

December 27, 2025

1 Introduction

Este projeto tem como objetivo apoiar os utilizadores na escolha da melhor forma de deslocação no Grande Porto, permitindo a seleção de um ponto de partida e de chegada e a combinação de diferentes meios de transporte: deslocação a pé, metro e autocarro (STCP). O sistema desenvolvido considera que o utilizador pode deslocar-se a pé desde qualquer ponto até uma paragem de metro ou autocarro e que, em estações multimodais, pode escolher o meio de transporte a utilizar.

Atualmente, os utilizadores apresentam múltiplos objetivos, tais como minimizar o tempo total de viagem, reduzir a emissão de gases com efeito de estufa (CO₂). Adicionalmente, poderão existir restrições como por exemplo a limitação do número de transbordos ou um tempo máximo admissível de deslocação a pé.

2 Formulação do Problema

Este projeto enquadra-se, no domínio da otimização multiobjetivo, sendo particularmente adequado à aplicação de algoritmos evolutivos capazes de gerar múltiplas soluções de compromisso. Neste trabalho, foi adotada uma abordagem inspirada no algoritmo MOEA/D.

O problema foi modelado como um problema de otimização de caminhos num grafo multimodal. Dado um nó de origem e um nó de destino, pretende-se encontrar percursos válidos que minimizem simultaneamente dois objetivos contraditórios principais:

- Tempo total de viagem;
- Emissões de CO₂ associadas à deslocação.

Assim como respeitar as restrições apresentadas, nomeadamente:

- Número de transbordos;
- Número de trocas entre linhas
- Tempo máximo de deslocação a pé.

Cada solução corresponde a um caminho válido no grafo, respeitando as ligações entre os diferentes modos de transporte.

3 Dados e Construção do Grafo

3.1 Dados Utilizados

Foram utilizados dados reais disponibilizados pelas seguintes entidades:

STCP – Sociedade de Transportes Colectivos do Porto;
Metro do Porto.

Os dados estão organizados em ficheiros gtfs (General Transit Feed Specification) e incluem informação sobre paragens, estações, linhas e respetiva localização geográfica, permitindo a construção de um grafo representativo da rede de transportes do Grande Porto.

De forma geral, foram utilizados os seguintes tipos de ficheiros:

Ficheiros de paragens/estações (stops.txt): contêm a identificação única de cada paragem ou estação, bem como a sua localização geográfica (latitude e longitude). Estes ficheiros são a principal fonte para a criação dos nós do grafo.

Ficheiros de linhas/rotas(trips.txt & routes.txt): descrevem as linhas de autocarro (STCP) e de metro, especificando a sequência ordenada de paragens ou estações servidas por cada linha. Estes ficheiros são utilizados para criar as arestas associadas aos diferentes meios de transporte.

Ficheiros de horários e datas especiais (calendar.txt & calendar_dates.txt): fornecem informação temporal que pode ser usada para estimar o tempo de deslocação entre paragens consecutivas.

3.2 Modelação Multimodal

O sistema é modelado através de um grafo multimodal ponderado, no qual diferentes meios de transporte coexistem e se interligam. Formalmente, o grafo é definido como $G = (V, E)$, onde V representa o conjunto de nós e E o conjunto de arestas.

3.2.1 Nós do Grafo

Os nós do grafo representam diferentes tipos de pontos:

- Nós de paragens de autocarro, associados à rede STCP;
- Nós de estações de metro, associados à rede do Metro do Porto.

Cada nó contém informação adicional, nomeadamente:

- Identificador único;
- Tipo de nó (autocarro ou metro);
- Coordenadas geográficas;
- Conjunto de modos de transporte disponíveis.

Ao se incializar o processo de procura de um caminho mais curto criamos os nodes `_START_` e `_END_` que correspondem às coordenadas de onde o utilizador começa e de onde quer chegar. Geralmente tem de andar do seu ponto inicial até uma paragem.

3.2.2 Nós Multimodais

Os nós multimodais desempenham um papel central na integração dos diferentes meios de transporte. Um nó é considerado multimodal quando permite a troca entre dois ou mais modos de transporte, por exemplo:

- Estações de metro com paragens de autocarro nas proximidades;

A construção dos nós multimodais é realizada através de um critério de proximidade espacial. Paragens e estações cuja distância geográfica seja inferior a um determinado limite (no nosso caso foi escolhido 50 metros) são agregadas logicamente, permitindo a criação de ligações de transbordo entre os diferentes modos de transporte.

