

UM ALGORITMO EVOLUCIONÁRIO PARA UM PROBLEMA DE CONFIGURAÇÃO DE REDES DE DISTRIBUIÇÃO DE GÁS NATURAL.

Cristine C. Schmidt¹ (UFRN), Elizabeth F. G. Goldbarg² (UFRN), Marco C. Goldbarg³(UFRN)

¹cristines@gmail.com

²beth@dimap.ufrn.br

³gold@dimap.ufrn.br

O problema da configuração de redes de distribuição de gás natural requer a definição de um traçado por onde dutos devem ser colocados, para que a empresa fornecedora do serviço possa atender a um determinado conjunto de clientes. A definição deste traçado é de extrema importância na construção destas redes, pois, se o traçado for planejado de maneira eficiente, a empresa pode atender a demanda tendo um baixo custo de construção da rede. Problemas como este podem ser representados computacionalmente de diversas maneiras. Este trabalho apresenta uma modelagem onde o foco é o atendimento de clientes cuja demanda por gás natural compense a construção de uma infra-estrutura para atendê-los. Através deste enfoque, a presente proposta proporciona ao fornecedor de serviços uma visão econômica diferenciada. Enquanto as abordagens tradicionais se preocupam em atender a todos os clientes em uma rede, esta procura atender apenas a clientes que são financeiramente atrativos. A modelagem utilizada é baseada no problema da árvore de Steiner com coleta de prêmios, que é uma generalização do conhecido Problema da Árvore de Steiner. Este é um problema de otimização combinatória da classe dos NP-Árduos. A solução exata deste problema é inviável para casos reais, onde o número de clientes pode ser muito grande. Foram encontradas na literatura apenas duas abordagens metaheurísticas para resolver este problema. Este trabalho apresenta uma abordagem metaheurística baseada em Algoritmos Transgenéticos (ATs), que se enquadram na categoria dos algoritmos evolucionários. A metáfora dos ATs propõe um processo endossimbiótico de troca de informações genéticas, segundo a teoria da endossimbiose. Esta teoria foi popularizada por Lynn Margulis, e diz que um novo organismo pode surgir da fusão de dois ou mais seres independentes. Além disso, também são consideradas nos ATs outras transformações do fluxo intracelular biológico. No algoritmo, uma população de cromossomos interage com uma população de vetores transgenéticos, segundo um conjunto de regras. Para a geração da população inicial foi utilizado um algoritmo primal-dual. Path-relinking foi utilizado como intensificador na busca, caracterizando a hibridização do algoritmo. Foram realizados experimentos computacionais comparando o algoritmo proposto com as outras duas abordagens baseadas em metaheurísticas encontradas na literatura. Na comparação foram utilizadas 114 instâncias, e os resultados mostram que a abordagem proposta é eficiente na resolução do problema.

Redes urbanas de distribuição, Gás natural, Algoritmo evolucionário, Árvore de Steiner com coleta de prêmios.

1. INTRODUÇÃO

O gás natural é um combustível fóssil com constituição semelhante a do petróleo, e pode ser encontrado no subsolo ou no fundo do mar, em depósitos naturais de rochas sedimentares, associado ou não ao petróleo. Sua origem é da decomposição de plantas e animais, resultado de um processo de milhões de anos. Nos últimos 40 anos, pelas grandes vantagens que oferece quando comparado com outras fontes primárias de energia, o gás natural tem se transformado em uma alternativa energética em diversos países do mundo. É utilizado nos setores industrial, residencial, comercial, na geração de energia e nos transportes. O gás natural promove no Brasil uma revolução de consumo, devido ao aumento da demanda pelo crescente número de indústrias e empresas usuárias. Além das vantagens de sua utilização para o meio ambiente, a sua grande disponibilidade como energia irá promover um consumo ainda maior durante os próximos anos.

Para que esta fonte de energia chegue até seus consumidores finais são necessários diversos processos, entre eles o processo de transporte. A melhor opção em termos econômicos para o transporte de gás natural no estado gasoso são os gasodutos, Borelli *et al.* (ANO). Eles são os responsáveis por levar o gás desde seu ponto de produção até as empresas distribuidoras (públicas ou privadas), que distribuem o gás através de linhas ou redes urbanas de distribuição, as quais são as conexões até o consumidor (residência, comércio ou indústria).

O projeto de uma rede urbana de distribuição de gás natural envolve, entre outros problemas, o da definição de um *layout* para a rede, ou seja, um traçado físico por onde os dutos serão colocados, levando em conta questões econômicas e ambientais. Sendo a tubulação, o componente de maior custo do sistema, Borelli *et al.* (2001), são estudadas formas de diminuir o consumo deste material.

Em um cenário real, uma empresa que deseja construir uma rede de distribuição de gás possui uma estimativa de demanda do seu produto. Deste modo, ela pode estimar o lucro advindo de atender cada cliente em

potencial. A empresa pode, então, avaliar a relação entre o benefício e o custo da instalação da rede uma vez que a mesma possua uma estimativa destes custos.

Diversos modelos podem ser utilizados para tratar o problema da definição do traçado da rede. Nesse trabalho, optou-se por modelar o problema considerando o custo de uma configuração em relação ao atendimento da demanda. Esta abordagem proporciona uma vantagem ao fornecedor de serviços, pois existe a possibilidade de os clientes serem selecionados pela companhia de acordo com o lucro esperado com a adição dos mesmos. Nesta abordagem, apenas clientes financeiramente atrativos são adicionados à rede. O objetivo é traçar uma rede conectando um subconjunto de clientes, maximizando o lucro e minimizando os custos de construção da rede.

