CNMAC2012
XXXIV CONGRESSO NACIONAL DE MATEMÁTICA APLICADA E COMPUTACIONAL
17 a 21 de setembro de 2012 – Águas de Lindóia/SP

ISSN 1984-8218

Seleção de Atributos com Novas Metaheurísticas na Teoria de Conjuntos Aproximativos

Alex Sandro Aguiar Pessoa,
INPE - Programa de Pós-graduação em Computação Aplicada (CAP)
12.227-010, São José dos Campos, SP
E-mail: asapessoa@gmail.com

Stephan Stephany
INPE – Laboratório Associado de Computação e Matemática Aplicada (LAC)
12.227-010, São José dos Campos, SP
E-mail: stephan@lac.inpe.br

Resumo: *Técnicas de seleção de atributos são aplicadas em algoritmos de aprendizagem de máquina para reduzir a dimensionalidade de uma base de dados com o objetivo de melhorar a qualidade dos resultados e o desempenho computacional. A Teoria dos Conjuntos Aproximativos é empregada em mineração de dados com ênfase no tratamento de informações incertas e imprecisas. No caso da classificação, esta teoria implicitamente calcula reduções de atributos, eliminando aqueles que são supérfluos. Neste contexto, o cálculo das reduções é tipicamente efetuado por um algoritmo genético, sendo que o presente trabalho propõe o uso inovador de algumas metaheurísticas para este cálculo. Os resultados apresentados mostram que foi possível obter reduções com menor cardinalidade, ou seja, com menor número de atributos, para algumas bases de dados comuns na literatura da área.*

Palavras-chave: *seleção de atributos, teoria dos conjuntos aproximativos, metaheurísticas.*

1. Introdução

A Teoria dos Conjuntos Aproximativos (TCA), do termo em inglês *Rough Set Theory*, foi proposta por Pawlak [1], vem se mostrando uma ferramenta muito eficiente e eficaz no tratamento de incerteza em bases de dados, que surgem com inexatidão, ruídos ou informações incompletas. A TCA vem se difundindo como técnica de mineração de dados, em particular na classificação [2]. Suas principais características são o bom formalismo matemático, facilidade de uso, não requerer informações adicionais, tais como grau de pertinência ou probabilidade *a priori* e compactação de bases de dados. Esta última característica é decorrente de dois conceitos que são inerentes à TCA: a relação de indiscernibilidade e a redução de atributos. Enquanto a indiscernibilidade diminui o número de objetos ou elementos de uma base de dados, a redução provê a diminuição do número de atributos ou variáveis redundantes, em perda de informações.

A indiscernibilidade é uma relação de equivalência que particiona o universo (conjunto finito contendo todos os elementos de uma base de dados) de acordo com classes. Estas classes são rótulos atribuídos a elementos que possuem determinadas características em comum, expressas por meio de um atributo de decisão.

Para cada classe e para o conjunto de atributos considerado, particiona-se o universo em 3 regiões: uma região interna composta por objetos que certamente pertencem à classe, uma região externa, de elementos que certamente não pertencem a ela e uma região de borda ou fronteira, de elementos que, embora indiscerníveis em relação a esse conjunto de atributos, estão rotulados ou não como pertencentes à classe considerada. Considerando-se a região interna, todos os elementos pertencem à classe considerada para o conjunto de atributos definido, mas pode-se subdividi-los em classes de equivalência em função dos valores de seus atributos. Assim, um único elemento pode representar todos os demais de sua classe de equivalência, reduzindo significativamente o número de objetos da base de dados.

ISSN 1984-8218

A redução de atributos em TCA é outro conceito que compacta a base de dados através da eliminação de atributos supérfluos. Embora esteja implícita na TCA, a redução ou seleção de atributos na classificação é também pesquisada em outras abordagens.

Uma redução constitui um subconjunto de atributos que preserva a indiscernibilidade dos objetos em relação às classes definidas, ou seja, que permite obter a mesma classificação que seria conseguida com o conjunto completo de atributos. Entretanto o cálculo das reduções não é uma tarefa trivial, sendo considerado como um problema NP-difícil [1], agravado por bases de dados cada vez maiores. Assim, a computação das reduções na TCA é normalmente feito empregando-se uma metaheurística, como um algoritmo genético, ou outras, como será exposto a seguir.

