

Avaliação de Técnicas de Estimação da Matriz Origem-Destino do Tráfego de Veículos em Cidades

Luciano Urgal Pando
Orientador: Ricardo Lüders

Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial
UTFPR

14/06/2018

Sumário

- 1 Introdução
- 2 Descrição do problema
- 3 Técnicas de estimação da matriz OD
- 4 Resultados
 - Estudo de caso do Porto (Portugal)
 - Estudo de caso de Curitiba (Brasil)
- 5 Conclusão

Introdução

Introdução

- **Compreender o padrão de deslocamento** é fundamental para o planejamento urbano. Fluxos entre diferentes origens e destinos (OD) podem ser representados pela chamada **matriz origem destino**;
- A matriz origem destino não é disponível diretamente, e então, **diversas técnicas** têm sido aplicadas para a sua **estimação**;
- A dissertação realiza um **comparativo entre técnicas** de estimação da matriz origem-destino.

Introdução

Motivação

- **Planejamento urbano** utilizando dados provenientes de **monitoração**, como sensores de tráfego. Geração crescente de **dados**;
- **Interesse econômico**: Pesquisas de campo sobre origem destino são caras, demoradas e trabalhosas. (Pesquisa OD de Curitiba com custo aproximado de 6 milhões de reais);
- **Diversas técnicas** de estimação da matriz OD disponíveis na literatura. Auxiliar na decisão de **qual técnica utilizar** em trabalhos futuros.

Introdução

Objetivo geral

- Avaliar diferentes técnicas de estimação da matriz OD que utilizam dados de sensores, usando métricas já definidas na literatura.

Objetivos específicos

- Comparar **PSO**, **GA**, **LS** e **MILP** na estimação da matriz OD;
- Considerar as cidades do **Porto e Curitiba** como estudos de caso, pois utilizam diferentes fontes de dados de fluxo de veículos para a estimação da matriz origem-destino;
- Para o estudo de caso de Porto, **Identificar origens e destinos**, assim como a **distribuição do fluxo** de veículos **na rede viária**, quando se utilizam dados de viagens de táxi para a estimação da matriz origem-destino;
- Avaliar a influência da qualidade e quantidade de informação proveniente dos sensores (**número e posicionamento dos sensores**).

Matriz Origem-Destino

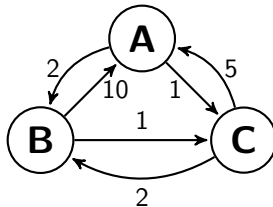
Tabela: Informações de fluxos OD entre nós A, B e C em veículos/min da Matriz origem-destino X

	A	B	C
A		2	1
B	10		1
C	5	2	

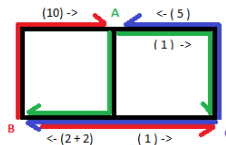
- Linhas representam as origens;
- Colunas representam os destinos.

Matriz Origem-Destino

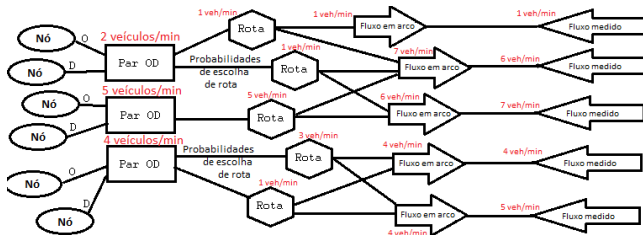
Representação em grafo direcionado da Matriz OD



Matriz OD distribuída em uma rede viária



Trafégo da Matriz Origem-Destino (ODM) na rede viária



- Conjunto de pares origem-destino (OD) $w \in W$
- Fluxo de veículos por par OD (Matriz OD) q_w
- Conjunto de rotas por par OD
- Probabilidade de uso de uma rota por par OD
- Prob. de uso de um arco por par OD: **(traffic) Assignment Matrix**

Estimação da Matriz Origem-Destino

Distribuição do tráfego de acordo com q_w e p_{wz}

$$v_z = \sum_{w \in W} q_w \cdot p_{wz} + e \quad (1)$$

$$\text{s.t. } q_w \geq 0 \quad (2)$$

Estimação de \bar{q}_w através \bar{v}_z

$$\forall z (v_z \approx \bar{v}_z) \quad (3)$$

$$\forall w (q_w \approx \bar{q}_w) \quad (4)$$

Estimação da Matriz Origem-Destino

Dados de entrada

- Conjunto de pares origem-destino (OD) $w \in W$
- Topologia da rede (Grafo direcionado)
- Medições de sensores de tráfego v_z
- Probabilidade de uso de um arco por par OD: **(traffic)**
Assignment Matrix

Saída

- Fluxo de veículos q_w por par OD (Matriz OD)

Algoritmo genético (GA)

Cada indivíduo é uma solução candidata contendo np valores reais;
Utiliza parâmetros $P_{crossover}$, $P_{mutacao}$, pop , t_{ga} .

