

# UMA PROPOSTA PARA OTIMIZAÇÃO DE SISTEMAS DE TRANSPORTE COLETIVO POR ALGORITMOS GENÉTICOS

**Glauber O. Santana**<sup>1</sup> (glauber.santana@gmail.com)

**Daniele S. Silva**<sup>1</sup> (danielless.sp@gmail.com)

**Fernando K. Correia**<sup>1</sup> (jpkoyama@gmail.com)

**Álvaro Luiz. Fazenda**<sup>1</sup>

**Luis Fernando Almeida**<sup>1</sup> (luis.almeida@unitau.br)

<sup>1</sup>UNITAU – Universidade de Taubaté - Departamento de Informática

***Resumo.** O transporte coletivo geralmente é formado por unidades de transporte que seguem itinerários pré-determinados. Nestes sistemas de transporte dois problemas se destacam: unidades com poucos passageiros e insatisfação dos usuários devido à demora no atendimento. Este artigo propõe uma abordagem baseada em Algoritmos Genéticos com o objetivo de otimizar a utilização destas unidades, tornando os itinerários dinâmicos de tal modo que elas transitem apenas por onde são solicitadas. Os resultados obtidos em simulação apontam a viabilidade de um sistema inteligente capaz de auxiliar no melhor aproveitamento das unidades de transporte.*

## 1. INTRODUÇÃO

Um problema comum nas grandes metrópoles atualmente é o deslocamento diário da população de casa para o trabalho e do trabalho para casa, já que geralmente as áreas residenciais encontram-se distantes dos centros comerciais e dos pólos industriais. Muitas iniciativas têm sido adotadas pelos governantes para enfrentar essa dificuldade, como a construção de linhas metroviárias e de vias especiais para tráfego de ônibus, gerando uma possível melhora desses meios de transporte. Entretanto, é comum a ocorrência de reclamações quanto à demora no atendimento aos passageiros, à superlotação e à inexistência de linhas que atendam às populações em bairros mais distantes dos grandes centros. Por outro lado, o aumento das frotas de ônibus e a construção de infra-estruturas de transportes coletivos demandam uma grande quantidade de investimentos e nem sempre garantem um retorno financeiro ou social recompensador. A busca de soluções viáveis para minimizar problemas decorrentes do transporte coletivo tem sido explorada por alguns pesquisadores através de propostas de sistemas baseados em consultas telefônicas (McCormack, 1999) e de sistemas de informações eletrônicos em tempo real (Caetano, 2005). Este artigo propõe um diferente paradigma para os transportes coletivos de vias públicas, no qual as unidades que efetuam o transporte dos passageiros possam deixar de seguir itinerários pré-fixados e passem a percorrer apenas as vias necessárias para atender os usuários que aguardam por um ônibus em diferentes instantes. Para tanto, uma estrutura capaz de registrar e de transmitir a existência de usuários e de pontos de espera seria necessária, além de um meio de determinar, de forma dinâmica e rápida, o melhor caminho a ser tomado pelo veículo para buscar e levar os passageiros em seus destinos. Aqui, parte-se do princípio da existência desta estrutura, e é proposto um método de otimização de rotas para os pontos que devem ser visitados baseado em Algoritmos Genéticos (AG). Nas demais seções serão apresentadas o que segue: na seção 2 serão apresentados os conceitos básicos sobre AG; a seção 3 descreve o problema e a formulação proposta nesta pesquisa; a seção 4 ilustra os resultados do método proposto; e a seção 5 apresenta as considerações finais sobre a pesquisa realizada.

## 2. ALGORITMOS GENÉTICOS

A Computação Evolutiva (CE) é uma área da Ciência da Computação que engloba um amplo conjunto de técnicas de resolução de problemas que se baseiam na simulação computacional de processos derivados da biologia como reprodução e evolução, à luz da teoria de Darwin (Davis, 1991), sem a necessidade do conhecimento de uma sistemática prévia de resolução.