Em Hedar et al. [3] implementou-se a metaheurística de Busca Tabu no cálculo de reduções em TCA, comparando-a com outras três metas-heurísticas (colônia de formigas, recozimento simulado e algoritmo genético). Esse estudo motivou o presente trabalho, que implementou de maneira inédita três metaheurísticas para o cálculo das reduções em TCA: *Variable Neighborhood Search* (VNS), *Variable Neighborhood Decent* (VND) e *Decrescent Cardinality Search* (DCS). As duas primeiras metaheurísticas são bem conhecidas na literatura, enquanto que a última foi criada para este trabalho.

2. Metodologia

2.1. Redução de Atributos na Teoria dos Conjuntos Aproximativos (TCA)

O conhecimento em TCA é representado na forma tabular, representado pelo par ordenado $S = (U; A)$, chamado de *sistema de informação*, onde U é um conjunto finito não-vazio de objetos chamado de universo e A é um conjunto finito não-vazio de atributos condicionais ou condições, tal que $a: U \rightarrow V_a$ para todo $a \in A$. O conjunto V_a é chamado de conjuntos de valores de a. Uma forma particular do sistema de informação é adicionando um atributo distinto dos atributos condicionais, com o objetivo de criar classes. Essa forma do sistema de informação é chamada de *sistema de decisão*, onde $S = (U; A \cup \{d\})$ e $d \notin A$ é o atributo de decisão.

O modo em que a TCA enxerga os dados e suas incertezas é através da indiscernibilidade, que é uma relação de equivalência que cria partições do universo U, de acordo com as incertezas associadas aos dados. Segue a definição formal:

Dado $S = (U; A)$ como sistema de informação, então com qualquer $B \subseteq A$ existe uma relação de equivalência $\mathrm{IND}_A(B)$, ou simplesmente $IND(B)$:

$$IND(B) = \{(x, x') \in U \mid \forall a \in B, a(x) = a(x')\} \qquad (1)$$

$IND(B)$ é chamada de relação de *B-indiscernibilidade*. Se $(x, x') \in IND(B)$, então objetos x e x' são indiscerníveis ou iguais a qualquer atributo de B. A classe de equivalência da relação determinada por x pertencente a $X \subseteq U$ é denotado $[x]_B$ [1]. As partições criadas por $[x]_B$ no universo U são denotadas por $U/IND(B)$.

Conforme discutido na seção anterior, para cada classe e para o conjunto de atributos considerado, há 3 regiões ou aproximações determinadas pela relação de indiscernibilidade. Para o cálculo das reduções ou a seleção de atributos, será abordada apenas a aproximação inferior, que é constituída por objetos que certamente pertencem a uma determinada classe.. Seja $B \subseteq A$, $x \in X$ e $X \subseteq U$, a aproximação inferior é dada por:

$$\underline{B}X = \{x \mid [x]_B \subseteq X\} \quad (2)$$

Para ilustrar os conceitos mostrados acima é tomado como exemplo o sistema de decisão da Tabela 1, onde os atributos condicionais são $A=\{a,b,c\}$, o atributos de decisão $\{d\}$ e o conjunto universo $U = \{e_1,e_2,e_3,e_4,e_5,e_6\}$. Se $B = \{a,b\}$ então os objetos e_1, e_2 e e_3 são indiscerníveis, assim como os objetos e_4 e e_6. Então $IND(B)$ tem as seguintes partições:

$$U/IND(B) = \{\{e_1,e_2,e_3\},\{e_4,e_6\},\{e_5\}\}$$

ISSN 1984-8218

Para a classe "1", o conjunto $X_1 = \{x \in X \mid d(x) = 1\}$ é igual a $X_1 = \{e_1,e_4,e_5\}$ e com a aproximação inferior $\underline{B}X_1 = \{e_5\}$. De forma análoga, para a classe "0", $X_0 = \{e_2,e_3,e_6\}$, com $\underline{B}X_0 = \varnothing$.

