GRASP-PR encontrou 24 melhores resultados referentes às médias dos valores e o DM-GRASP-PR encontrou um. Esses resultados mostram que a proposta MDM-GRASP-PR conseguiu superar os resultados obtidos pelo DM-GRASP-PR.

Pode-se observar que o DM-GRASP-PR foi mais rápido em 19 instâncias e MDM-GRASP-PR foi mais rápido em 6. Porém, observa-se que o MDM-GRASP-PR foi, na média, apenas 1,34% mais lento que o DM-GRASP-PR, o que não é muito significativo. Pode-se concluir que ambas propostas híbridas tiveram um comportamento similar em termos de tempo computacional.

Tabela 2. Resultados do DM-GRASP-PR e MDM-GRASP-PR

Instâncias	DM-GRASP-PR					MDM-GRASP-PR					Tempo %
	Melhor	Média	Desv	Tempo	TDesv	Melhor	Média	Desv	Tempo	TDesv	
a100-1	676	682,00	3,55	37,39	0,65	674	681,90	5,28	38,50	0,91	-2,96
a100-10	662	668,70	2,83	36,14	0,54	659	665,20	3,22	37,54	1,38	-3,87
a100-100	666	670,30	2,10	38,89	0,32	667	670,00	2,41	40,41	0,83	-3,91
a100-1000	641	647,00	4,31	36,11	0,79	640	646,70	3,95	37,51	0,51	-3,89
a100-10000	661	666,50	3,58	36,87	0,58	658	665,40	3,56	38,41	1,13	-4,19
a200-1	1379	1384,60	3,80	161,87	1,77	1380	1383,90	4,16	163,19	6,41	-0,81
a200-10	1362	1376,10	8,19	166,02	1,85	1362	1372,50	5,80	167,06	3,42	-0,63
a200-100	1354	1362,00	4,80	157,37	1,96	1352	1360,70	6,63	162,58	6,85	-3,31
a200-1000	1358	1367,90	8,63	158,63	2,35	1356	1364,00	7,87	160,25	6,65	-1,02
a200-10000	1369	1377,50	7,57	166,49	1,75	1363	1374,30	7,85	166,61	6,77	-0,07
a300-1	2081	2102,40	9,36	401,89	3,01	2082	2099,30	9,23	409,38	12,47	-1,86
a300-10	2122	2133,70	7,89	401,34	4,00	2125	2132,10	5,05	410,17	12,15	-2,20
a300-100	2072	2082,30	7,04	412,27	29,71	2069	2076,30	5,40	404,22	10,37	1,95
a300-1000	2080	2094,50	8,69	398,99	4,41	2076	2090,30	7,09	395,88	18,55	0,78
a300-10000	2067	2078,20	7,70	399,88	5,29	2060	2075,10	10,38	403,97	14,56	-1,02
a400-1	2788	2797,50	4,76	749,77	10,15	2786	2791,40	4,52	749,77	21,24	2,59
a400-10	2833	2847,80	7,17	780,44	17,78	2819	2844,10	11,35	811,97	30,30	-4,04
a400-100	2803	2818,90	13,36	854,99	97,01	2803	2808,90	4,39	799,67	27,00	6,47
a400-1000	2800	2816,40	10,28	824,99	78,34	2793	2810,90	7,91	797,91	47,62	3,28
a400-10000	2844	2857,20	8,94	808,78	56,69	2793	2810,90	10,37	797,91	36,19	1,34
a500-1	3571	3579,60	6,04	1330,40	28,97	3567	3576,90	7,27	1349,39	73,55	-1,43
a500-10	3573	3580,70	7,55	1302,53	38,48	3566	3580,10	10,49	1346,80	86,90	-3,40
a500-100	3576	3584,70	7,99	1396,26	125,10	3572	3583,10	9,42	1413,63	65,44	-1,24
a500-1000	3554	3564,20	5,84	1332,65	26,15	3554	3564,90	4,95	1382,99	64,73	-3,78
a500-10000	3580	3597,90	11,26	1337,25	20,10	3573	3596,10	13,44	1420,02	114,77	-6,19
Média											-1,34

Para verificar se as diferenças entre os valores médios obtidos pelas estratégias possuem significância estatística, empregou-se a técnica teste t de Student não pareado. Comparando o DM-GRASP-PR com o GRASP-PR, observa-se que o DM-GRASP-PR encontrou melhores resultados em todas as 25 instâncias e, em 19 delas, com significância estatística, considerando *p-value* menor do que 0,01. Quando compara-se o MDM-GRASP-PR com GRASP-PR, o MDM-GRASP-PR também encontra resultados melhores em todas as 25 instâncias e, em 21 delas, com significância estatística. Esses resultados mostram a superioridade das estratégias com mineração de dados e, principalmente, o bom comportamento do MDM-GRASP-PR.