Dentre os algoritmos que compõem a CE destacam-se a Programação Evolutiva (Fogel et al, 1966), as Estratégias Evolutivas (Rechenberg, 1973), a Programação Genética (Koza, 1992) e os Algoritmos Genéticos (Holland, 1972). Atualmente, os mais beneficiados com as técnicas de CE estão sendo os problemas de otimização, além de muitos problemas de engenharia que, segundo Michalewicz e Fogel (2000), podem ser adequados como problemas de otimização, garantindo que o escopo de abordagem se mantenha bem amplo. Goldberg define os AG como "algoritmos baseados nos mecanismos de seleção e genética natural" (Goldberg, 1989). São comumente usados em processos de otimização, ou seja, em processos cujo objetivo é alcançar um ou vários pontos ótimos buscando sempre melhorias no desempenho. Os AG empregam uma terminologia originada da teoria da evolução natural e da genética, na qual um indivíduo da população (uma solução em potencial para o problema) é representado por um único cromossomo, o qual contém a codificação de uma possível solução do problema. Um Algoritmo Genético básico consiste no seguinte processo: geração da população inicial e a execução de um ciclo subdividido na avaliação da população atual, seleção de indivíduos para o processo de reprodução e geração da nova população. A população inicial consiste no conjunto de soluções iniciais para o problema, podendo ser geradas as seguintes diferentes abordagens: randômica uniforme, randômica não-uniforme, randômica com *dope* e parcialmente enumerativa. A avaliação de cada indivíduo é realizada por meio da função de aptidão que mede o grau de aptidão de um indivíduo e fornece a probabilidade do indivíduo sobreviver para a próxima geração. A seleção dos indivíduos para o processo de reprodução pode ser conduzida segundo os critérios: roleta, torneio e determinismo. A reprodução para geração da nova população é realizada por meio de dois operadores genéticos: cruzamento e mutação. O operador de cruzamento gera novos indivíduos através da combinação de dois ou mais indivíduos. O cruzamento mais comum é o de um ponto de corte, conforme ilustrado na figura 1. Existem outros tipos de cruzamentos como o cruzamento de dois ou mais pontos de corte e o cruzamento uniforme (Syswerda, 1989).

|             |   |   |   |   |   |   |   |   |
|-------------|---|---|---|---|---|---|---|---|
| Pai         | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| Pai         | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Descendente | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 0 |
| Descendente | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 |

Figura 1. Exemplo de cruzamento de um ponto.

O operador de mutação modifica aleatoriamente um ou mais genes de um cromossomo e a probabilidade dessa modificação em um gene é denominada taxa de mutação. A idéia é criar uma variabilidade extra na população.

### 3. PROBLEMA DE TRANSPORTE COLETIVO

Analisando o sistema de transporte coletivo de uma cidade, pode-se verificar que uma das formas de agilizar o atendimento aos usuários e maximizar o aproveitamento das unidades de transporte seria fazer com que cada uma dessas unidades trafegasse estritamente por onde são necessárias, ou seja, nos pontos onde os usuários a estivessem aguardando. Para tanto, seria necessário que a quantidade de usuários aguardando em cada ponto de parada pudesse ser obtida e que cada unidade de transporte fosse capaz de determinar a melhor forma de atender a esses usuários. Esse trabalho propõe uma maneira de definir o caminho de menor custo que parte do ponto onde se encontra a unidade de transporte e passa por cada um dos pontos onde ela é solicitada, qualquer que seja a ordem. Porém, os pontos que uma unidade de transporte deve visitar variam de acordo com a entrada de novos usuários no sistema ou com o atendimento aos usuários que já estão nele, ou seja, a definição da rota da unidade de transporte deve ser dinâmica e aceitar alterações mesmo durante a sua execução. Para simplificar o modelo a ser criado, considerou-se que cada unidade de transporte atue em uma determinada região da cidade formada por um subconjunto de pontos de paradas próximos ou adjacentes, cujos usuários devem ser atendidos apenas por esta unidade. Assumindo essa idéia, há pontos de parada que pertencem à fronteira entre duas regiões e servem como ponte quando o usuário de uma região deseja ir até um ponto pertencente à outra região.

