

Relatório #3 de Computação Evolucionária da otimização de rotas em uma malha de metro

Thiago Brandenburg

¹Universidade do Estado de Santa Catarina

1. Introdução

O problema da otimização de rotas em uma malha de metro consiste em encontrar o caminho mais rápido entre duas estações, considerando que os três andam à 40 km/h e que a troca de estação leva 5 minutos. Para este trabalho, foi utilizado uma malha com 14 estações, enumeradas de E1 à E14, conforme a figura 1

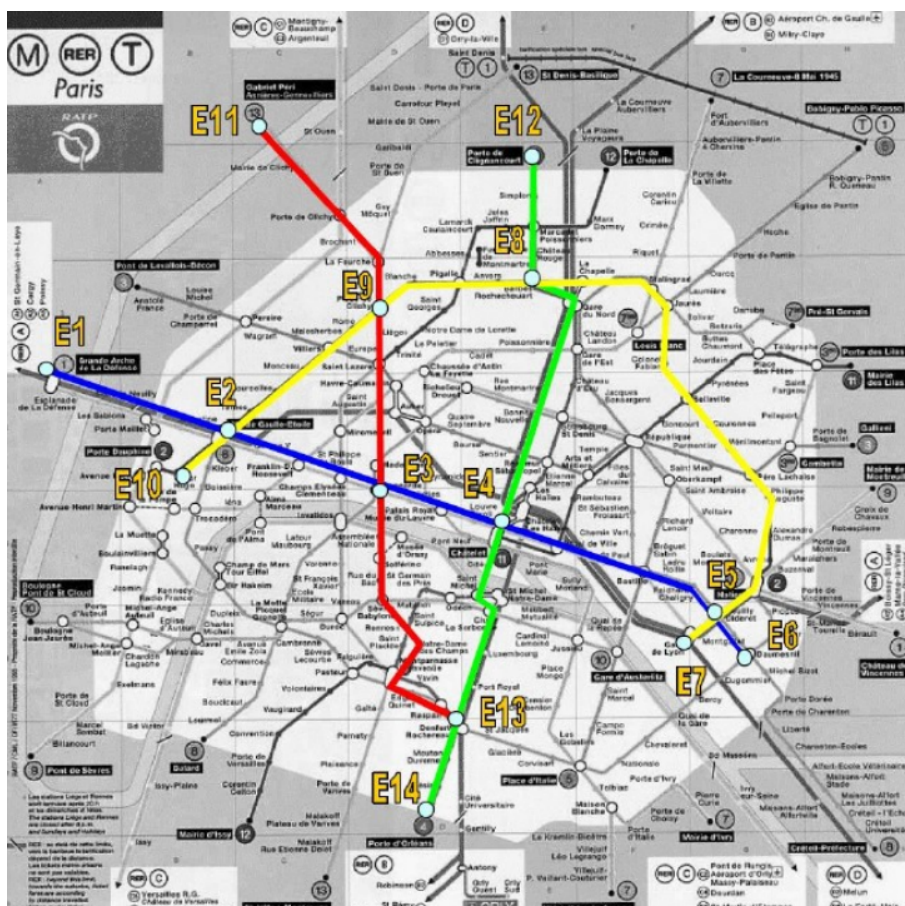


Figura 1. Malha de metrô com 14 estações

A representação computacional deste problema foi feita por meio de duas matrizes de vizinhança, a matriz 1 fornece as distâncias para estações que estão conectadas entre si, enquanto a matriz 2 fornece as distâncias em linha reta entre as estações. As matrizes 1 e 2 estão representadas nas figuras 2 e 3 a seguir:

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14
E1	-	10												
E2		-	8,5						10	3,5				
E3			-	6,3					9,4				18,7	
E4				-	13			15,3					12,8	
E5					-	3	2,4	30						
E6						-								
E7							-							
E8								-	9,6			6,4		
E9									-		12,2			
E10										-				
E11											-			
E12												-		
E13													-	5,1
E14														-

Figura 2. Distâncias entre as estações conectadas

	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12	E13	E14
E1	-	10	18,5	24,8	36,4	38,8	35,8	25,4	17,6	9,1	16,7	27,3	27,6	29,8
E2		-	8,5	14,8	26,6	29,1	26,1	17,3	10	3,5	15,5	20,9	19,1	21,8
E3			-	6,3	18,2	20,6	17,6	13,6	9,4	10,3	19,5	19,1	12,1	16,6
E4				-	12	14,4	11,5	12,4	12,6	16,7	23,6	18,6	10,6	15,4
E5					-	3	2,4	19,4	23,3	28,2	34,2	24,8	14,5	17,9
E6						-	3,3	22,3	25,7	30,3	36,7	27,6	15,2	18,2
E7							-	20	23	27,3	34,2	25,7	12,4	15,6
E8								-	8,2	20,3	16,1	6,4	22,7	27,6
E9									-	13,5	11,2	10,9	21,2	26,6
E10										-	17,6	24,2	18,7	21,2
E11											-	14,2	31,5	35,5
E12												-	28,8	33,6
E13													-	5,1
E14														-

Figura 3. Distâncias em linha reta entre as estações

Para esse problema, foi proposto encontrar o melhor caminho entre as estações E14 para E7 e E11 para E6.