É efetuado um clustering das paragens dentro desse raio. É calculado um centro que passa a ser as coordenadas de um node multimodal. Desta forma fazemos com que as arestas que ligavam às paragens do cluster agora ligam ao ponto multimodal.

3.2.3 Arestas e Tipos de Ligações

As arestas do grafo representam deslocações possíveis entre nós e são classificadas de acordo com o modo de transporte:

- Arestas de autocarro: ligam paragens consecutivas pertencentes à mesma linha STCP;
- Arestas de metro: ligam estações consecutivas na rede do Metro do Porto;

- Arestras de transbordo: ligam nós multimodais, permitindo a mudança de meio de transporte.

Cada arestra possui pesos associados, nomeadamente:

- Tempo estimado de deslocação;
- Distância percorrida;
- Emissões de CO₂ (quando aplicável).

A distância é calculada usando a Distância Haversine. É a forma de calcular a distância entre dois pontos na Terra, dados as suas latitudes e longitudes, tendo em conta a curvatura do planeta.

3.2.4 Cálculo dos Pesos

O tempo de deslocação é estimado com base no tipo de transporte, distância e a sua velocidade média:

- WALK SPEED KMH = 4.5
- BUS SPEED KMH = 18.0
- METRO SPEED KMH = 30.0

As emissões de CO₂ são calculadas apenas para os segmentos realizados em transporte público, utilizando os valores fornecidos no enunciado:

- STCP: 109,9 gCO₂ por passageiro-km;
- Metro do Porto: 40 gCO₂ por passageiro-km.

Ao serem criados os nodes de inicio e fim do percurso, criamos arestras que os ligam a todas as paragens num determinado raio. Estas arestras não têm associadas um valor de emissão de CO₂.

3.2.5 Grafo final

4 Metodologia de Otimização

4.1 Algoritmo Utilizado

Foi utilizada uma abordagem baseada no MOEA/D (Multi-Objective Evolutionary Algorithm based on Decomposition), um algoritmo multiobjetivo que decompõe o problema em vários subproblemas.

A escolha do MOEA/D justifica-se pela sua capacidade de lidar eficientemente com objetivos contraditórios e pela sua ampla utilização em problemas de otimização multiobjetivo.

No contexto deste projeto, cada subproblema procura otimizar um compromisso diferente entre:

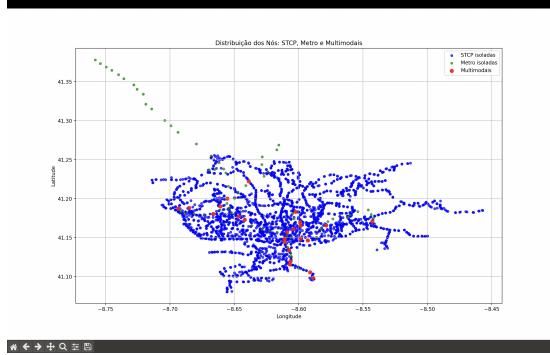


Figure 1: Paragens de transportes Públicos do Porto.

- minimizar o tempo total de viagem;
- minimizar as emissões de CO₂.

A escolha do MOEA/D justifica-se por várias razões:

- Permite obter um conjunto diversificado de soluções de compromisso (fronteira de Pareto);
- Apresenta boa escalabilidade para problemas com grafos de grande dimensão;
- Facilita a incorporação de heurísticas e operadores com conhecimento do domínio;

4.2 Codificação das Soluções

Cada indivíduo da população representa um caminho completo entre origem e destino, codificado como uma sequência ordenada de nós do grafo.

Esta representação garante:

- A validade estrutural das soluções;
- Facilidade na aplicação de operadores evolutivos;
- Integração natural de diferentes modos de transporte.
- Permite identificar facilmente transbordos e segmentos a pé;

4.3 Função Objetivo

O problema abordado neste projeto é formulado como um problema de *otimização multiobjetivo*, no qual se pretende encontrar percursos que estabeleçam compromissos adequados entre critérios contraditórios. Em particular, são considerados dois objetivos principais: o tempo total de viagem e as emissões de CO₂.