Algorithm 2 Pseudo-código do GA

- 1: Inicializa a população com pop indivíduos aleatórios
 - 2: Avalia a adequação de cada indivíduo de acordo com a função $fitness$
 - 3: $t=1$
 - 4: **while** $t \leq t_{ga}$ **do**
 - 5: Seleção de pais
 - 6: Realiza *crossover*. População passa a ser ($2 \cdot pop$)
 - 7: Realiza mutação dos novos indivíduos
 - 8: Avalia a adequação dos novos indivíduos
 - 9: Somente os (pop) indivíduos de melhor adequação sobrevivem
 - 10: $t = t + 1$
 - 11: **end while**
-

Otimização por nuvem de partículas (PSO)

Cada partícula é uma solução candidata contendo np valores reais;
Utiliza parâmetros w_i , w_f , $c1$, $c2$, s , t_{ps0} .

Algorithm 1 Pseudo-código do PSO

- 1: Inicializa o enxame com s partículas aleatórias
 - 2: Avalia a adequação de cada partícula de acordo com a função *fitness*
 - 3: \mathbf{best}^{global} é a partícula de melhor adequação
 - 4: $t=1$
 - 5: **while** $t \leq t_{ps0}$ **do**
 - 6: **for** Cada partícula \mathbf{p}^t **do**
 - 7: Calcula velocidade da partícula \mathbf{p}^t (Equação 25)
 - 8: Desloca partícula \mathbf{p}^t (Equação 26)
 - 9: Reavalia a adequação da partícula \mathbf{p}^t
 - 10: Atualiza a melhor posição da partícula (Melhor entre \mathbf{best}_p e \mathbf{p}^t)
 - 11: Atualiza a melhor posição global (Melhor entre \mathbf{best}^{global} e \mathbf{best}^t)
 - 12: **end for**
 - 13: $t = t + 1$
 - 14: **end while**
-

Funções *fitness* do GA e PSO

$$\min_{q_w} \sum_{z \in Z} |v_z - \bar{v}_z| \quad (5)$$

$$\min_{q_w} \sum_{z \in Z} (v_z - \bar{v}_z)^2 \quad (6)$$

$$\min_{q_w} \sum_{z \in Z} GEH(z) \quad (7)$$

$$GEH(z) = \frac{\sqrt{2(\bar{v}_z - v_z)^2}}{\bar{v}_z + v_z} \quad (8)$$

Mínimos quadrados (LS)

Dados de entrada

- *Assignment matrix* A ($ns \times np$);
- Fluxos b de veículos observados por sensores ($ns \times 1$).

Dados de saída - elementos da matriz OD

- Fluxos x de veículos estimados por par OD ($np \times 1$)

Solução pela pseudo-inversa $(A^T A)^{-1}$

$$Ax = b \quad (9)$$

$$x = (A^T A)^{-1} A^T b \quad (10)$$

Modelo de programação linear inteira mista (MILP)

Tabela 18: Variáveis do modelo de programação linear inteira mista

Variável	Tipo	Símbolo	Quantidade
Fluxo estimado por par OD w	\mathbb{R}	q_w	np pares OD
Fluxo estimado no arco z	\mathbb{R}	\bar{v}_z	ns sensores
Folga superior no arco z	\mathbb{R}	F_{z+}	ns sensores
Folga inferior no arco z	\mathbb{R}	F_{z-}	ns sensores
Total			$(np + 3 \cdot ns)$

Função *fitness*

$$\min_{q_w} \sum_{z \in Z} (F_{z+} + F_{z-}) \quad (11)$$

Modelo de programação linear inteira mista (MILP)

Restrições do modelo MILP

$$\bar{v}_z = \sum_{w \in W} p_{wz} \cdot q_w \quad (12)$$

$$v_z = (\bar{v}_z + F_{z+} - F_{z-}) \quad (13)$$

$$q_w \geq 0 \quad (14)$$

$$F_{z+} \geq 0 \quad (15)$$

$$F_{z-} \geq 0 \quad (16)$$

Métricas de desempenho

Tabela 16: Métricas utilizadas para avaliar a qualidade da ODME

	MAE	RMSE	R^2 ODM	TDD	GEH	pGEH5	R^2 arcs
Objetivo	Minimizar	Minimizar	Maximizar	Minimizar	Minimizar	Maximizar	Maximizar
Variável analisada	Matriz OD	Matriz OD	Matriz OD	Matriz OD	Arcos	Arcos	Arcos
Valores	$[0, +\infty)$	$[0, +\infty)$	$[0, 1]$	$[0, +\infty)$	$[0, +\infty)$	$[0, 1]$	$[0, 1]$