1.1. Modelagem

Inicialmente, verificou-se a possibilidade de aplicar a a codificação de inteiro permutado para o problema, onde uma solução seria representa pela sequencia aleatória de estações entre a origem e o destino. Entretanto, considerando que uma estação tem muito menos vizinhos do que o número total de estações, a chance de cada alelo possuir um trecho inválido é alta, o que se propaga no cromossomo, tornando a chance da rota não ter violações remota. Portanto, essa codificação foi considerada imprópria.

A outra codificação testada foi a codificação real. Nesta codificação, cada alelo representa uma escolha dentre as estações vizinhas (que estão conectadas diretamente por uma linha) . Esta codificação funciona da seguinte forma:

1. Uma estação qualquer possui n estações vizinhas.
2. Gera-se um número aleatório de 0 à 1 (alelo).
3. Ordena-se as estações vizinhas de 0 à n-1, para uma estação vizinha qualquer k, se o valor da coordenada cai dentro do intervalo $[\frac{k}{n}, \frac{k+1}{n})$, esse é o trecho escolhido.

4. Exemplo: Se para uma estação há 4 vizinhos, a vizinha 1 é selecionada se o valor do alelo está contido no intervalo $[0.0, 0.25)$, para as outras estações, os intervalos seriam $[0.25, 0.5)$, $[0.5, 0.75)$ e $[0.75, 1.0)$ respectivamente.
5. Caso a estação seja a destino, finaliza-se a decodificação. Caso contrário, passa-se para o valor seguinte.

Essa codificação permite a minimização do número de soluções inválidas, pois o caminho é construído a partir de uma sequência válida de estações. Também foi realizado uma adaptação onde estações já percorridas são removidas das vizinhanças, o que permite maior eficiência nos caminhos gerados. O tamanho do caminho foi definido como o número total de estações, o que funciona para este problema de pequena escala, mas que talvez precise de uma redução para problemas com mais estações.

A função objetivo é definida como a distância até a estação destino, caso as estações sejam vizinhas é dado peso 1, caso as estações não sejam vizinhas o peso é aumentado duas vezes. Apesar do peso para a distância ser uma característica mais própria de penalidade, ele foi alocado na função objetivo para simplificar o algoritmo, concentrando os conceitos de distância na função objetivo. A normalização da função objetivo é feita utilizando a maior distância entre as estações. Como função de penalidade, foi utilizado o tempo de viagem, em horas. O tempo gasto para trocar de linha é 5 minutos, e a velocidade média dos trens é 40 km/h. Portanto a função *fitness*, consiste em minimizar a distância até a estação destino e o tempo de viagem, de forma que a solução ótima é o caminho mais rápido até a estação destino.

No que se diz ao algoritmo genético, foi utilizado uma população de tamanho 10, com 1 elitista. Como método de seleção, foi utilizado o método do torneio, com 90% de chance de sucesso para melhor indivíduo. Como método de *crossover* foi utilizado o método de 1 ponto, com chance de 80%. Para mutação, foi utilizado mutação aleatória com chance de 1%. O peso da penalidade foi de 0.5. Foi utilizado 100 gerações.

2. Resultados

Para os dois problemas fornecidos (E14 à E7 e E11 à E6), o algoritmo foi executado 30 vezes. Inicialmente, será analisado o comportamento dos dois problemas pelo gráfico de convergência das 3 primeiras execuções, em seguida será analisado o gráfico de caixa das duas execuções. Para o problema 1, a imagem 4 nos fornece os gráficos de convergência:

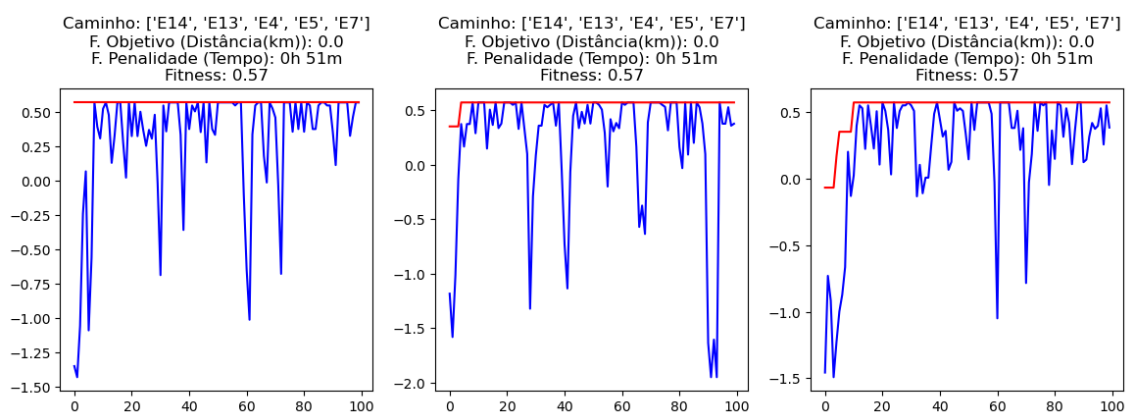


Figura 4. Gráficos de convergência das 3 primeiras execuções do problema 1

Verifica-se que nas três execuções encontrou-se a solução $E14 \rightarrow E13 \rightarrow E4 \rightarrow E5 \rightarrow E7$, alcançando a estação final em 51 minutos, passando por 5 estações. Enquanto na primeira execução a solução apareceu na população inicial, para as outras duas apresenta-se o comportamento evolutivo, com convergência nas primeiras iterações.

Para o segundo problema, a figura 5 apresenta os gráficos de convergência das 3 primeiras soluções:

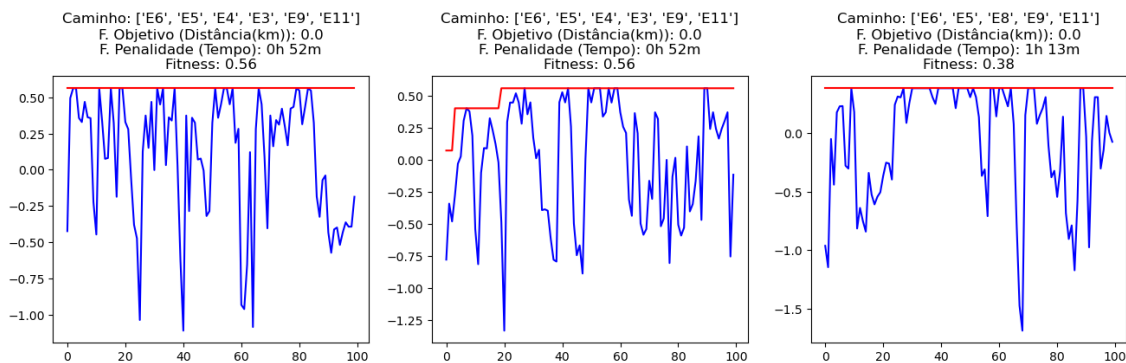


Figura 5. Gráficos de convergência das 3 primeiras execuções do problema 2

Diferente do problema 1, agora temos execuções diferentes alcançando valores diferentes, enquanto as execuções 1 e 2 alcançaram o trecho válido de 52min de duração, a execução 3 encontrou outro trecho válido de 1h13min.

Agora que temos um conhecimento do comportamento das execuções, é interessante analisar as distribuições das execuções. A figura 6 mostra o gráfico de caixa da *fitness* das duas execuções.

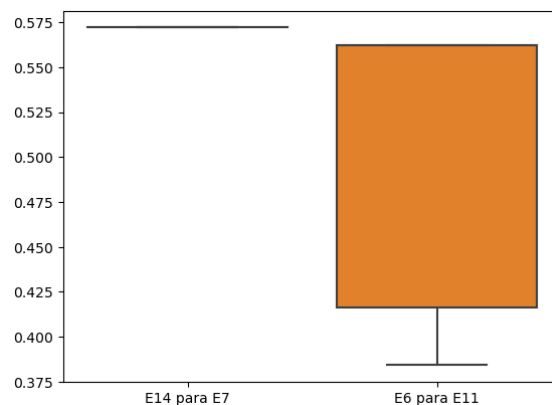


Figura 6. Boxplot da *fitness* das execuções

Observa-se que para o problema 1, todas as execuções convergem para o caminho $E14 \rightarrow E13 \rightarrow E4 \rightarrow E5 \rightarrow E7$ (51 minutos), enquanto para o problema existem múltiplos caminhos ótimos locais. Para o problema 2, a melhor solução encontrada foi $'E6' \rightarrow 'E5' \rightarrow 'E4' \rightarrow 'E3' \rightarrow 'E9' \rightarrow 'E11'$ (52 minutos), limitando superiormente o gráfico de caixa do problema 2. É interessante observar que para a mesma malha, dependendo das estações origem e destino podem obter características bastante distintas para o problema.

3. Conclusão

O problema da alocação de rotas de metro pode ser codificado para o algoritmo genético. Enquanto diversas codificações são possíveis, a codificação real a partir de decisões do caminho em cada estação se apresentou adequada para o problema. Foi possível codificar a *fitness* definindo a distância até a estação destino como função objetivo e o tempo de viagem como penalidade.

Enquanto o algoritmo sempre obteve caminhos válidos para os problemas testados, nem sempre o caminho escolhido é o mais rápido. A complexidade do problema não é só dependente do número de estações, pois dependendo das estações origem e destino, o problema pode ter múltiplos ótimos locais, portanto é necessário levar esse fator em consideração para problemas mais complexos com mais estações

No geral, a aplicação do algoritmo genético no problema da alocação de rotas é recomendada, com adequações para cenários com muitas estações e trechos com muitas paradas.