O problema do traçado pode ser modelado através da Teoria dos Grafos. Considerando um grafo $G = (V, E)$ conexo não direcionado, $V = \{v_1, \dots, v_n\}$ é o conjunto de nós (clientes e nós de passagem), $E = \{e_1, \dots, e_m\}$ é o conjunto de arestas (ruas ou vias), $c: E \rightarrow \mathbb{R}^+$ uma função custo não negativa associada às arestas e $p: V \rightarrow \mathbb{R}^+$ uma função prêmio não negativa associada aos nós, o objetivo é definir uma estrutura de custo mínimo em relação ao comprimento dos trechos de tubulação e o atendimento da demanda. Desse modo, o objetivo do problema é encontrar uma sub-árvore $T = (V_T, E_T)$ de G que minimize a soma dos custos das arestas mais a soma dos prêmios dos nós que não estão na árvore. A função objetivo adotada nesse trabalho é conhecida como função de minimização de Goemans e Williamson, GW , e é apresentada na equação (I).

$$GW = \sum_{e \in E_T} c(e) + \sum_{v \notin V_T} p(v) \quad (I)$$

Neste trabalho serão consideradas redes do tipo ramificadas, onde existe um único caminho ligando um nó v a qualquer outro nó w da rede. Na modelagem, os clientes são vistos como nós de demanda do grafo, onde o valor da demanda (ou prêmio) é o lucro estimado pela empresa em atender ao cliente. Intersecções de ruas são vistas como nós de passagem do grafo, cuja demanda é zero. As ruas por onde os dutos serão colocados são vistos como as arestas do grafo, cujos pesos são os custos estimados dos dutos.

A modelagem computacional adotada para o problema é conhecida como problema da árvore de Steiner com coleta de prêmios – PSP, sendo este um problema NP-Árduo, Karp (1972). Por se tratar de um problema combinatório NP-árduo, escolheu-se uma abordagem metaheurística para tratá-lo. A abordagem utilizada para a solução do problema se baseia em duas técnicas metaheurísticas: Algoritmos Transgenéticos (ATs), Goldberg e Goldberg (2002), e Path-relinking, Glover (1996). Os ATs pertencem à classe dos algoritmos evolucionários e são inspirados na Teoria da Endossimbiose e nas propriedades do fluxo intracelular. A Teoria da Endossimbiose baseia sua formulação na união de indivíduos de naturezas distintas para a constituição de saltos adaptativos ou a formação de espécies híbridas ou novas, Morowitz (1992). Os ATs, adaptando estes conceitos ao contexto computacional, propõem a evolução através de um processo de troca de informações realizado entre populações de espécies diferentes.

De modo a verificar a potencialidade do algoritmo proposto para o problema, o mesmo é comparado com duas abordagens propostas na literatura para o PSP.

Este trabalho apresenta a modelagem computacional do problema no item 2. O item 3 introduz os conceitos da metáfora empregada, além de formalizar o algoritmo proposto no item 3.1. O item 4 relata um experimento computacional e analisa seus resultados. Finalmente, o item 6 apresenta as considerações finais do trabalho.

2.MODELAGEM COMPUTACIONAL

Um modelo intuitivo para formalizar o problema da configuração do traçado de redes ramificadas é o problema da árvore de Steiner, Hwang *et al.* (1992). Considere $G = (V, E)$ um grafo onde $V = \{v_1, \dots, v_n\}$ representa os possíveis pontos de passagem da rede de distribuição, pontos provavelmente associados a uma planta urbana e $E = \{e_1, \dots, e_m\}$ que representa as vias de ligação entre os pontos de V . O conjunto de pontos é composto por três tipos de nós: nós de demanda, nó fonte e nós de passagem. Às arestas do conjunto E estão associados custos $c(e_i)$, $i = 1, \dots, m$, que quantificam a distância entre os nós. O objetivo é encontrar uma configuração em forma de árvore, com custo mínimo, que conecte os nós de demanda ao nó fonte. O problema da árvore de Steiner com coleta de prêmios é uma generalização do problema da árvore de Steiner. Esse problema considera prêmios $p(v_j)$ $j = 1, \dots, n$, associados aos nós de demanda do grafo, além dos custos associados às arestas. O objetivo é otimizar o traçado de uma rede ramificada, em forma de árvore, de modo a minimizar os custos do traçado enquanto maximiza o lucro obtido em atender à demanda da rede. Um exemplo pode ser visto na Figura 1, que apresenta as ruas do bairro de Candelária na cidade de Natal e os clientes em potencial para fornecimento de gás natural. O bairro está dividido em duas etapas para implantação da rede, sendo Fase 1 e Fase 2. A empresa distribuidora de gás natural da cidade, a Potigás, forneceu os dados da Fase 2 deste projeto. Cada ponto em preto no mapa representa um possível cliente com uma demanda de gás estimada. Estão implícitos no mapa os pontos de passagem, que possuem demanda zero. A Figura 2(a) mostra o modelo

computacional para a Fase 2, e a Figura 2(b) exibe uma possível solução para a configuração da rede, onde os pontos em preto representam os clientes, e os pontos em branco representam os pontos de passagem.