$$MAE = \frac{\sum_{w \in W} |\bar{q}_w - q_w|}{np} \quad (17)$$

$$RMSE = \frac{\sqrt{\frac{1}{np} \sum_{w \in W} (\bar{q}_w - q_w)^2}}{\sum_{w \in W} q_w} \quad (18)$$

$$TDD = \frac{|\sum_{w \in W} \bar{q}_w - \sum_{w \in W} q_w|}{\sum_{w \in W} q_w} \quad (19)$$

Métricas de desempenho

Tabela 16: Métricas utilizadas para avaliar a qualidade da ODME

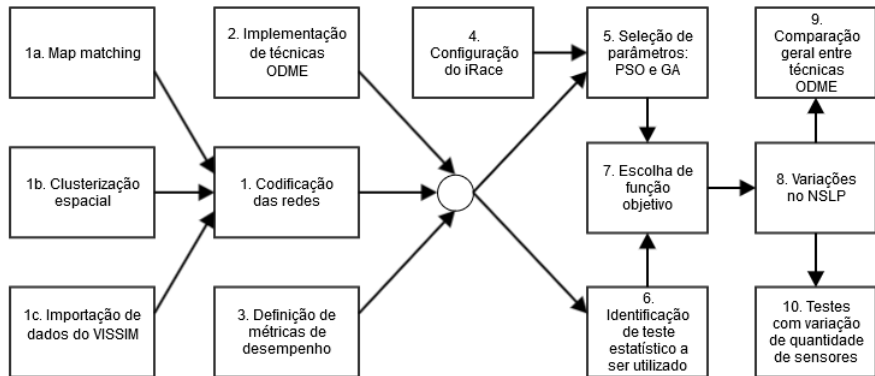
	MAE	RMSE	R^2 ODM	TDD	GEH	pGEH5	R^2 arcos
Objetivo	Minimizar	Minimizar	Maximizar	Minimizar	Minimizar	Maximizar	Maximizar
Variável analisada	Matriz OD	Matriz OD	Matriz OD	Matriz OD	Arcos	Arcos	Arcos
Valores	$[0, +\infty)$	$[0, +\infty)$	$[0, 1]$	$[0, +\infty)$	$[0, +\infty)$	$[0, 1]$	$[0, 1]$

$$R^2 = \left(\frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{(n \sum x^2 - (\sum x)^2)(n \sum y^2 - (\sum y)^2)}} \right)^2 \quad (20)$$

$$GEH(z) = \frac{\sqrt{2(\bar{v}_z - v_z)^2}}{\bar{v}_z + v_z} \quad (21)$$

$$GEH = \frac{\sum_{z \in Z} GEH(z)}{ns} \quad (22)$$

Organização das atividades desenvolvidas



Resumo dos estudos de caso: Porto e Curitiba

	Curitiba (Brasil)	Porto (Portugal)
Fontes dos dados <i>Assign. matrix</i>	IPPUC PTV VISUM	OSM e UCI (2015) Agrupamento espacial e Map matching
Nós	191	153840
Arcos	594	221446
Sensores	71	Agrupamento espacial, map matching e NSLP
Pontos OD	29	Agrupamento espacial
Pares OD	841	Agrupamento espacial
$q_w > 0$	183	Agrupamento espacial

Estudo de caso do Porto

- Agrupamento espacial;
- *Map matching*;
- Configuração dos algoritmos GA e PSO;
- Influência da função *fitness* na ODME;
- Influência da localização dos sensores na ODME;
- Comparação das técnicas de OMDE;
- Influência do número de sensores na ODME.

Dataset de corridas de taxi de Porto

Tabela 22: Atributos dos dados de corridas de táxi do Porto

Atributo	Tipo	Descrição
TRIP ID	String	Identificador da corrida
CALL TYPE	Caractere	A (corrida iniciada através da central de táxis) B (corrida iniciada em ponto de táxi) C (Outros casos)
ORIGIN CALL	Inteiro	ID do ponto de táxis, se CALL TYPE = B
TAXI ID	Inteiro	ID do motorista
TIMESTAMP	Inteiro	Início da viagem (segundos).
DATATYPE	Caractere	B (Feriado ou outro dia especial) C (dia anterior a um dia tipo B) A (demais dias)
MISSING DATA	Booleano	Completude do atributo POLYLINE
POLYLINE	String	Coordenadas GPS a cada 15 segundos para a corrida, no formato WGS84.