As Figuras 1(a) e 1(b) mostram outra comparação entre as três estratégias, baseada nos gráficos *time-to-target* (TTT) [Aiex et al. (2007)], que são usados para fazer análise de comportamento de algoritmos aleatórios. Esses gráficos basicamente mostram as distribuições de probabilidade acumulada de tempos de execução, ou seja, $p(\text{tempo_computacional} < x)$ vs. x .

Para gerar um gráfico TTT, executa-se um algoritmo algumas vezes e mede-se o tempo necessário para atingir uma solução com um valor melhor ou igual a um valor alvo. Nos experimentos apresentados neste trabalho, cada estratégia foi executada 100 vezes. Os tempos obtidos são ordenados de forma crescente e ao i -ésimo tempo t_i é associado o valor de probabilidade $p_i = (i - 1/2)/100$. Os pontos $z_i = (t_i, p_i)$, para $i = 1, \dots, 100$, são traçados no gráfico, indicando a probabilidade (eixo vertical) de uma estratégia atingir uma solução alvo em um tempo indicado (eixo horizontal). Os gráficos apresentados nas Figuras 1(a) e 1(b) foram gerados a partir das execuções do GRASP-PR, DM-GRASP-PR e MDM-GRASP-PR, para a instância a400-100, usando um alvo de dificuldade média (2834) e um alvo difícil (2820).

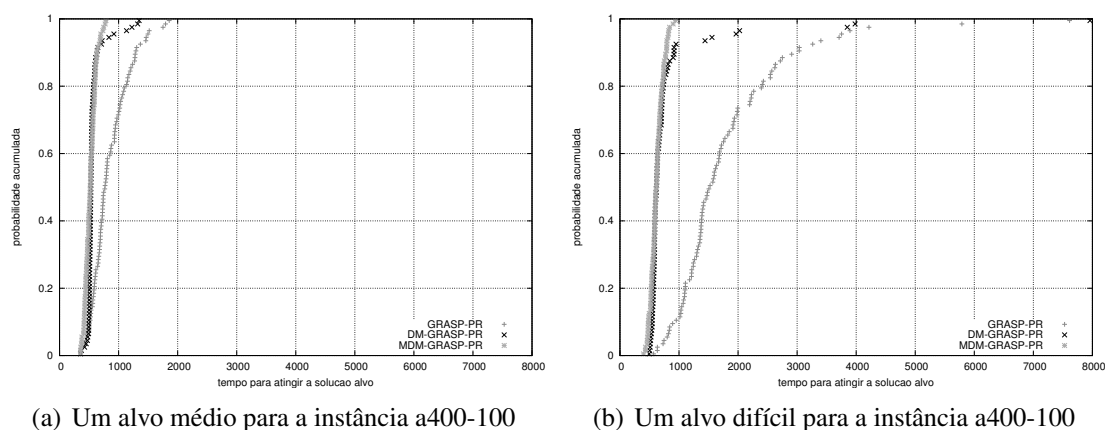


Figura 1. Gráficos Time-to-target

Para o alvo médio, pode-se observar na Figura 1(a) que o GRASP-PR apresentou o pior comportamento e que o MDM-GRASP-PR apresentou um melhor desempenho que o DM-GRASP-PR. Pode-se verificar que a probabilidade do MDM-GRASP-PR atingir o alvo médio em 800 segundos é 100%, do DM-GRASP-PR é aproximadamente 95% e do GRASP-PR é de aproximadamente 58%. Para o alvo difícil, a Figura 1(b) mostra que o MDM-GRASP-PR obteve o melhor comportamento. Esse gráfico indica que MDM-GRASP-PR consegue encontrar soluções difíceis mais rapidamente que o DM-GRASP-PR e muito mais rapidamente que o GRASP-PR, demonstrando que aplicar a mineração de dados mais de uma vez, sempre que o conjunto elite se torna estável, fornece maior robustez para a estratégia híbrida.

5. Conclusões

Metaheurísticas híbridas GRASP que incorporam processos de mineração de dados tem sido utilizadas com sucesso em diferentes problemas de otimização combinatória. Neste trabalho, propôs-se combinar mineração de dados com a metaheurística GRASP com reconexão por caminhos para mostrar que não só o GRASP pode se beneficiar do uso de

padrões para guiar sua busca, mas também o GRASP que utiliza a técnica de reconexão por caminhos.

Os resultados dos experimentos mostraram que a primeira versão da estratégia híbrida com reconexão por caminhos, chamada DM-GRASP-PR, conseguiu obter soluções melhores em um tempo computacional menor que o GRASP com reconexão por caminhos original desenvolvido para resolver o problema da síntese de redes a 2-caminhos, que era considerado o método estado da arte para esse problema.