U	*a*	*b*	*c*	*d*
e_1	0	0	0	1
e_2	0	0	1	0
e_3	0	0	2	0
e_4	1	0	0	1
e_5	1	1	1	1
e_6	1	0	2	0

Tabela 1: Conjunto de dados

A aproximação inferior é usada para se calcular o chamado grau de dependência que avalia se um conjunto de atributos *B* tem informação mútua em relação a outro conjunto *C*, ou seja, em termos de informação, o quanto um conjunto de atributos contribui no sistema de informação. Sejam *IND*(*B*) e *IND*(*C*) relações de indiscernibilidade em *U* e os conjuntos de atributos *B* e *C*, o grau de dependência entre *B* e *C* é dado por:

$$\gamma_B(C) = \frac{|POS_B(C)|}{|U|} \quad (3)$$

onde $POS_B(C) = \bigcup_{\Theta} \underline{B}\Theta$, com $\Theta \in U / IND(C)$, é conhecida como região positiva, $|POS_B(C)|$ e $|U|$, denotam a cardinalidade ou número de elementos da região positiva e do conjunto universo, respectivamente. Se $\gamma_B(C) = 1$ então *C* é totalmente dependente de *B*. Se $\gamma_B(C) < 1$ diz-se que *C* e parcialmente dependente de B. Voltando novamente para o exemplo da Tabela 1 e tomando como $B = \{a,b\}$ e $C=\{d\}$, tem-se:

$$U/IND(C) = \{\{e_1,e_4,e_5\},\{e_2,e_3,e_6\}\}$$

A região positiva é dada por e o grau de dependência são dados, respectivamente, por:

$$POS_B(C) = POS_{\{a,b\}}(\{d\}) = \text{U}\ \{\{e_5\},\varnothing\} = \{e_5\}$$

$$\gamma_B(C) = \gamma_{\{a,b\}}(\{d\}) = \frac{|POS_{\{a,b\}}(\{d\})|}{|U|} = \frac{1}{6}$$

Uma das formas de aplicação da TCA é a seleção de atributos, que consiste na eliminação de atributos considerados redundantes ou supérfluos. Considerando-se o conjunto completo de atributos *A*, uma determinada redução *B* pode ser avaliada pelo seu grau de dependência (Eq. 3) em relação ao conjunto *C* composto unicamente pelo atributo de decisão $\{d\}$. Assim, denotando-se *R* uma redução, tem-se a seguinte definição:

$$R = \{B : B \subseteq A \mid \gamma_B(\{d\}) = \gamma_A(\{d\})\}$$

Então, finalmente, as reduções podem ser expressas pelos mínimos subconjuntos de *R*, dados por:

$$R_{\min} = RED = \{T \in R : |T| \le |S|, \forall S \in R\}$$

Abaixo seguem os graus de dependências para todos os subconjuntos de $B = \{a,b,c\}$:

$$\gamma_B(\{d\}) = 1,$$
$$\gamma_{\{a,b\}}(\{d\}) = 1/6,\ \gamma_{\{b,c\}}(\{d\}) = 1,\ \gamma_{\{a,c\}}(\{d\}) = 1$$
$$\gamma_{\{a\}}(\{d\}) = 0,\ \gamma_{\{b\}}(\{d\}) = 1/6,\ \gamma_{\{c\}}(\{d\}) = 1/6$$

Desta forma as possíveis reduções, de cardinalidade mínima (no caso 2), para o sistema de decisão apresentado na Tabela 1 são:

$$R_{\min} = \{\{a,c\},\{b,c\}\}$$

ISSN 1984-8218

Assim a Tabela 1 pode ser representada tanto pela Tabela 2(i) como pela Tabela 2(ii):

(i)

U	*a*	*c*	*d*
e_1	0	0	1
e_2	0	1	0
e_3	0	2	0
e_4	1	0	1
e_5	1	1	1
e_6	1	2	0

(ii)

U	*b*	*c*	*d*
e_1	0	0	1
e_2	0	1	0
e_3	0	2	0
e_4	0	0	1
e_5	1	1	1
e_6	0	2	0

Tabela 2: Sistemas de decisão reduzidos

2.2. Metaheurísticas Propostas

Em TCA as metaheurísticas são utilizadas, tipicamente, para o cálculo das reduções e consequentemente para a seleção de atributos. Outros trabalhos utilizaram metaheurísticas tais como algoritmo genético, colônia de formigas, recozimento simulado ou busca tabu, com a mesma finalidade. Metaheurísticas são métodos computacionais em que o problema é formulado iterativamente e soluções candidatas são refinadas como num problema de otimização. Metaheurísticas são geralmente aplicadas a problemas para os quais não há soluções conhecidas, como na caso de problemas NP-difíceis. Aborda-se a seguir as metaheurísticas utilizadas neste trabalho, a VNS (*Variable Neighbourhood Search*), a VND (*Variable Neighbourhood Descent*), e a DCS (*Descrecent Cardinality Search*).