4.3.1 Minimização do Tempo Total de Viagem

O primeiro objetivo consiste em minimizar o tempo total necessário para realizar a deslocação entre o node *start* e o *end*. O tempo total de um percurso é calculado como a soma dos tempos associados a cada segmento do caminho:

$$f_1(\text{caminho}) = \sum_{e \in E_{\text{caminho}}} t(e)$$

onde $t(e)$ representa o tempo estimado para percorrer a aresta e .

O cálculo do tempo depende do tipo de deslocação associado à aresta:

- **Deslocações a pé:** o tempo é estimado com base numa velocidade média de caminhada, assumida constante;
- **Autocarro (STCP):** o tempo é estimado a partir da distância percorrida e de uma velocidade média;
- **Metro do Porto:** o tempo é estimado a partir da distância percorrida e de uma velocidade média;

4.3.2 Minimização das Emissões de CO₂

O segundo objetivo visa minimizar o impacto ambiental da deslocação, medido através das emissões totais de CO₂ associadas ao percurso:

$$f_2(\text{caminho}) = \sum_{e \in E_{\text{caminho}}} \text{CO}_2(e)$$

As emissões são calculadas exclusivamente para os segmentos realizados em transporte público, sendo as deslocações a pé consideradas neutras em termos de emissões.

Os valores utilizados seguem os dados fornecidos no enunciado do projeto:

- **Autocarro (STCP):** 109.9 gCO₂ por passageiro·km;
- **Metro do Porto:** 40 gCO₂ por passageiro·km;
- **Deslocações a pé:** não têm emissão de CO₂.

Assim, para cada aresta correspondente a um segmento de transporte público, a emissão é calculada como:

$$\text{CO}_2(e) = d(e) \times \text{fator}_{\text{modo}}$$

onde $d(e)$ representa a distância percorrida na aresta e $\text{fator}_{\text{modo}}$ corresponde ao fator de emissão do meio de transporte utilizado.

4.3.3 Tratamento das Restrições e Penalizações

Para além dos objetivos principais, o problema inclui restrições adicionais relevantes do ponto de vista do utilizador, nomeadamente:

- limite máximo de tempo de deslocação a pé;
- limitação do número de transbordos;
- limitação do número de linhas utilizadas.

Estas restrições são tratadas através de penalizações adicionadas às funções objetivo, permitindo desencorajar soluções inviáveis sem restringir excessivamente o espaço de pesquisa do algoritmo evolutivo.

As funções objetivo penalizadas são então definidas como:

$$f'_1(\text{caminho}) = f_1(\text{caminho}) + P_{\text{pé}}(\text{caminho}) + P_{\text{trans}}(\text{caminho}) + P_{\text{lin}}(\text{caminho})$$

$$f'_2(\text{caminho}) = f_2(\text{caminho}) + \gamma \cdot (P_{\text{trans}}(\text{caminho}) + P_{\text{lin}}(\text{caminho}))$$

4.3.4 Natureza Multiobjetivo e Decomposição

As duas funções objetivo consideradas são contraditórias: percursos mais rápidos tendem a recorrer a meios de transporte mais poluentes, enquanto percursos ambientalmente mais sustentáveis podem implicar maior tempo de viagem ou maior componente pedonal.

No contexto do MOEA/D, estas funções são combinadas através de vetores de pesos, permitindo definir subproblemas escalares da forma:

$$g(\text{caminho} | \lambda) = \lambda_1 f'_1(\text{caminho}) + \lambda_2 f'_2(\text{caminho})$$

onde λ_1 e λ_2 representam os pesos associados ao tempo e às emissões, respetivamente.

Esta decomposição permite explorar diferentes compromissos entre os objetivos e construir uma fronteira de Pareto diversificada.

4.3.5 Justificação das Escolhas

A escolha do tempo de viagem e das emissões de CO₂ como objetivos principais justifica-se por representarem critérios relevantes e contraditórios em contextos reais de mobilidade urbana. A utilização de penalizações para o tratamento das restrições permite manter a flexibilidade do algoritmo evolutivo, assegurando simultaneamente a viabilidade prática das soluções obtidas.

4.4 Inicialização da População

A inicialização da população desempenha um papel fundamental no desempenho de algoritmos evolutivos, em particular em problemas de otimização combinatoria com um espaço de soluções vasto e altamente estruturado, como é o caso do planeamento de percursos multimodais.

Antes sequer de inicializar a população, através do calendário e horários fornecidos, encontramos e calculamos o grafo de paragens disponíveis no dia em questão. A seguir criaram-se os nodes de inicio e fim de percurso e as arestas respetivas.