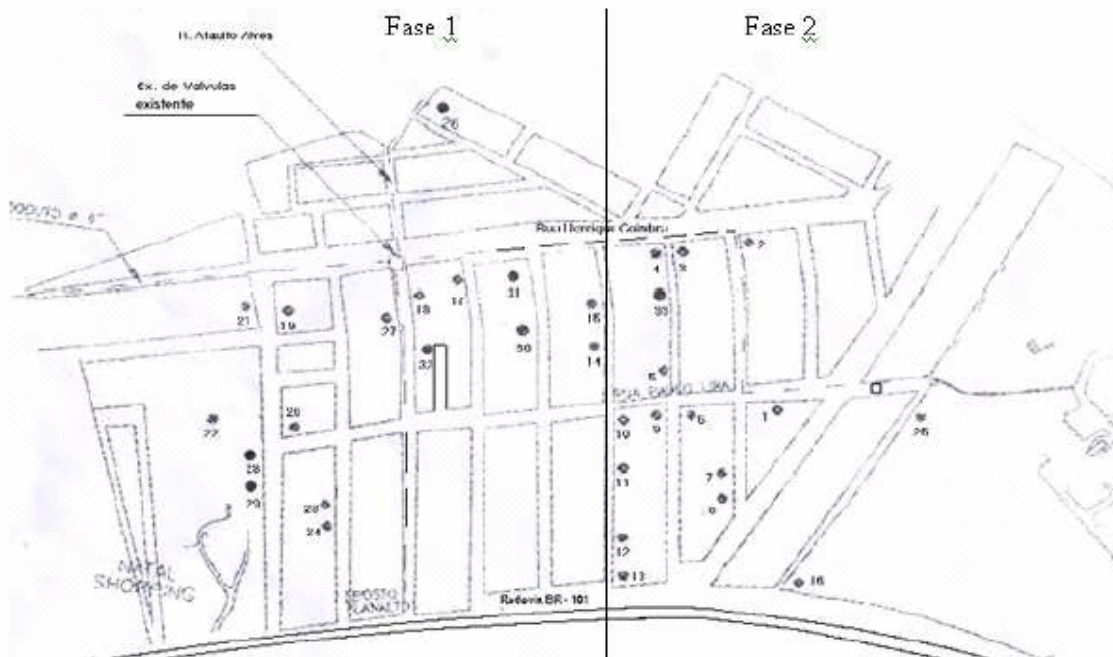


Figura 1: Clientes potenciais da rede de distribuição de gás natural para o bairro Candelária, Natal.

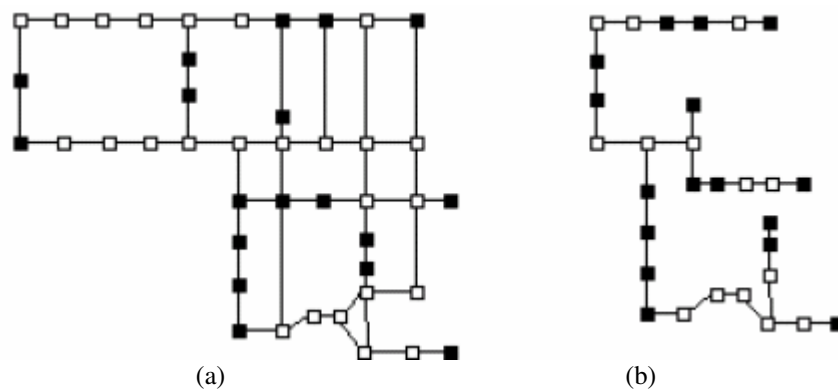


Figura 2: (a) Representação computacional da Fase 2, (b) uma configuração para a rede.

3. ALGORITMO PROPOSTO

Os ATs pertencem à classe dos algoritmos evolucionários e baseiam sua metáfora em um processo de endossimbiose, considerando ainda outros mecanismos do fluxo intracelular e mecanismos de comunicação entre bactérias. A teoria da endossimbiose foi popularizada por Margulis (1991), e diz que um novo organismo pode surgir pela fusão de dois ou mais organismos independentes. Mais precisamente, a teoria afirma que organismos que evoluíram de forma independente são capazes de se unir em um sistema simbiótico e, eventualmente, constituir um só organismo. Diversas descobertas de mecanismos evolucionários naturais voltados para o compartilhamento direto de DNA entre microorganismos ou células vêm a reforçar a teoria de Margulis. Através do compartilhamento de DNA, estes mecanismos facilitam a emergência de saltos de adequação. Um conhecido conjunto desses processos está reunido sob o nome de "transferência horizontal". Dentre os mecanismos de transferência horizontal estão os plasmídios. Os plasmídios são partículas genéticas móveis que podem ser intercambiadas entre certas células. A transferência horizontal de genes funcionais entre organismos é o fundamento teórico da origem endossimbiótica das organelas celulares, Pierce *et al.* (2003).

Além da endossimbiose, outras vias podem resultar em alteração genética guiada dentro dos microorganismos ou células como, por exemplo, os mecanismos de transposição ou transposons, Shapiro (1999). Os transposons são seqüências de DNA que fazem parte de outros elementos genéticos como cromossomos ou plasmídios. Eles se movem de uma locação para outra no DNA de uma mesma célula caracterizando uma transposição ou permutação. As permutações podem ocorrer exclusivamente sobre a informação contida no cromossomo ou em

composição com informações oriundas do meio ambiente. Finalmente, entende-se hoje que bactérias mostram habilidades de comunicação, cooperação e coordenação de ações. A habilidade das bactérias se comunicarem e coordenar o desempenho através de sinais moleculares recebeu o nome de *Quorum Sensing*, Waters e Bassler (2005).

Tendo por base, os conceitos biológicos previamente descritos, o processo de evolução de um AT se dá mediante a interação de uma população de cromossomos e um grupo de atuadores chamados de vetores transgenéticos, consideradas as possibilidades de interferência de um dado meio ambiente.