As soluções codificadas deste trabalho representam subconjuntos de atributos e são representadas na forma de vetores de 0's e 1's, onde cada elemento do vetor é relativo a um atributo. O dígito 0 representa a ausência do atributo e o dígito 1, a sua presença. Por exemplo: dado o conjunto de atributos condicionais A = {*umidade*, *pressão*, *temperatura*, *precipitação*}. A cadeia A^* ={0,1,1,0} representa uma solução contendo somente os atributos *pressão* e *temperatura*. Uma solução vizinha de uma dada solução seria obtida pela permutação de um elemento (de 0 para 1 ou vice-versa).

Nos pseudocódigos abaixo, a função objetivo, ou medida de avaliação, denotada por f(s), avalia se uma solução candidata, dada por um subconjunto de atributos constitui uma redução. Aqui, esta função é dada pelo grau de dependência (Eq. 3), aplicada para a solução B em relação ao conjunto C constituído unicamente pelo atributo de decisão. Nestes pseudocódigos, $|\cdot|$ denota cardinalidade de um conjunto, ou seja, o número de membros do conjunto. Assim, s denota uma solução candidata e $|s|$ é seu número de atributos. Da mesma forma, N_s denota o conjunto de vizinhos de s, sendo o número de vizinhos expresso por $|N_s|$.

O VNS ou busca com vizinhança variável, é uma metaheurística baseada num procedimento iterativo que compreende a geração aleatória de um novo vizinho a partir de uma solução inicial ou corrente e a busca local que corresponde à exploração da vizinhança deste novo vizinho, cuja melhor solução passa a ser a solução corrente. Isto permite explorar gradativamente vizinhanças mais distantes [4], [5]. O pseudocódigo do VNS é mostrado abaixo:

```
Seja s uma solução inicial e Ns o conjunto de vizinhos de s

enquanto (critério de parada não satisfeito) faça
    k • 1;
    enquanto (k • |Ns|) faça
       s' • Gera um vizinho aleatório de s ∈ Ns
       s'' • Aplica Busca Local em s';
       se f(s'') < f(s) então
             s • s''; k • 1;
       senão
             k • k +1
    fim_enquanto
fim_enquanto
retorne s
fim VNS
```

Figura 1: Pseudocódigo do VNS

ISSN 1984-8218

O VND ou busca com vizinhança variável em descida é uma variante do VNS, ou, mais especificamente, um caso particular [4]. Este algoritmo explora a vizinhança de uma solução inicial, buscando o mínimo local através do gradiente. O pseudocódigo é mostrado abaixo:

```
Seja s uma solução inicial e Ns o conjunto de vizinhos de s
k ← 1;

enquanto (k ≤ |Ns|) faça
    s' ← Encontre o melhor vizinho de s ∈ Ns
    se f(s') < f(s) então
       s = s'
       k = 1
    senão
       k = k +1
fim_enquanto
retorne s
fim VND
```

Figura 2: Pseudocódigo do VND

O DCS ou busca de cardinalidade decrescente, é uma nova metaheurística proposta neste trabalho, derivada do VNS, cujo enfoque, diferentemente das outras metaheurísticas aqui abordadas, que buscam soluções melhores na vizinhança próxima, é buscar uma nova solução candidata (s') qualquer que tenha necessariamente menor cardinalidade do que a solução corrente. Segue o pseudocódigo do algoritmo proposto:

```
Seja s uma solução inicial

enquanto (critério de parada não satisfeito) faça
    s' • Gera um vizinho aleatório de s com |s'| = |s|-1
    s'' = Aplica Busca Local em x';
    se f(s'') < f(s) então
       s • s''
fim_enquanto
retorne s
fim DCS
```

Figura 3: Pseudocódigo do DCS

O pseudocódigo de busca local utilizado nas metaheurísticas VNS e DCS é mostrado abaixo:

```
Seja s uma solução inicial e |A| a cardinalidade do conjunto de
atributos condicionais

s_local = s
para i = 1 até i • |A| faça
     s' • muda um bit na posição i de s
     se f(s') < f(s) então
         s_local • s'
fim_para
retorne s_local
fim Busca_Local
```