A função `initialize_population` foi desenhada de forma a garantir a viabilidade e a diversidade das soluções iniciais. O processo segue os seguintes passos:

1. **Cálculo da Solução Mais Rápida:** É executado o algoritmo de Dijkstra para inserir na população um caminho de referência focado na rapidez. Escolhemos o algoritmo de Dijkstra pois ele ajuda o MOEA/D a convergir muito mais rapidamente, "ensinando" desde o início onde estão os caminhos mais eficientes.
2. **Amostragem Estocástica:** Os restantes indivíduos da população são calculados seguindo a lógica de decomposição por pesos do MOEA/D. Este processo garante que a população não convirja apenas para uma solução, mas sim para uma Fronteira de Pareto diversificada.
 - **Vetores de Pesos (λ):** São gerados N vetores de pesos uniformemente distribuídos. Cada vetor define a importância relativa entre o tempo de viagem e as emissões de CO₂. Por exemplo, um indivíduo pode ser incentivado a minimizar agressivamente o CO₂, enquanto outro busca equilibrar os dois objetivos.
 - **Escalarização Tchebycheff:** Para cada subproblema, a qualidade de um caminho é avaliada através de uma função de escalarização. Isto permite transformar os dois objetivos em um único valor que o algoritmo tenta minimizar para aquele indivíduo específico.
 - **Refinamento por Vizinhança:** Os restantes indivíduos não são gerados de forma isolada. O algoritmo define uma "vizinhança" para cada subproblema. Quando um novo caminho é descoberto (por exemplo, através de mutação), ele é testado nos vizinhos. Se o novo percurso for melhor para os pesos de um vizinho, ele substitui a solução antiga, promovendo a evolução de toda a população em paralelo.
 - **Penalizações de Viabilidade:** Durante este cálculo, a função objetivo (`path_objective`) aplica penalizações pesadas ao tempo total se o caminho gerado violar limites de mudanças de linha ou tempo de caminhada excessivo, garantindo que os indivíduos evoluam para soluções utilizáveis pelo passageiro.
3. **Validação de Restrições:** Cada caminho gerado é verificado quanto à sua continuidade temporal (horários das próximas passagens) e física (ligações entre paragens e clusters multimodais).

Esta abordagem evita o problema comum em algoritmos evolutivos aplicados a grafos, onde uma inicialização puramente aleatória resultaria numa população composta maioritariamente por indivíduos inválidos (caminhos interrompidos).

4.5 Operadores Genéticos e Procura Local

Dada a natureza do problema em grafos, os operadores genéticos foram desenhados de forma a manter a validade dos caminhos (conectividade) e não puramente aleatórios.

4.5.1 Operador de Mutação de Caminho

A mutação é o motor de exploração do algoritmo. Em vez de alterar um "gene" isolado, o operador de mutação atua sobre a estrutura do trajeto:

1. **Seleção de Segmento:** Escolhem-se aleatoriamente dois pontos, A e B , dentro de um caminho existente.
2. **Re-routing:** O segmento original entre A e B é removido e o algoritmo tenta encontrar um novo sub-caminho viável no grafo.
3. **Diversificação:** Este processo permite que o sistema descubra novas alternativas (ex: trocar uma parte realizada de autocarro por um de metro) que o operador de crossover dificilmente encontraria sozinho. Permite também que não fique preso em possíveis máximos locais.

4.5.2 Algoritmos de Procura Local (Local Search)

Após a geração de novos indivíduos, aplica-se uma procura local para refinar a solução antes de a reinserir na população. Esta atua em três frentes:

- **Substituição de Segmentos:** O algoritmo verifica se, para um par de paragens no trajeto, existe uma linha alternativa ou uma combinação multimodal que reduza o custo (seja em tempo ou CO_2) sem alterar o resto do percurso.
- **Otimização de Transbordos:** A procura local analisa pontos de transferência. Se detetar uma caminhada curta (cluster de 50m) que permita apanhar um transporte mais rápido ou menos poluente, o trajeto é ajustado localmente.
- **Ajuste Pedonal:** Verifica se a distância a pé no início e fim da viagem pode ser reduzida ao ligar o utilizador a paragens vizinhas alternativas dentro do raio de 800m.

Esta combinação de procura global (MOEA/D) com procura local (ajuste de segmentos) é o que permite encontrar soluções que não são apenas matematicamente ótimas, mas logicamente viáveis para um passageiro real.