#### 3.1. Implementação do AG

Primeiramente, para representação do ambiente a ser explorado foi utilizado um grafo direcionado como estrutura de dados para representar as ruas e cruzamentos de uma cidade, conforme exemplificado pela figura 2, armazenado na forma de listas de adjacências (Drozdek, 2002), tal qual apresentado na tabela 1.



Figura 2. Exemplo de grafo com aresta bi-direcionadas.

Cada indivíduo para o problema é representado na forma de um vetor contendo uma seqüência dos índices dos vértices na ordem em que devem ser percorridos pela unidade de transporte. Essa codificação, além de ser de simples implementação, mostrou-se também a mais eficiente para se realizar as operações de cruzamento e mutação, por fornecer um objeto de comparação direta mesmo quando um elemento do vetor é analisado isoladamente.

Tabela 1: Exemplo de lista de adjacência do grafo.

| Vértice |      |       | Adjacências |       |      |       |      |       |
|---------|------|-------|-------------|-------|------|-------|------|-------|
| Índice  | Nome | Custo | Nome        | Custo | Nome | Custo | Nome | Custo |
| 0       | A    | 1     | B           | 8     | K    | 25    |      |       |
| 1       | B    | 1     | A           | 8     | C    | 10    |      |       |
| 2       | C    | 4     | B           | 10    | D    | 2     | F    | 9     |
| 3       | D    | 2     | C           | 2     | E    | 5     |      |       |
| 4       | E    | 5     | D           | 5     | F    | 5     | G    | 7     |
| 5       | F    | 5     | C           | 9     | E    | 5     | I    | 8     |
| 6       | G    | 1     | E           | 7     | H    | 10    |      |       |
| 7       | H    | 2     | G           | 10    | I    | 6     |      |       |
| 8       | I    | 4     | F           | 8     | H    | 6     | J    | 10    |
| 9       | J    | 2     | I           | 10    | K    | 5     |      |       |
| 10      | K    | 1     | A           | 25    | J    | 5     |      |       |

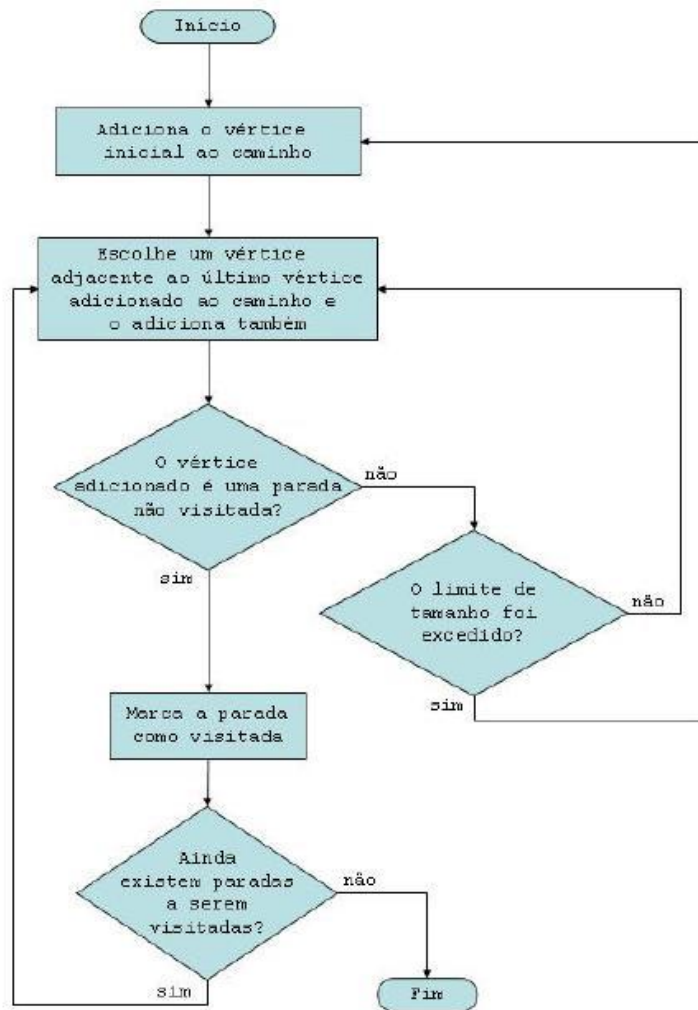


Figura 3. Procedimento para geração de um indivíduo.