Agrupamento espacial baseado em densidade

Problema

- Cada corrida possui um ponto de origem e um de destino;
- É necessário definir áreas para serem OD;
- É necessário definir um número de clusters.

Solução

- Novo dataset contendo somente as origens e destinos;
- CascadeKMeans para definir número de clusters (Weka);
- Áreas definidas a partir de agrupamento baseado em densidade (Weka).

Agrupamento espacial baseado em densidade



Resultado

- 13 áreas/regiões de origem ou destino;
- Cada instância da base de dados de corrida de táxis têm sua origem e seu destino identificados a uma das 13 áreas;
- **Matriz OD real** para o estudo de caso do Porto está definida.

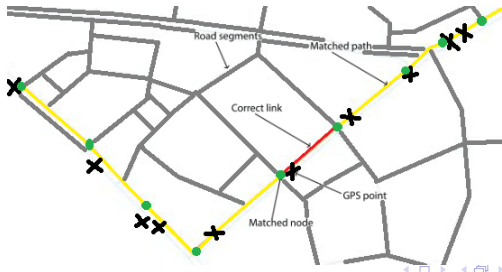
Agrupamento espacial baseado em densidade



Map Matching

Descrição do problema

- Dados georreferenciados (**trajetórias**)
- Dados topográficos (**rede viária**)
- O objetivo é **mapear** os dados georreferenciados à malha viária. (Imagem modificada de Xie et al. 2017)



Map matching para o estudo de caso de Porto

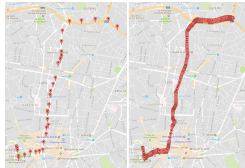
Problema

- Existe um dataset contendo dados georreferenciados, UCI (2015);
- Existe uma base de dados geográfica (rede viária, OSM);
- Os dados não estão relacionados.

Solução

- Identificar trajetórias na topologia utilizando *map matching*;
- Variações no processo de identificação dos trajetos;
- Métricas propostas para identificar a melhor solução:
Número de nós identificados e conectividade.

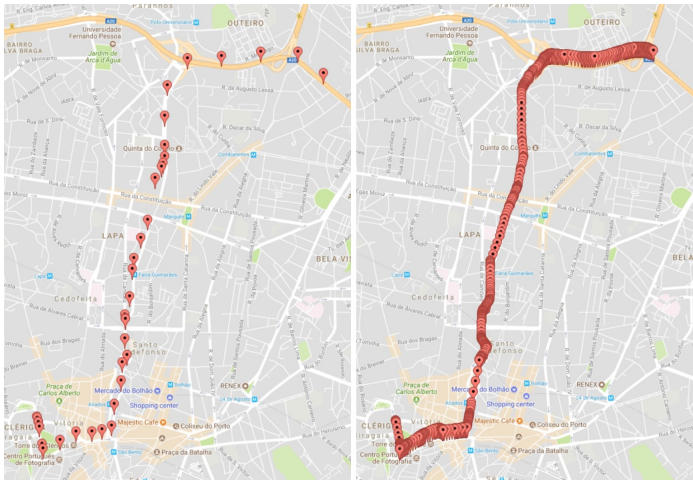
Map matching de uma corrida de táxi



Resultado

- Cada instância da base de dados de corrida de táxis têm seu trajeto na rede viária definido;
- A combinação da matriz OD real com essa nova informação permite obter uma **assignment matrix** e **fluxos** v_z ;
- Os dados de entrada para estimação da matriz OD estão disponíveis. Os experimentos de estimação da matriz OD podem ser realizados.

Map matching de uma corrida de táxi



Funções *fitness* para GA e PSO

$$\min_{q_w} \sum_{z \in Z} |v_z - \bar{v}_z| \quad (27)$$

$$\min_{q_w} \sum_{z \in Z} (v_z - \bar{v}_z)^2 \quad (28)$$

$$\min_{q_w} \sum_{z \in Z} GEH(z) \quad (29)$$

Influência da função *fitness* na ODME (GA)