Figura 4: Pseudocódigo da busca local

CNMAC2012
XXXIV CONGRESSO NACIONAL DE MATEMÁTICA APLICADA E COMPUTACIONAL
17 a 21 de setembro de 2012 – Águas de Lindóia/SP

ISSN 1984-8218

3. Resultados

A aplicação das metaheurísticas propostas (VNS, VND e DCS) na redução de atributos em TCA foi testada para as bases de dados utilizadas em Hedar et al. [3] e comparada aos resultados obtidos nesse trabalho para outras metaheurísticas: Busca Tabu, colônia de formigas, recozimento simulado e algoritmo genético, sendo a primeira delas a metaheurística proposta pelo autor. Essas bases de dados são descritas na Tabela 3, onde $|A|$ denota a cardinalidade do conjunto de atributos e $|U|$ o número de objetos/elementos em cada base de dados.

Conj. dados	$\|A\|$	$\|U\|$
Vote	16	300
Credit	20	1000
Mushroom	22	8124
Derm	34	366
Lung	56	32

Tabela 3: Conjunto de dados

Neste trabalho, de forma a se comparar o desempenho das metaheurísticas com os resultados anteriores supracitados, cada metaheurística foi executada 20 vezes com soluções iniciais diferentes que foram geradas aleatoriamente. Utilizou-se um computador pessoal com 4 núcleos de processamento.

Abaixo, na Tabela 4 são mostrados os resultados para o problema de seleção de atributos utilizando a TCA com as diversas metaheurísticas abordadas: colônia de formigas (Ant), recozimento simulado (SA), algoritmo genético (AG), Busca Tabu (TS), busca com vizinhança variável (VNS), busca com vizinhança variável em descida (VND), e busca com cardinalidade decrescente (DCS). Os resultados de Ant, SA, AG e TS são provenientes de Hedar et al. [3] servindo de referência para as metaheurísticas VNS, VND e DCS, utilizadas neste trabalho.

Dados	$\|A\|$	TCA – Seleção de Atributos						
		Ant	**SA**	**AG**	**TS**	**VNS**	**VND**	**DCS**
Vote	16	$8^{(20)}$	$8^{(15)}9^{(5)}$	$8^{(2)}9^{(18)}$	$8^{(20)}$	$8^{(20)}$	$5^{(1)*}6^{(1)*}7^{(2)*}8^{(3)}$ $9^{(6)}10^{(4)}11^{(3)}$	$8^{(3)}9^{(13)}$ $10^{(4)}$
Credit	20	$8^{(12)}9^{(4)}$ $10^{(4)}$	$8^{(18)}9^{(1)}$ $11^{(1)}$	$10^{(6)}$ $11^{(14)}$	$8^{(13)}9^{(5)}$ $10^{(2)}$	$7^{(17)}8^{(3)}$	$5^{(1)*}9^{(4)}10^{(9)}11^{(5)}$ $12^{(1)}$	$7^{(5)}8^{(15)}$
Mushroom	22	$4^{(20)}$	$4^{(20)}$	$5^{(1)}6^{(5)}$ $7^{(14)}$	$4^{(17)}5^{(3)}$	$3^{(4)}4^{(15)}$ $5^{(1)}$	$8^{(9)}9^{(6)}10^{(4)}14^{(1)}$	$3^{(14)}4^{(6)}$
Derm	34	$6^{(17)}7^{(3)}$	$6^{(12)}7^{(8)}$	$10^{(6)}$ $11^{(14)}$	$6^{(14)}7^{(6)}$	$10^{(1)}11^{(19)}$	$11^{(1)}14^{(3)}15^{(5)}$ $16^{(8)}17^{(2)}18^{(1)}$	$10^{(5)}11^{(12)}$ $12^{(3)}$
Lung	56	$4^{(20)}$	$4^{(7)}5^{(12)}$ $6^{(1)}$	$6^{(8)}7^{(12)}$	$4^{(6)}5^{(13)}$ $6^{(1)}$	$4^{(10)}5^{(10)}$	$18^{(2)}19^{(3)}21^{(2)}$ $22^{(6)}24^{(2)}25^{(3)}$ $27^{(1)}33^{(1)}$	$3^{(7)}4^{(12)}$ $5^{(1)}$

Tabela 4: Resultados

onde $Q^{(n)}$, Q denota a cardinalidade da solução e n número de vezes que soluções com tal cardinalidade foi encontrada. * refere-se a soluções com graus de dependência menor do que 1 ($\gamma_B < 1$). Na Tabela 5 são mostrados os tempos médios gasto pelas metaheurísticas propostas neste trabalho.