A população inicial é gerada por meio de um processo aleatório com *dope*, pois, além do vértice inicial de cada indivíduo ser o vértice onde a unidade de transporte se encontra, todos os vértices que obrigatoriamente devem ser visitados pela unidade de transporte (deste ponto em diante chamados de paradas) também estarão presentes em cada um dos indivíduos. O procedimento para geração de um indivíduo é descrito na figura 3. O tempo limite utilizado corresponde à quantidade relacionada ao dobro do número de vértices existentes no grafo, pois se um caminho visita uma quantidade de vértices maior do que o dobro da quantidade de vértices no grafo, então esse caminho certamente tem ciclos e vértices repetitivos, o que o torna pouco qualificado para a busca pelo caminho de menor custo. Foi utilizado o cruzamento de ponto único com corte aleatório. No entanto, para que não ocorresse um filho com um caminho impossível, o ponto de corte é selecionado aleatoriamente apenas no primeiro pai. No segundo, é procurado o mesmo vértice do ponto de corte definido para o primeiro, para que a emenda dos pedaços não gere falhas de continuidade. Mesmo assim, há a possibilidade de gerar como filhos caminhos que não visitem todas as paradas. Esses filhos são então descartados durante a seleção de indivíduos que farão parte da próxima geração. Para a seleção dos indivíduos foi utilizada uma roleta em sua forma tradicional. A mutação foi implementada com dois pontos de corte aleatórios que definem o trecho do indivíduo a ser substituído por uma nova formação aleatória de genes.

#### 4. TESTES E RESULTADOS

O AG proposto foi aplicado a um ambiente representado pelo grafo ilustrado pela figura 4. Esse ambiente é composto por 100 vértices interligados por arestas bidirecionadas que recebem os custos de acordo com a legenda representada pela figura 5. Como se pode observar nesse grafo, os caminhos são melhores quando percorrem os vértices na direção para baixo e para a direita e piores quando percorrem na direção para cima e para a esquerda. Intencionou-se, dessa forma, facilitar a visualização dos resultados obtidos e também validar o método de solução proposto.

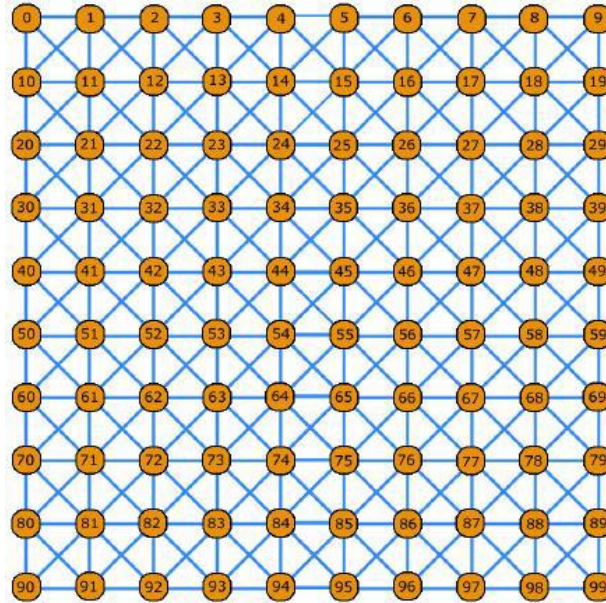


Figura 4. Grafo do ambiente simulado.

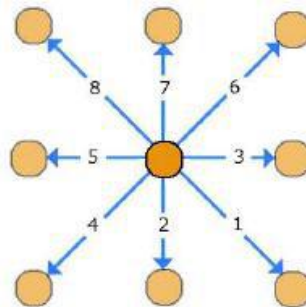


Figura 5. Regra para custo das arestas.