Um vetor transgenético, λ , é uma dupla $\lambda=(I,\Phi)$, onde I é uma cadeia de informação e Φ é um método de manipulação. $\Phi = (p_1, \dots, p_s)$, $p_j, j=1, \dots, s$, são procedimentos que definem a atuação do vetor. Ao infiltrarem sua cadeia de informação, I , os vetores transgenéticos provavelmente alterarão um cromossomo S e sua adequação. São previstos vários tipos de vetores transgenéticos, porém apenas dois são empregados neste algoritmo: plasmídios e transposons. A Tabela 1 resume os procedimentos que compõem o método de manipulação desses vetores. Um vetor, λ , é dito um plasmídio quando sua cadeia de informação, I , é traduzida no formato genético - uma subcadeia de DNA - e seu método utiliza somente os procedimentos p_1 e p_2 . Um vetor, λ , é dito transposon quando seu método utiliza os procedimentos p_1 , p_2 , p_3 e p_4 .

Tabela 1: Procedimentos dos vetores transgenéticos.

p_1 - Ataque (A)	Define se o cromossomo S é suscetível à manipulação do vetor λ $A: S, \lambda \rightarrow \{\text{falso}, \text{verdadeiro}\}$
p_2 - Transcrição	Se $A(S, \lambda) = \text{verdadeiro}$, I será transferida para o cromossomo S .
p_3 - Identificação	Identifica o intervalo para operação do vetor.
p_4 - Recombinação	Identifica a origem e o comprimento da informação do vetor de manipulação.

No processo evolucionário transgenético três contextos interagem: uma população de cromossomos, uma população de vetores transgenéticos e o contexto extracelular rico em informações ambientais capazes de influenciar vetores e cromossomos.

O banco de informações genéticas é o repositório geral de informações. Ele pode somar informações obtidas a priori com as que são alcançadas durante a execução do processo evolucionário. O controle da evolução das populações de vetores, cromossomos e do meio ambiente é realizado através de três classes de regras denominadas "regras transgenéticas". Existem regras dos tipos 1, 2 e 3. As regras do tipo 1 dirigem a construção da cadeia de informação I que será transportada pelos vetores transgenéticos. Elas podem utilizar qualquer conhecimento armazenado no banco de informações. As regras do tipo 2 definem como a informação I será transcrita no cromossomo - o operador que será utilizado por λ . As regras de transcrição podem evoluir em conformidade com a resistência à manipulação demonstrada pelos cromossomos. Em um algoritmo poderão existir várias regras do tipo 1 e 2. As regras tipo 3 orquestram o processo completo. Elas definem que tipos de vetores serão empregados, quantos cromossomos serão atacados em uma iteração, quantos vetores de manipulação serão criados, a condição de parada, etc. A Figura 3 apresenta o pseudo-código de um algoritmo transgenético genérico. Inicialmente é criada uma população de cromossomos. Este passo pode ser feito da mesma forma que em algoritmos evolucionários clássicos: geração aleatória, geração através de uma heurística de baixo custo computacional, etc. As regras transgenéticas serão definidas conforme a aplicação, e irão administrar a manipulação dos cromossomos. O banco de informações genéticas é obtido através de informações teóricas sobre a aplicação, e pode ser atualizado durante a evolução. O processo de troca de informações entre as populações de vetores e de cromossomos é realizado até que algum critério de parada seja atingido, sendo que as regras transgenéticas e o banco de informações podem ser atualizados durante a evolução.

```

1 Criar uma população inicial de cromossomos;
2 Carregar as regras transgenéticas;
3 Carregar o Banco de Informações Genéticas;
4 repita
5   Gerar vetores de manipulação;
6   Selecionar cromossomos para manipulação;
7   Manipular os cromossomos conforme as RET;
8   Atualizar as RET e o BIG;
9 até atender algum critério de parada.
```

Figura 3: Meta-algoritmo transgenético.

3.1 AT para o PSP

As soluções do PSP são representadas através de cromossomos. Um cromossomo é um vetor, onde cada posição representa um nó do grafo. Os nós contidos em um cromossomo, ligados através de uma árvore geradora

mínima, definirão uma solução. É utilizado o algoritmo de Prim (1957) para a definição da árvore geradora mínima. A adequação do cromossomo é calculada somando-se os custos das arestas de ligação entre os pontos da árvore com o valor dos prêmios de cada ponto fora da árvore. A Figura 4 mostra um exemplo numérico.

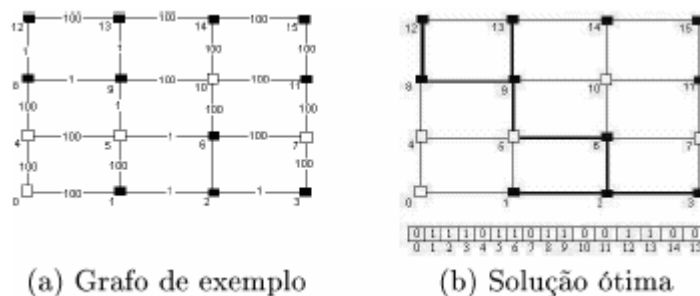


Figura 4: Exemplo numérico.

Na Figura 4(a), os nós em preto representam clientes, e possuem um valor prêmio associado. Suponha que todos os clientes têm prêmio igual a 10. Os nós em branco representam intersecções de ruas, e possuem prêmio igual a zero. O valor sobre cada aresta representa seu custo. A Figura 4(b) representa a solução ótima para a entrada da Figura 4(a), conectando à rede apenas um sub-conjunto dos clientes. A soma dos custos das arestas é igual a 8, e a soma dos prêmios dos nós não conectados é 30. Portanto o valor da função objetivo para este grafo é 38.

A presente proposta utiliza simultaneamente dois vetores transgenéticos: plasmídios e transposons. A frequência com que cada vetor é utilizado procura mimetizar o processo biológico, empregando mais intensamente a informação obtida no contexto extracelular no início da evolução, e a do contexto extracelular mais intensamente no final da evolução.