Dados	$\|A\|$	VNS	VND	DCS
Vote	16	495	0,5	21
Credit	20	1825	4	448
Mushroom	22	559	6	1458
Derm	34	298	1	298
Lung	56	473	0,7	55

Tabela 5. Tempo médio de execução (segundos)

ISSN 1984-8218

4. Comentários Finais

Este trabalho aborda o uso das reduções, um conceito da TCA, aplicadas ao problema de seleção de atributos. Neste contexto, foi proposta a aplicação inovadora de três metaheurísticas, VNS, VND e DCS. Anteriormente, Hedar e al. [3] havia proposto a Busca Tabu na redução de atributos em TCA como alternativa a outras metaheurísticas, no caso, colônia de formigas, recozimento simulado e algoritmo genético. Este trabalho utiliza as mesmas bases de dados deste trabalho para comparar o desempenho obtido pelas novas metaheurísticas propostas com as anteriores.

Os resultados mostram que o uso das metaheurísticas propostas na redução de atributos apresentou desempenhos similares, mas em muitos casos, superiores às demais metaheurísticas, para as bases de dados consideradas. Este desempenho refere-se à obtenção de reduções com menor número de atributos, ou seja, com menor cardinalidade. Além disso, os tempos de processamentos das metaheurísticas propostas foram também analisados. Embora não tenha sido possível comparar esses tempos com aqueles referentes às demais metaheurísticas, pode-se inferir que sejam melhores, uma vez que, tipicamente, métodos determinísticos de busca demandam menos tempo de processamento que os estocásticos. Naturalmente, há toda uma discussão acerca dos primeiros métodos serem mais propensos a encontrar mínimos locais, mas isso não pareceu se refletir nos resultados apresentados.

Analisando as reduções obtidas, nota-se que, em particular, a metaheurística VNS mostrou-se mais robusta, em termos de obter boas reduções com regularidade. Por outro lado, em comparação à VNS, a metaheurística VND sempre apresentou menor tempo de processamento, embora suas reduções geralmente sejam piores, apresentando maior cardinalidade. Finalmente, a metaheurística DCS, criada neste trabalho, foi derivada da VNS pela adoção de um esquema de busca mais agressivo, que busca somente soluções vizinhas com menor cardinalidade. Isso resultou em reduções com baixa cardinalidade, às vezes melhores que as obtidas por qualquer outra metaheurística. Os tempos de processamento da DCS foram comparáveis àqueles da VNS, o que demonstra ser esta a metaheurística mais apropriada, dentre as três sugeridas neste trabalho.

Agradecimentos

O autor Alex Sandro Aguiar Pessoa agradece ao CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico pelo auxílio financeiro, sob a forma de bolsa de doutorado (processo n. 140161/2010-4).

Referências

[1] J. Komorowski, L. Polkowski, A. Skowron, "Rough sets: A tutorial". in: S.K. Pal and A. Skowron (eds.), Rough fuzzy hybridization: A new trend in decision-making, Springer-Verlag, Singapore, 1999.

[2] A. S. A. Pessoa, S. Stephany, L. M. G. Fonseca, "Feature selection and image classification using rough sets theory". In: IEEE International Geoscience and Remote Sensing Symposium (IGARSS-2011), 2011, Vancouver. Symposium Proceedings of IGARSS-2011, 2011. v. 1. p. 2904-2907.

[3] A. Hedar, J. Wang; M. Fukushima, "Tabu search for attribute reduction in rough set theory", Technical Report 2006-008, Department of Applied Mathematics and Physics, Kyoto University. 2006

[4] N. Mladenovic, P. Hansen, "Variable Neighborhood Search". Computers and Operations Research, 24:1097-1100, 1997.

[5] P. Hansen, N. Mladenovic, "A Tutorial on Variable Neighborhood Search", Les Cahiers du GERAD, HEC Montreal and GERAD, 2003.