Tabela 27: Resultados do GA para diferentes funções *fitness*, para o Porto

Métrica	Função objetivo	Mínimo	Média	Mediana	Máximo	Desvio padrão	p-value ANOVA
MAE	Equação 27	10,86767	13,30929	12,76822	19,53625	1,945	0,017
	Equação 28	11,14313	14,31201	13,92296	19,17522	2,086	
	Equação 29	10,34039	12,8993	12,54277	16,7681	1,723	
RMSE	Equação 27	97,79227	145,27042	131,43525	295,12815	47,618	0,416
	Equação 28	99,9748	153,97209	146,645	260,3932	44,895	
	Equação 29	89,36184	138,50392	122,89202	248,24941	42,793	
GEH	Equação 27	0,30809	0,32847	0,32479	0,38935	0,018	0,000
	Equação 28	0,34937	0,3728	0,37158	0,40176	0,011	
	Equação 29	0,28768	0,30949	0,30504	0,36069	0,018	
R^2 Arcos	Equação 27	0,98996	0,99136	0,99146	0,99228	$5,8 \cdot 10^{-4}$	0,000
	Equação 28	0,99384	0,99415	0,99415	0,99441	$1,2 \cdot 10^{-4}$	
	Equação 29	0,98878	0,99068	0,9908	0,9915	$6,1 \cdot 10^{-4}$	
R^2 ODM	Equação 27	0,43062	0,77613	0,81555	0,88717	0,113	0,485
	Equação 28	0,51011	0,75684	0,77293	0,88481	0,105	
	Equação 29	0,53249	0,79023	0,82798	0,90281	0,102	
TDD	Equação 27	0,20769	8,6391	7,84774	31,85538	6,484	0,948
	Equação 28	0,81897	9,10822	7,17805	27,46652	6,886	
	Equação 29	0,12278	8,62535	6,82644	25,82952	6,158	

Resultados

- Equação 28 superior em R^2 Arcos
- **Equação 29** superior em GEH e MAE (empate com Eq. 27)
- Empates em RMSE, R^2 ODM e TDD

Influência da função *fitness* na ODME (PSO)

Tabela 28: Resultados do PSO para diferentes funções *fitness*, para o Porto

Métrica	Função objetivo	Mínimo	Média	Mediana	Máximo	Desvio padrão	p-value ANOVA
MAE	Equação 27	9,36434	11,63505	10,81801	21,98346	2,486	0,321
	Equação 28	10,73506	12,04275	11,52271	15,26974	1,320	
	Equação 29	8,75937	11,23562	10,19799	17,25993	2,195	
RMSE	Equação 27	84,07547	151,39068	123,79863	475,95861	78,009	0,647
	Equação 28	108,79722	137,97824	116,72612	267,79211	45,614	
	Equação 29	88,36997	152,44179	117,49723	339,31932	72,466	
GEH	Equação 27	0,27226	0,27808	0,27745	0,28872	0,004	0,000
	Equação 28	0,33203	0,33877	0,33847	0,34838	0,003	
	Equação 29	0,26371	0,26762	0,26743	0,27409	0,002	
R^2 Arcos	Equação 27	0,9925	0,99301	0,99297	0,99352	$2,8 \cdot 10^{-4}$	0,000
	Equação 28	0,99467	0,99474	0,99474	0,99478	$2,2 \cdot 10^{-4}$	
	Equação 29	0,99113	0,9916	0,99155	0,99222	$2,7 \cdot 10^{-4}$	
R^2 ODM	Equação 27	0,23103	0,78012	0,84005	0,91617	0,149	0,605
	Equação 28	0,50981	0,80535	0,85714	0,86741	0,102	
	Equação 29	0,37854	0,77007	0,85442	0,91178	0,161	
TDD	Equação 27	0,12483	6,09107	4,23159	41,0173	8,025	0,595
	Equação 28	0,89732	5,05938	4,95665	14,014	3,148	
	Equação 29	0,04891	5,89024	3,41138	25,68239	6,343	

Resultados

- Equação 28 superior em R2 Arcos
- **Equação 29** superior em GEH
- Empates nas métricas relacionadas à matriz OD

Network Sensor Location Problem (NSLP)

Definir o número e localização dos sensores. A proposta é utilizar nove regras, segundo Ye e Wen (2017).

- **R1, Cobertura de todos os pares OD;**
- R2, Fração de fluxo OD em relação ao arco (q_w/v_a);
- R3, Cobertura de rotas;
- **R4, Maximização de fluxo (v_a);**
- R5, Fração de fluxo OD em relação a rotas (p_{wa});
- R6, Maximização de rotas;
- **R7, Maximização de pares OD;**
- R8, Não redundância dos arcos;
- **R9, Minimização do número de sensores.**

Influência da alocação de sensores na ODME (GA)