As informações a serem infiltradas pelo vetor plasmídio têm origem de duas fontes de informação, dividindo a atuação do vetor em duas etapas. No início da evolução a preferência é dada à inserção de informações do banco de informações genéticas. Dois nós são sorteados formando um intervalo no cromossomo, e os nós de demanda dentro desse intervalo que ainda não estiverem na solução são incluídos através de plasmídios. Os fragmentos genéticos transportados pelos plasmídios são os caminhos mais curtos ligando cada nó i a ser inserido até um nó w presente solução. Estes caminhos estão armazenados no banco de informações, BIG. No final da evolução, a preferência é dada para a inserção de informações obtidas de trechos de uma solução de qualidade, contida no *pool* de soluções elite. É sorteada do *pool* uma solução qualquer, e desta solução é sorteado um trecho de no máximo $\lfloor n/4 \rfloor$ posições. Os nós deste trecho são inseridos, e então é construída a AGM do grafo induzido pelo conjunto de nós.

A probabilidade com que cada tipo de informação é infiltrada é definida através dos parâmetros u e β . Um valor aleatório de zero até o número máximo de gerações é atribuído a u . O valor de u tende a ser maior do que β no início da evolução, pois β representa o número da geração atual. Conforme o número de gerações evolui, maior é a probabilidade de u ser menor do que β . A variável x recebe aleatoriamente um valor de 1 a 100. Se o valor sorteado for menor que 80, serão infiltradas informações do contexto extracelular. Caso contrário serão infiltradas informações do contexto intracelular. Essas probabilidades se invertem conforme o valor de u . O pseudo-código do procedimento do vetor plasmídio é mostrado na Figura 5.

O vetor transposon realiza re-arranjos do código genético em determinados trechos do cromossomo, com o objetivo de encontrar nestes re-arranjos uma solução de qualidade. O fragmento a ser modificado é obtido sorteando-se duas posições quaisquer formando um intervalo no cromossomo. A proposta do transposon verifica uma vizinhança que é obtida do exame de possíveis retiradas de nós existentes no trecho sorteado e a construção da árvore geradora mínima do conjunto de nós da solução corrente, em busca de uma economia de traçado. É retornada a melhor solução gerada. O pseudo-código é mostrado na Figura 6.

Mimetizando o processo biológico, o algoritmo transgenético emprega mais intensamente a informação obtida no contexto extracelular no início da evolução. Na medida em que a população de cromossomos evolui, a informação de origem endógena passa a ser privilegiada. Conseqüentemente, a arquitetura geral do algoritmo da Figura 3 pode ser adaptada na solução do problema como sugerido no pseudo-código da Figura 7.

No passo 1 é gerada uma população inicial de cromossomos através de uma implementação do algoritmo primal dual de Johnson *et al.* (2000) com perturbações, Canuto *et al.* (2001). Este algoritmo determina um conjunto de nós que devem compor uma solução. Em seguida é construída a árvore geradora mínima do grafo induzido por este conjunto de nós, formando uma solução viável.

No passo 2 é organizado o banco de informações que será utilizado pelos plasmídios. O banco contém, a princípio, informações obtidas a priori. Posteriormente serão incorporadas informações alcançadas no processo evolucionário. As informações obtidas a priori formam uma matriz que contém os caminhos mais curtos entre

todos os nós do grafo de entrada. As informações do processo evolucionário são representadas por trechos dos cinco melhores cromossomos da população. No passo 3 é armazenada a melhor solução gerada na população inicial. Enquanto o critério de parada não for atendido o algoritmo realiza as etapas de interação entre as populações de vetores e a população de cromossomos. A cada iteração, um tipo de vetor, plasmídio ou transposon, é selecionado. Como a metáfora sugere que a infiltração de informações extracelulares tende a ser mais útil nas etapas iniciais de um processo evolucionário, a probabilidade de escolha do tipo de vetor varia ao longo das iterações. Na fase inicial do processo o fluxo de informações extracelular é privilegiado, de modo que o ataque via plasmídios possui maior probabilidade de ocorrência. Ao final do processo, a probabilidade de atuação é maior para os transposons.

```

Ataque_Plasmidio( $C, \beta$ )
1  $u \leftarrow \text{rand}() \% \text{ngerações};$ 
2  $x \leftarrow \text{rand}() \% 100;$ 
3 se  $u > \beta$  então
4   se  $x < 80$  então
5      $v_1 \leftarrow \text{rand}() \% n; v_2 \leftarrow \text{rand}() \% n;$ 
6     para  $i \leftarrow v_1$  até  $v_2$  faça
7       se  $p_i > 0$  e  $i \notin C$  então
8          $C \leftarrow C + \text{BIG}[i][w]$ 
9   senão
10     $x \leftarrow \text{rand}() \% \text{tampool};$ 
11     $C \leftarrow C + \text{trecho}(\text{pool}[x]);$ 
12     $C \leftarrow \text{AGM}(C);$ 
13 senão
14   se  $x < 20$  então
15      $v_1 \leftarrow \text{rand}() \% n; v_2 \leftarrow \text{rand}() \% n;$ 
16     para  $i \leftarrow v_1$  até  $v_2$  faça
17       se  $p_i > 0$  e  $i \notin C$  então
18          $C \leftarrow C + \text{BIG}[i][w]$ 
19   senão
20     $x \leftarrow \text{rand}() \% \text{tampool};$ 
21     $C \leftarrow C + \text{trecho}(\text{pool}[x]);$ 
22     $C \leftarrow \text{AGM}(C);$ 

```

Figura 5: Algoritmo de ataque através do vetor plasmídio.

```

Ataque_Transposon( $C$ )
1  $v_1 \leftarrow \text{rand}() \% n; v_2 \leftarrow \text{rand}() \% n;$ 
2 para  $i \leftarrow v_1$  até  $v_2$  faça
3    $C' \leftarrow C - i;$ 
4    $C' \leftarrow \text{AGM}(C');$ 
5 se  $C'$  melhor que  $C$  então
6    $C \leftarrow C';$ 

```

Figura 6: Algoritmo de ataque através do vetor transposon.