Na primeira versão híbrida do GRASP com reconexão por caminhos, a mineração de dados ocorria apenas uma vez. Para explorar a evolução gradativa do conjunto elite de soluções e permitir a extração de padrões de melhor qualidade, foi proposta uma outra versão híbrida do GRASP com reconexão por caminhos, chamada MDM-GRASP-PR. Essa estratégia extrai novos conjuntos de padrões sempre que o conjunto elite é alterado e se torna estável. Os experimentos realizados mostraram que o MDM-GRASP-PR obteve resultados ainda melhores que os encontrados pelo DM-GRASP-PR.

Esses resultados mostraram que a incorporação da técnica de mineração de dados é eficiente não só para heurísticas que não fazem uso de memória, mas também para métodos que usam troca de informações sobre as soluções obtidas anteriormente como a estratégia de reconexão por caminhos.

6. Comentários

Este trabalho é parte de um projeto de pesquisa sobre metaheurísticas híbridas com mineração de dados. O estudante Hugo Barbalho desenvolveu em 2010, sob supervisão dos pesquisadores Simone Martins e Alexandre Plastino, as estratégias DM-GRASP-PR e MDM-GRASP-PR baseadas no GRASP-PR desenvolvido pela pesquisadora Isabel Rosseti na sua tese de doutorado. Uma versão estendida deste artigo foi aceito para publicação na edição especial *GRASP with Path Relinking* do periódico *Computers and Operation Research Journal*.

Referências

- R. Aiex, M. G. C. Resende e C. C. Ribeiro, *TTT plots: a perl program to create time-to-target plots*, *Optimization Letters*, 4 (2007), pp. 355–366.
- G. Dahl e B. Johannessen, *The 2-path network problem*, *Networks*, 43 (2004), pp. 190–199.
- T. A. Feo e M. G. C. Resende, *A probabilistic heuristic for a computationally difficult set covering problem*, *Operations Research Letters*, 8 (1989), pp. 67–71.
- T. A. Feo e M. G. C. Resende, *Greedy randomized adaptive search procedures*, *Journal of Global Optimization*, 6 (1995), pp. 109–133.
- P. Festa e M. G. C. Resende, *An annotated bibliography of GRASP Part II: Applications*, *International Transactions in Operational Research*, 16 (2009), pp. 131–172.
- F. Glover, M. Laguna e R. Martí, *Fundamentals of scatter search and path-relinking*, *Control and Cybernetics* 39, pp. 653–684, 2000.
- J. Han e M. Kamber, *Data Mining: Concepts and Techniques*, 2nd Ed., Morgan Kaufmann Publishers, 2006.

- M. Laguna e R. Martí, *GRASP and path relinking for 2-layer straight line crossing minimization*, *INFORMS Journal on Computing*, 11 (1999), pp. 44–52.
- A. Plastino, E. R. Fonseca, R. Fuchshuber, S. L. Martins, A. A. Freitas, M. Luis e S. Sahli, *A hybrid data mining metaheuristic for the p -median problem*, *Proceedings of the SIAM International Conference on Data Mining*, pp. 305-316, 2009.
- M. G. C. Resende e C. C. Ribeiro, *GRASP with path-relinking: Recent advances and applications*, *Metaheuristics: Progress as Real Problem Solvers* (T. Ibaraki, K. Nonobe e M. Yagiura, editores), (2005), 29–63.
- M. G. C. Resende e C. C. Ribeiro, *Greedy randomized adaptive search procedures*, *Handbook of Metaheuristics*, Kluwer Academic Publishers, 2003.
- C. C. Ribeiro, S. L. Martins e I. Rosseti, *Metaheuristics for optimization problems in computer communications*, *Computer Communications*, 30 (2007), pp. 656-669.
- C. C. Ribeiro e I. Rosseti, *Efficient parallel cooperative implementations of GRASP heuristics*, *Parallel Computing* 33 (2007), pp. 21-35.
- M. H. F. Ribeiro, V. F. Trindade, A. Plastino e S. L. Martins, *Hybridization of GRASP metaheuristic with data mining techniques*, *Proceedings of the ECAI Workshop on Hybrid Metaheuristics*, pp. 69-78, 2004.
- L. F. Santos, C. V. Albuquerque, S. L. Martins e A. Plastino, *A hybrid GRASP with data mining for efficient server replication for reliable multicast*, *Proceedings of the IEEE GLOBECOM Conference*, 2006.
- L. F. Santos, S. L. Martins e A. Plastino, *Applications of the DM-GRASP heuristic: A survey*, *International Transactions in Operational Research*, 15 (2008), pp. 387–416.
- L. F. Santos, M. H. F. Ribeiro, A. Plastino e S. L. Martins, *A hybrid GRASP with data mining for the maximum diversity problem*, *Proceedings of the International Workshop on Hybrid Metaheuristics, LNCC 3636*, pp. 116–127, 2005.