Tabela 29: Resultados do GA para critérios de escolha de sensores, para o Porto

Métrica	Critério de escolha	Mínimo	Média	Mediana	Máximo	Desvio padrão	p-value ANOVA
MAE	Menos rotas	30,86833	37,05226	35,24839	47,15096	4,493	0,000
	Mais rotas	12,2536	15,83797	15,8158	20,99185	2,474	
	Mais fluxo	17,96483	21,39172	20,99966	29,81795	2,533	
	Mais fluxo/rotas	13,75931	19,07086	18,55761	23,67368	2,326	
RMSE	Menos rotas	412,22141	496,39805	485,61495	584,72518	47,133	0,000
	Mais rotas	108,68122	177,44101	166,46044	292,95476	55,379	
	Mais fluxo	137,17443	201,78858	190,50677	344,72919	48,583	
	Mais fluxo/rotas	105,75229	195,70995	187,65252	310,59334	53,148	
R^2 ODM	Menos rotas	0,05869	0,47947	0,49546	0,61922	0,103	$1,6 \cdot 10^{-10}$
	Mais rotas	0,45407	0,71021	0,73401	0,86673	0,124	
	Mais fluxo	0,30913	0,61991	0,64508	0,78369	0,114	
	Mais fluxo/rotas	0,38355	0,63487	0,64399	0,8642	0,127	
TDD	Menos rotas	55,89925	81,94593	79,09295	127,95449	16,697	0,000
	Mais rotas	1,9005	13,18891	9,84184	32,95365	8,554	
	Mais fluxo	9,11579	22,97202	22,52626	45,90968	7,605	
	Mais fluxo/rotas	5,58078	18,54723	15,56353	36,59141	8,884	

Resultados

- "Mais rotas" superior em MAE, R^2 ODM e TDD
- Empate em RMSE entre três critérios de escolha

Influência da alocação de sensores na ODM (PSO)

Tabela 30: Resultados do PSO para critérios de escolha de sensores, para o Porto

Métrica	Critério de escolha	Mínimo	Média	Mediana	Máximo	Desvio padrão	p-value ANOVA
MAE	Menos rotas	36,11661	58,35329	57,41902	81,62047	11,599	0,000
	Mais rotas	11,30395	14,5663	14,05105	20,23376	2,275	
	Mais fluxo	16,98548	21,42774	21,33242	29,16685	3,103	
	Mais fluxo/rotas	14,35948	18,20084	18,18849	22,62569	2,281	
RMSE	Menos rotas	458,45051	666,43096	662,50109	829,64584	107,921	0,000
	Mais rotas	108,27874	165,61693	144,16464	347,30729	59,844	
	Mais fluxo	142,59114	237,54566	223,74847	376,60357	69,416	
	Mais fluxo/rotas	120,72471	191,41823	167,19605	315,60038	60,694	
R^2 ODM	Menos rotas	0,00766	0,12384	0,13682	0,29446	0,084	0,000
	Mais rotas	0,40654	0,75441	0,80823	0,86851	0,123	
	Mais fluxo	0,32226	0,57129	0,58536	0,78829	0,142	
	Mais fluxo/rotas	0,4088	0,66846	0,7203	0,83572	0,134	
TDD	Menos rotas	20,0167	129,81258	123,32993	226,2199	47,263	0,000
	Mais rotas	0,17986	5,6951	4,72531	28,39905	5,420	
	Mais fluxo	1,47423	11,95243	9,22528	43,7242	10,299	
	Mais fluxo/rotas	0,07921	6,29968	3,12204	20,11625	5,705	

Resultados

- "Mais rotas" superior em MAE e R^2 ODM
- Empate em RMSE e TDD: "Mais rotas" e "Mais fluxo/rotas"

Influência da alocação de sensores na ODME (MILP e LS)

Tabela 31: Resultados do MILP para critérios de escolha de sensores, Porto

	Menos rotas	Mais rotas	Mais fluxo	Mais fluxo/rotas
MAE	16,378	9,16	10,86	10,773
RMSE	376,61	99,003	97,506	107,435
R^2 ODM	0,738	0,894	0,892	0,875
TDD	0,691	2,839	5,459	2,635

Tabela 32: Resultados do LS para critérios de escolha de sensores, Porto

	Menos rotas	Mais rotas	Mais fluxo	Mais fluxo/rotas
MAE	13,671	16,259	24,589	19,651
RMSE	238,184	164,606	327,058	191,925
R^2 ODM	0,809	0,753	0,437	0,694
TDD	8,731	2,096	4,899	2,616

Resultados

- "Mais rotas" superior em MAE e R^2 ODM para MILP
- "Mais rotas" superior para RMSE e TDD para LS