Após o ataque com os vetores, a solução resultante é submetida a um procedimento de path-relinking, Glover (1996), para intensificação, passo 12. Se o cromossomo resultante deste processo já existir na população, é realizado um processo de perturbação, para manter a diversidade entre os cromossomos. Caso o novo indivíduo represente uma solução melhor que todas as encontradas até o momento, a melhor solução da população e o *pool* de soluções elite são atualizados, passo 16 e 17. Se o indivíduo gerado é melhor que o original, então ele substitui o antigo na população, passo 19. Caso contrário, o indivíduo só substituirá o antigo com 5% de probabilidade, passo 23. Essa possibilidade está presente pois uma piora momentânea pode acarretar em uma melhoria futura.

No presente trabalho, a idéia de Laguna e Marti (1999) e Canuto *et. al.* (2001) foi adaptada para o uso conjunto do path-relinking com o AT. A Figura 8 resume a seqüência de passos. Um conjunto de soluções elite é formado por um *pool* com as melhores soluções encontradas até então. Para fazer a conexão de caminhos, é selecionada do *pool* a solução X mais próxima do cromossomo C que está sendo manipulado, através de um critério de distância (passo 1). A distância é calculada através da subtração do número de nós que as duas soluções têm em comum do número de nós da união das duas soluções. O conjunto de movimentos possíveis de

C' até X é armazenado no vetor M (passos 2 a 8). São realizados então todos os movimentos possíveis (passos 9 a 14). Os movimentos consistem na inserção e exclusão de nós da solução C' , de modo que ela se torne mais próxima da solução X . A viabilidade da solução é garantida pois é construída a AGM do grafo induzido pelo conjunto de nós em C' (passo 15). A melhor solução gerada nesta série de movimentos é armazenada em C (passo 17) e é retornada.

```

1 Gerar uma população inicial de cromossomos –  $P$ ;
2 Carregar o banco de informações genéticas –  $BIG$ ;
3  $C^* \leftarrow$  melhor cromossomo de  $P$ ;
4  $\beta \leftarrow 1$ ;
5 enquanto  $\beta < n_{gerações}$  faça
6   para  $j \leftarrow 1$  até  $tam_{pop}$  faça
7      $u \leftarrow \text{rand}()\%n_{gerações}$ ;
8     se  $u > \beta$  então
9        $C' \leftarrow \text{Ataque\_Plasmidio}(C_j, \beta)$ ;
10      senão
11         $C' \leftarrow \text{Ataque\_Transposon}(C_j)$ ;
12         $C' \leftarrow \text{PathRelinking}(C')$ ;
13        se  $C' \in P$  então
14           $C' \leftarrow \text{Perturba}(C')$ ;
15        se  $C'$  melhor que  $C^*$  então
16           $C^* \leftarrow C'$ ;
17          AtualizaPoolElite( $C^*$ );
18        se  $C'$  melhor que  $C_j$  então
19           $C_j \leftarrow C'$ ;
20      senão
21         $x \leftarrow \text{rand}()\%100$ ;
22        se  $x \leq 5$  então
23           $C_j \leftarrow C'$ ;
24   $\beta++$ ;

```

Figura 7: Algoritmo transgenético híbrido.

```

PathRelinking( $C$ )
1 Encontrar solução elite  $X$  mais próxima de  $C$ ;
2 para  $i \leftarrow 1$  até  $n$  faça
3   se  $C_i \neq X_i$  então
4     se  $C_i = 1$  então
5        $M_i \leftarrow 1$ ;
6     senão
7        $M_i \leftarrow 2$ ;
8   senão  $M_i \leftarrow 0$ ;
9 para  $i \leftarrow 1$  até  $n$  faça
10  se  $M_i \neq 0$  então
11    se  $M_i = 1$  então
12       $C'_i \leftarrow 0$ ;
13    senão
14       $C'_i \leftarrow 1$ ;
15     $C' \leftarrow \text{AGM}(C')$ ;
16  se  $C'$  melhor que  $C$  então
17     $C \leftarrow C'$ ;

```

Figura 8: Algoritmo de Path-Relinking.

4. RESULTADOS COMPUTACIONAIS

Os testes computacionais foram executados em uma plataforma Pentium IV 2.8GHz, com 256 Mb de RAM, sistema operacional Kubuntu Linux 6.06 e compilador gcc. Foram utilizadas 114 instâncias teóricas de um banco disponível eletronicamente em <http://www.research.att.com/flmger/data>. O valor da solução ótima é disponibilizado, sendo que o mesmo foi obtido pelos métodos de Lucena e Resende (2000) e de Klau *et al.*

(2004). A escolha dos parâmetros do algoritmo foi definida através de testes preliminares. O número máximo de gerações é igual a 60, o tamanho da população é 25. Como a cada geração são manipulados os 25 indivíduos da população, tem-se um total de 1500 iterações ($60 \times 25 = 1500$). É utilizada uma condição de parada adicional, com 1200 iterações sem melhora o algoritmo pára retornando a melhor solução obtida até então. O parâmetro β recebe 1 no início do algoritmo e é incrementado de um a cada geração. O *pool* de soluções elite tem tamanho igual a 5.

Tabela 2: Resultados obtidos para a série K.