4.6 Parametrização e Configuração Experimental

A eficácia do algoritmo MOEA/D depende da calibração dos seus hiperparâmetros, que foram ajustados para refletir a realidade da rede de transportes do Porto:

- **Tamanho da População (N):** Definido para permitir uma representação fiel da Frente de Pareto (ex: entre 50 a 100 indivíduos), garantindo soluções que variam entre o foco total em tempo e o foco total em sustentabilidade.
- **Vetores de Peso (λ):** Gerados para decompor o problema multiobjetivo em subproblemas monoobjetivo. Cada indivíduo na população tenta minimizar a função de escalarização para o seu respetivo peso.
- **Penalizações Físicas:** Utilização de constantes rigorosas no cálculo do custo, como a penalização por mudança de modo (MODE_CHANGE_PENALTY_MIN) e de linha (LINE_CHANGE_PENALTY_MIN), essenciais para evitar trajetos com transbordos excessivos e pouco práticos.

5 Metodologia de Avaliação

A robustez do sistema foi testada utilizando a rede real de transportes da Área Metropolitana do Porto, combinando os sistemas STCP e Metro.

- **Cenários de Teste:** Foram selecionados pares origem-destino de forma estratégica (por exemplo, zonas periféricas sem Metro versus polos universitários). O sistema gera nós virtuais (_START_ e _END_) e utiliza a fórmula de Haversine para conectar o utilizador às paragens num raio de 800 m.
- **Métricas de Qualidade:**
 1. **Convergência e Hipervolume:** Avaliou-se a capacidade do MOEA/D em aproximar-se da Frente de Pareto real, garantindo que as soluções de "caminho mais rápido" e "caminho mais ecológico" fossem distintas e competitivas.
 2. **Análise de Estabilidade:** Verificou-se a sensibilidade do algoritmo em relação às penalizações de transbordo (MODE_CHANGE_PENALTY), assegurando que as rotas sugeridas fossem logisticamente confortáveis para um passageiro humano.

6 Software

A implementação foi realizada em **Python**, focada na eficiência de processamento de grafos de grande escala:

- **Pandas:** Utilizado para o tratamento massivo de dados GTFS (mais de 900.000 arestas geradas no pré-processamento).

- **Heapq:** Implementação de filas de prioridade para o algoritmo de Dijkstra, permitindo uma inicialização rápida da população.
- **Matplotlib:** Utilizado para a validação visual do grafo multimodal, permitindo identificar a correta ligação dos 24 nós multimodais (clusters) criados entre as redes de autocarro e metro.
- **Fórmula de Haversine:** Implementada para garantir que as distâncias espaciais entre coordenadas GPS fossem precisas ao criar os clusters de 50 metros para transferências.

7 Discussão de Resultados

Os testes demonstraram que a inclusão de Procura Local e de uma Mutação por Re-encaminhamento foi decisiva para a viabilidade do projeto:

- **Refinamento da Mutação:** Ao contrário de mutações aleatórias que ”quebram” caminhos em grafos, a mutação de sub-caminhos permitiu explorar alternativas viáveis entre paragens, resultando numa convergência mais rápida para a Frente de Pareto.
- **Equilíbrio Multimodal:** O algoritmo demonstrou eficácia ao sugerir o uso do Metro para longas distâncias (otimizando tempo e CO_2) e o uso de STCP para a ”última milha” ou zonas sem cobertura ferroviária.
- **Impacto das Restrições:** No documento com alguns dos teste corridos podemos ver que as restrições têm o seu devido impacto pois se não limitarmos o tempo a pé este será sempre o caminho com menos emissões de CO_2 , o que é o resultado espectável.

8 Conclusão

O sistema desenvolvido revelou-se uma ferramenta eficaz para a mobilidade sustentável no Porto. A combinação do algoritmo MOEA/D com o processamento de dados GTFS permitiu gerar soluções de compromisso reais entre rapidez e impacto ambiental.

A integração de nós multimodais foi o avanço técnico mais significativo, e inicialmente mais difícil, permitindo a integração/compatibilidade entre operadores distintos. Como trabalho futuro, destaca-se:

1. A inclusão do custo monetário através das tarifas do metro.
2. A adaptação para dados em tempo real (**GTFS-Realtime**) para considerar atrasos e greves.
3. A expansão do grafo para incluir o formato específico das estradas.