Comparação das técnicas de ODME

Tabela 33: Comparação das técnicas de estimação da matriz OD para Porto

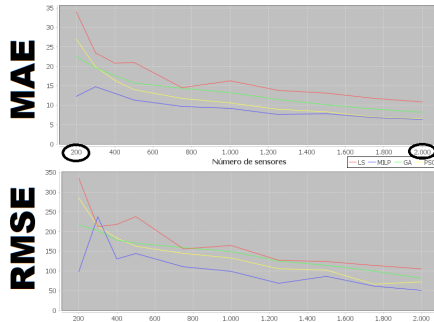
Métrica	Técnica	Mínimo	Média	Mediana	Máximo	Desvio padrão	p-value ANOVA
MAE	PSO	8,75937	11,23562	10,19799	17,25993	2,195	0,000
	GA	10,34039	12,8993	12,54277	16,7681	1,723	
	LS	16,25877	16,25877	16,25877	16,25877	0	
	MILP	9,15974	9,15974	9,15974	9,15974	0	
RMSE	PSO	88,36997	152,44179	117,49723	339,31932	72,466	$9,7 \cdot 10^{-8}$
	GA	89,36184	138,50392	122,89202	248,24941	42,793	
	LS	164,60557	164,60557	164,60557	164,60557	0	
	MILP	99,00264	99,00264	99,00264	99,00264	0	
GEH	PSO	0,26371	0,26762	0,26743	0,27409	0,002	0,000
	GA	0,28768	0,30949	0,30504	0,36069	0,018	
	LS	0,97954	0,97954	0,97954	0,97954	0	
	MILP	0,27001	0,27001	0,27001	0,27001	0	
R^2 Arcos	PSO	0,99113	0,9916	0,99155	0,99222	$2,7 \cdot 10^{-4}$	0,000
	GA	0,98878	0,99068	0,9908	0,9915	$6,1 \cdot 10^{-4}$	
	LS	0,95856	0,95856	0,95856	0,95856	0	
	MILP	0,99328	0,99328	0,99328	0,99328	0	
R^2 ODM	PSO	0,37854	0,77007	0,85442	0,91178	0,161	$2,0 \cdot 10^{-7}$
	GA	0,53249	0,79023	0,82798	0,90281	0,102	
	LS	0,75296	0,75296	0,75296	0,75296	0	
	MILP	0,89398	0,89398	0,89398	0,89398	0	
TDD	PSO	0,04891	5,89024	3,41138	25,68239	6,343	$9,3 \cdot 10^{-8}$
	GA	0,12278	8,62535	6,82644	25,82952	6,158	
	LS	2,09616	2,09616	2,09616	2,09616	0	
	MILP	2,83939	2,83939	2,83939	2,83939	0	

Comparação das técnicas de ODME

Resultados

- MILP superior em MAE, RMSE, R2 ODM e R2 Arcos
 - PSO superior em GEH
 - LS superior em TDD
-
- MILP superior em 3 de 4 métricas para ODM
 - Empate 1x1 entre PSO e MILP para métricas de arcos
 - Entre as 6 métricas, MILP é superior em 4.

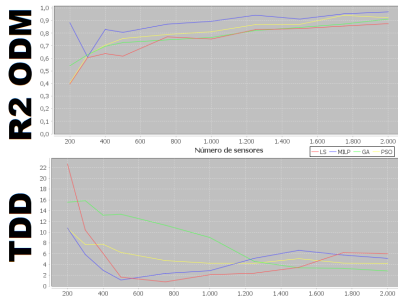
Influência do número de sensores na ODME



Resultados

- MAE e RMSE menores para maior número de sensores
- PSO mais próximo de MILP com mais sensores

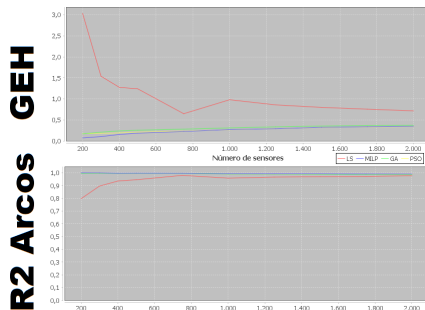
Influência do número de sensores na ODME



Resultados

- R2 ODM com comportamento similar a RMSE
- TDD com ponto mínimo para MILP e LS em 750 sensores
- PSO e GA com menor TDD para mais sensores

Influência do número de sensores na ODME



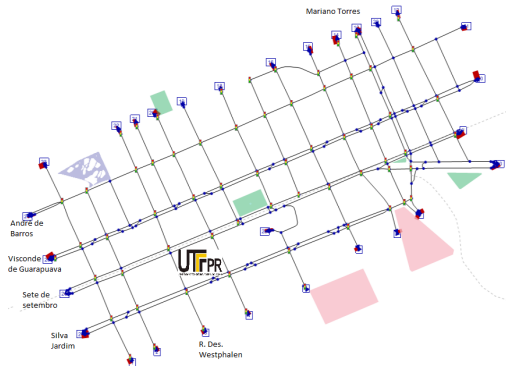
Resultados

- GEH menor para menos sensores para PSO, GA e MILP
- R2 Arcos com valores muito próximos para PSO, GA e MILP

Estudo de caso de Curitiba

- Definição da *Assignment Matrix*;
- Configuração dos algoritmos GA e PSO;
- Influência da função *fitness* na ODME;
- Comparação das técnicas de OMDE.