Instância	BLP		AM		AT	
	%gap	tempo	%gap	tempo	%gap	tempo
K100.8	0	2	2,3	1,5	0	0
K400.1	0	194	0	6,7	0	0
K400.2	0,2	234	0,2	7,1	0	1,5
K400.3	0	140	0	7,4	0	0
K400.4	0	204	0,1	7,7	0	0,8
K400.5	0	122	0,3	7,1	0	0,2
K400.6	0	60	0	8,3	0	0,1
K400.7	0,1	306	0,1	7,7	0	0,9
K400.8	0	42	0	6,7	0	0
K400.9	0	76	0,1	7,5	0	1
K400.10	0,4	231	0,7	8,7	0,2	0,4

Tabela 3: Resultados obtidos para a série P.

Instância	BLP		AM		AT	
	%gap	tempo	%gap	tempo	%gap	tempo
P400	0	397	0,2	29,3	0,2	1,4
P400.4	0	565	0,5	21,4	0,2	1,1

Tabela 4: Resultados obtidos para a série C.

Instância	BLP		AM		AT	
	%gap	tempo	%gap	tempo	%gap	tempo
C04-B	0	387	0,36	39	0,09	10,9
C07-B	0,98	167	3,92	5,7	0,98	0,9
C08-B	0	404	0,4	29	0	8
C09-B	0	583	0,76	38,5	0	24,3
C13-A	0,42	1050	0,42	17,9	0,84	6,8
C16-A	0	1920	9,09	10,6	0	0,7
C16-B	0	1758	9,09	11,5	0	0,9
C17-A	0	549	5,55	11,2	0	0,2
C19-B	0	3390	2,05	15,8	0	12,3

É realizada uma comparação do desempenho do algoritmo com o de outros dois algoritmos baseados em metaheurísticas, para concluir sobre a competitividade da proposta, e avaliar sua eficiência em relação ao estado da arte. São eles o algoritmo de busca local com perturbações (BLP) de Canuto *et al.* (2001) e o algoritmo memético (AM) de Klau *et al.* (2004). O tempo e o desempenho dos algoritmos comparados foram transcritos dos trabalhos publicados, e da tese de Ljubic (2004). Foram realizadas 50 execuções independentes do algoritmo transgnético para cada instância. As Tabelas 2, 3, 4 e 5 reportam os resultados para as instâncias onde resultados distintos quanto à qualidade da solução foram encontrados pelos algoritmos. Na primeira coluna é mostrado o nome da instância. As três colunas seguintes identificam os algoritmos a serem comparados. Estas colunas são subdivididas em duas colunas. A coluna "%gap" descreve o afastamento percentual entre o resultado obtido e o valor ótimo. A coluna tempo relata o tempo médio em segundos de processamento nas instâncias examinadas.

Pode-se notar através da comparação mostrada nas tabelas que o AT apresenta o melhor desempenho para as instâncias da série K, tanto em qualidade de soluções quanto em tempo computacional. Para as séries P e C, o BLP apresenta melhor qualidade de soluções, porém o tempo computacional do AT é o melhor. Para a série D, o AT apresenta melhor qualidade de soluções, porém o melhor tempo computacional é o do AM.

5. CONCLUSÕES

No presente trabalho, o problema da definição do traçado de redes de distribuição de gás é solucionado através de uma abordagem evolucionária intitulada Algoritmo Transgenético. Anteriormente, apenas duas abordagens metaheurísticas foram propostas para a modelagem utilizada para o problema, sendo uma delas evolucionária. Pode-se concluir, através das tabelas apresentadas, que o AT proposto obteve resultados promissores em relação às abordagens comparadas, apesar de não utilizar nenhum pacote de software pronto e não utilizar procedimentos de redução dos grafos de entrada, como fazem os outros dois algoritmos.

O tempo computacional do AT é, em média, menor que os tempos reportados no experimento BLP, e maior que os tempos reportados no AM apenas quando nos referimos às instâncias da série D. Todavia ressalte-se que o AT foi executado em plataforma superior a do primeiro trabalho, e inferior a do segundo trabalho. Adicionalmente, o AM realiza uma considerável redução das instâncias, e pode-se concluir que diminuindo o tamanho da entrada diminui-se o tempo necessário para o processamento. Estes fatores impedem uma comparação conclusiva com respeito ao tempo.

Com os experimentos realizados, concluiu-se que a mimetização do comportamento dos vetores do fluxo intra e extracelular pode inspirar algoritmos computacionais eficientes quantitativa e qualitativamente.

Tabela 5: Resultados obtidos para a série D.

Instância	BLP		AM		AT	
	%gap	tempo	%gap	tempo	%gap	tempo
D02-B	4,58	486	0	7,3	0	7,1
D03-B	0,06	2184	0,47	51	0,66	255,3
D04-A	0	1263	0	10,4	0,41	99,1
D04-B	0	2233	0,22	49,6	0,21	537,1
D05-A	0	3352	0	29,1	0,27	669,7
D05-B	0	2555	0,08	65,1	0,09	1200,2
D06-B	4,47	702	8,35	10,5	1,49	8,7
D07-B	1,94	711	1,94	9,5	0	0
D08-A	0	1727	0	19,1	0,39	160,7
D08-B	0,19	3175	0,93	123,8	0,19	346,1
D09-A	0,18	4109	0,43	52,1	0,56	365,5
D09-B	0	2754	1,15	151,2	0,07	224,8
D10-A	0	4193	0,2	122,2	0,11	878
D10-B	0	2644	0,51	107,3	0,04	1108,6
D11-B	3,44	1280	0	17,4	3,44	2,2
D13-A	0	5047	0,38	58,7	0	195,4
D13-B	0	4288	1,17	97,2	0,2	359,3
D14-A	0	6388	0,59	102,3	0	302,9
D14-B	0	6178	1,38	102,8	0	452,9
D15-A	0	7840	0,64	145,7	0	274,4
D15-B	0	5220	0,6	95,6	0	389,6
D16-A	0	1397	7,69	23,1	0	0
D18-A	0	30044	1,28	81,4	0,45	215,5
D18-B	0,44	36643	3,22	98,7	1,34	312,5
D19-A	0,65	40955	3,82	87,6	1,3	525,2
D19-B	0,32	38600	2,51	81,9	0,64	700,5
D20-A	0	28139	0,18	18,4	0	111,9

6. AGRADECIMENTOS

Ao Maurício Resende, por ter disponibilizado detalhes dos resultados computacionais de seu artigo. A presente pesquisa foi apoiada pelo programa ANP PRH-22 e pelo CNPq.