Tabela 2: Valores para os parâmetros genéticos.

| Parâmetro            | Valores                     |
|----------------------|-----------------------------|
| Tamanho da população | 50, 100, 150, 200, 250, 300 |
| Taxa de cruzamento   | 60, 70, 80, 90 (em %)       |
| Taxa de mutação      | 1, 5, 10, 15, 20 (em %)     |
| Taxa de elitismo     | 10, 20, 30 (em %)           |

Para os testes e análise de desempenho foi definido como vértice inicial o nó zero e como paradas obrigatórias os nós 13, 42, 56, 75 e 99. O custo mínimo necessário para se atender a essas condições é de 36, porém esse custo pode ser obtido com alguns caminhos diferentes.

Tabela 3: Média dos menores custos obtidos e do tempo de execução variando-se a quantidade de indivíduos na população.

| Tamanho da população | Custo do melhor caminho | Erro Relativo | Tempo   | Tamanho da população | Custo do melhor caminho | Erro Relativo | Tempo    |
|----------------------|-------------------------|---------------|---------|----------------------|-------------------------|---------------|----------|
| 50                   | 56,7                    | 57,5%         | 0,087 s | 350                  | 39,5                    | 9,7%          | 0,623 s  |
| 100                  | 41,7                    | 15,8%         | 0,187 s | 400                  | 39,1                    | 8,6%          | 0,775 s  |
| 150                  | 42,1                    | 16,9%         | 0,268 s | 450                  | 38,6                    | 7,2%          | 0,831 s  |
| 200                  | 40,9                    | 13,6%         | 0,338 s | 500                  | 38,1                    | 5,8%          | 0,879 s  |
| 250                  | 40,4                    | 12,2%         | 0,446 s | 550                  | 37,2                    | 3,3%          | 1,0316 s |
| 300                  | 39,2                    | 8,9%          | 0,504 s | 600                  | 36,8                    | 2,2%          | 1,179 s  |

Primeiramente, para se definir o melhor conjunto de valores para alguns parâmetros genéticos, foi estabelecido uma faixa de valores válidos para cada parâmetro e executadas simulações para todas as combinações possíveis destas. Os valores válidos fixados para os parâmetros são apresentados na tabela 2. Das 360 possíveis configurações a qual apresentou os melhores resultados em média define valor 200 para tamanho da população, 80% para taxa de cruzamento, 15% para taxa de mutação e 20% para taxa de elitismo. Partindo-se da configuração encontrada, uma segunda bateria de testes foi executada, agora somente variando-se o tamanho da população, e mantendo-se os outros valores definidos anteriormente. Estes testes visam medir a desempenho do algoritmo quanto a relação do custo mínimo encontrado e o tempo necessário de execução para encontrá-lo. Para esse teste foram utilizados de cinquenta a seiscentos indivíduos, com variação em intervalos de cinquenta. O resultado é apresentado na tabela 3, os quais mostram que, embora o tempo de execução aumente na mesma proporção em que a quantidade de paradas, o menor custo encontrado sofre uma grande queda nas primeiras instâncias do teste para depois entrar numa redução gradual e lenta. Por fim, um último conjunto de testes foi realizado considerando-se o número de nós que a unidade de transporte obrigatoriamente deve passar. Para isso, foram utilizados os seguintes valores para os parâmetros genéticos: taxa elitismo igual a 10%, taxa de mutação igual a 35%, taxa de cruzamento igual a 80% e 600 indivíduos. Os resultados obtidos, apresentados na tabela 4, ilustram o bom desempenho do algoritmo para instâncias do problema com até 15 paradas num tempo de execução considerado factível para as condições do problema proposto, considerando um equipamento com as seguintes especificações: processador Intel Pentium IV 3 GHz com 512 MB de RAM rodando o sistema operacional Microsoft Windows XP Professional. Porém, em situações onde o número de paradas se aproxima de ou ultrapassa 20% da quantidade de vértices do grafo utilizado, fica evidente a necessidade de utilização de maiores populações e maiores limites para a quantidade de vértices nos caminhos iniciais, em comparação aos utilizados nos testes, pois o tempo de execução aumenta exponencialmente, além de ser encontrada grande divergência nos melhores caminhos em relação ao caminho ótimo.

Tabela 4: Média dos menores custos e do tempo de execução obtidos variando a quantidade de paradas.

| Quantidade de paradas | Menor custo possível | Melhor custo encontrado | Erro relativo | Tempo de execução |
|-----------------------|----------------------|-------------------------|---------------|-------------------|
| 5                     | 80                   | 82,7                    | 3,4%          | 2,648 s           |
| 10                    | 96                   | 106,6                   | 11,0%         | 5,000 s           |
| 20                    | 143                  | 194,6                   | 36,1%         | 13,564 s          |
| 40                    | 187                  | 281,8                   | 50,7%         | 723,487 s         |

## 5. CONCLUSÃO

Embora não se possa afirmar que a utilização do sistema de transporte coletivo imaginado venha a trazer melhora significativa se aplicado na prática, a implementação de um algoritmo, desenvolvido na forma de protótipo para resolver uma variação de problemas típicos de busca por melhor caminho, permitiu ao presente trabalho agregar valiosas experiências acerca da definição e construção de um algoritmo genético. Muitos dos detalhes apenas mencionados na literatura básica sobre o assunto puderam

ser explorados de maneira mais aprofundada, trazendo à tona obstáculos que, de outra forma, não poderiam ser evidenciados.

Em alguns momentos, descobriu-se que as técnicas e estratégias que se esperava serem as mais adequadas não retornavam resultados práticos: a codificação dos cromossomos desenvolvida para diminuir a complexidade de algumas funções do código torna outras excessivamente demoradas; o operador de cruzamento elaborado com a intenção de gerar apenas filhos válidos revelou não compensar o esforço computacional gasto para alcançar seu objetivo, já que os testes executados preliminarmente mostraram que a geração de indivíduos em grande quantidade e posterior descarte dos indivíduos inválidos é mais rápida e permite retornar o mesmo resultado; taxas de mutação mais elevadas que o aconselhado (Haupt, 2000) retornaram resultados significativamente melhores. Futuramente, seria interessante acrescentar ao problema restrições até o momento desprezadas: a capacidade máxima de passageiros por unidade de transporte, o fato de os pontos de parada de ônibus ficarem localizados entre dois cruzamentos (entre dois vértices), o destino desejado por cada usuário ao acionar o sistema de transporte, entre outros.

## REFERÊNCIAS

- Caetano, D. J. Um Sistema informatizado de Apoio a Usuários de Transporte Coletivo”, Dissertação de mestrado, Universidade de São Paulo, 2005.
- Davis, L. Handbook of Genetic Algorithms, New York: Van Nostrand Reinhold, 1991.
- Drozdek, A. Estrutura de Dados e Algoritmos em C++, Tradução de Luiz Sérgio de Castro Paiva. São Paulo: Pioneira Thomson Learning, 2002.
- Fogel, L. J.; Owens, A. J.; Walsh, M. J. Artificial Intelligence through Simulated Evolution, New York: Wiley Publishing, 1966.
- Goldberg, D. E. Genetic Algorithms in Search, Optimization, and Machine Learning, Reading: Addison-Wesley Longman, Inc, 1989.
- Haupt, R. L. "Optimum Population Size and Mutation Rate for a Simple Real Genetic Algorithm that Optimizes Array Factors”, 2000 AP-S Symposium, Salt Lake City, UT, Jul 00.
- Holland, J. Adaptation in Natural and Artificial Systems, Ann Arbor: Univ. Michigan, 1975.
- Koza, J. R. Genetic Programming: on the Programming of Computers by Means of Natural Selection, Cambridge: The MIT Press, 1992.
- McCormack, E. “Using a GIS to Enhance the Value of travel Diaries”. Institute of Transportation Engineers. ITE Journal, v. 69, n.1, p.38-43, 1999.
- Michalewicz, Z.; Fogel, D. B. How to Solve it: Modern Heuristics. [S.I.], Springer, 2000.
- Rechenberg, I. Evolutionsstrategie: Optimierung technischer Systeme nach Prinzipien der biologischen Evolution. Stuttgart: Frommann-Holzboog, 1973.