7. REFERÊNCIAS

- Borelli, A.; Naka, H.; Oliveira, V. Gás Natural – Transporte e Distribuição. **Relatório técnico - Instituto de Eletrotécnica e Energia (PIPGE-IEE/USP)**, Universidade Federal de São Paulo, 2001.
- Canuto, S.; Resende, M.; Ribeiro, C. Local search with perturbations for the prize collecting Steiner tree problem in graphs. **Networks**, v.38, p.50-58, 2001.
- Glover, F. Tabu search and adaptive memory programming - advances, applications and challenges. **Interfaces in Computer Science and Operations Research**, p.1-75, 1996.

- Goldberg, M. C; Goldberg, .E. F. G. Transgenética computacional: Uma aplicação ao problema quadrático de alocação. **Pesquisa Operacional**, v.22, p.359-386, 2002.
- Hwang, F.K.; Richards, D.S; Winter, P. The Steiner Tree Problem. **Annals of Discrete Mathematics**, v.53, 1992.
- Johnson, D. S.; Minkoff, M.; Phillips, S. The prize collecting steiner tree problem: Theory and practice. **Proceedings of 11th ACM-SIAM Symposium on Discrete Algorithms**, p.760-769, San Francisco, CA, 2000.
- Karp, R. M. Reducibility among combinatorial problems. **Complexity of Computers Computations**, p.85-103, 1972.
- Klau, G. W.; Ljubic, I.; Moser, A.; Mutzel, P.; Neuner, P.; Pferschy, U.; Raidl, G.; Weiskircher, R. Combining a memetic algorithm with integer programming to solve the prize collecting steiner tree problem. **Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2004)**, v.3102 of LNCS, p.1304-1315, Springer-Verlag, 2004.
- Laguna, M.; Martí, R. Grasp and path relinking for 2-layer straight line crossing minimization. **INFORMS Journal on Computing**, v.11, p.44-52, 1999.
- Ljubic, I. Exact and Memetic Algorithms for Two Network Design Problems. PhD thesis, Technische Universität Wien, Fakultät für Informatik, 2004.
- Lucena, A.; Resende, M. G. C. Strong lower bounds for the prize collecting steiner tree problem in graphs. **AT&T Technical Report**, disponível em <http://www.research.att.com/flmgcr>, 2000.
- Margulis, L. Symbiosis as a source of evolutionary innovation: Speciation and morphogenesis. **The MIT Press**, 1991.
- Morowitz, H. J. Beginning of Cellular Life. New Haven, Conn, **Yale University Press**, 1992.
- Pierce, S. K., Massey, S. E., Hanten, J. J., and Curtis, N. E. Horizontal transfer of functional nuclear genes between multicellular organisms. **Biological Bulletin**, p.237-240, 2003.
- Prim, R. C. Shortest connection networks and some generalizations. **Bell System Tech.**, v.36, p.1389-1401, 1957.
- Shapiro, J. A. Transposable elements as the key to a 21st century view of evolution. **Genetica**, v.107, p.171-179, 1999.
- Waters, C.; Bassler, B. Quorum sensing: Cell-to-cell communication in bacteria. **Annual Review of Cell and Developmental Biology**, v.21, p.319-346, 2005.

AN APPROACH BY EVOLUTIONARY ALGORITHM FOR THE NATURAL GAS NETWORK CONFIGURATION PROBLEM.

The design network of a natural gas distribution problem requires a layout definition by where ducts can be laid, so that the company can attend a set of customers. This layout definition is of extreme importance in the construction of these networks, therefore, if the layout be planned in efficient way, the company can attend the demand having a low cost. This problem can be represented in a computer in diverse ways. This work presents a modeling where the focus is the attendance of customers whose demand for natural gas compensates the construction of an infrastructure to attend them. Through this approach, this work provides to the company a differentiated economic vision. While the traditional approaches wants to attend all the customers in a network, this approach want to attend only the customers who are financially attractive. The used modeling is based on the prize collecting Steiner tree problem, that is a generalization of the Steiner tree problem. This is a NP-hard problem. The accurate solution of these problems is impracticable for real cases, where the number of customers can be very large. Only two metaheuristics approaches had been found in literature to resolve this problem. This work presents a metaheuristic approach based in Transgenetic Algorithms (TAs), in the category of the evolutionary algorithms. The ATs metaphor considers an endosymbiosis process of genetic information exchange, according to the endosymbiosis theory. This theory was popularized by Lynn Margulis, and says that a new organism can appear by the fusing of two or more independent beings. Moreover, other transformations of the biological intracellular flow are considered in the TAs. In the algorithm, a population of chromosomes interacts with a population of transgenetics vectors, according to a set of rules. For the generation of the initial population a primal-dual algorithm was used. Path-relinking was used as intensifier in the search. Two approaches based on metaheuristics found in literature had been used to comparing the proposal algorithm in the computational experiments. In the comparison 114 instances had been used, and the results show that the proposal approach is efficient to resolve the problem.

Network design, natural gas, evolutionary algorithm, prize collecting Steiner tree.

Os autores são os únicos responsáveis pelo conteúdo deste artigo.