Descrição do estudo de caso



- Problema pré-definido pelo IPPUC usando o software PTV VISUM 17;
- Rede fechada com 191 nós e 594 arcos;
- 29 nós de origem ou destino nas fronteiras. 183 pares OD válidos;
- 71 sensores nos arcos.

Definição da *Assignment Matrix* (VISUM)

Problema: Forma como veículos transitam na rede é desconhecida

- Qual o **conjunto de rotas** para cada par OD?
- Qual a **probabilidade de uso** de arco por par OD?

Solução através do equilíbrio de usuário

- Motoristas buscam **minimizar a duração** de suas viagens;
- Motoristas têm **conhecimento pleno** do ambiente;
- Equilíbrio de tempo de viagem entre todas as rotas **de um mesmo par OD**.

Saída: Probabilidades de uso de arco por par OD

- (Traffic) Assignment matrix: $ns \times np$

T-Flow Fuzzy (VISUM)

- T-Flow Fuzzy é uma ferramenta integrada ao VISUM
- Uma sequência de procedimentos no VISUM é utilizada para obtenção da matriz OD

Algorithm 3 Pseudo-código da aplicação do *T-Flow Fuzzy*

- 1: Matriz OD resposta é inicialmente a matriz OD semente
 - 2: Procedimento de *equilíbrio de usuário* para definir a *assignment matrix*
 - 3: $t=1$
 - 4: **while** $t \leq 5$ **do**
 - 5: Atualização da matriz OD resposta através do *T-Flow Fuzzy*
 - 6: Procedimento de *equilíbrio de usuário* para redefinir a *assignment matrix*
 - 7: Reduz os desvios máximos aceitos nos sensores
 - 8: $t = t + 1$
 - 9: **end while**
-

Configuração dos algoritmos GA e PSO (iRace)

Tabela 34: Seleção de parâmetros do PSO, para Curitiba

Parâmetro	Intervalo avaliado	Valor selecionado
s	[20;2500]	1404
t_{psa}	3000000/ s	2136
$c1$	[0,01;1,0]	0,44
$c2$	[0,01;1,0]	0,89
w_i	[0,01;1,0]	0,90
w_f	[0,01;1,0]	0,07
$reset_t$	[20,500]	71

Tabela 35: Seleção de parâmetros do GA, para Curitiba

Parâmetro	Intervalo avaliado	Valor selecionado
pop	[20,1500]	50
t_{ga}	3000000/ pop	40000
P_c	[0,01;1,0]	0,85
P_m	[0,01;1,0]	0,34
vMut	[0,01;1,0]	0,10

- 1000 experimentos para cada seleção de parâmetros. Limite de 3.000.000 avaliações de *fitness* por experimento;
- Cada seleção de parâmetros levou cerca de 12 horas para cada algoritmo, utilizando 8 *threads* de um processador AMD Ryzen 1700 3 Ghz (lançado em 2017).

Funções *fitness* para GA e PSO

$$\min_{q_w} \sum_{z \in Z} |v_z - \bar{v}_z| \quad (27)$$

$$\min_{q_w} \sum_{z \in Z} (v_z - \bar{v}_z)^2 \quad (28)$$

$$\min_{q_w} \sum_{z \in Z} GEH(z) \quad (29)$$

Influência de função *fitness*

Tabela 36: Resultados do GA para diferentes funções *fitness*, para Curitiba

Métrica	Função objetivo	Mínimo	Média	Mediana	Máximo	Desvio padrão	p-value ANOVA
GEH	Equação 27	3,28577	3,90071	3,84033	5,29251	0,377	0,000
	Equação 28	2,31036	2,40437	2,39622	2,4810	0,051	
	Equação 29	3,03151	3,61156	3,52315	4,89015	0,417	
pGEH5	Equação 27	0,71831	0,77418	0,77465	0,83099	0,025	0,000
	Equação 28	0,85915	0,88357	0,88732	0,90141	0,011	
	Equação 29	0,70423	0,76714	0,77465	0,8169	0,033	
R^2 Arcos	Equação 27	0,82842	0,88432	0,88854	0,91597	0,021	0,000
	Equação 28	0,9572	0,95914	0,95947	0,96017	$8,0 \cdot 10^{-4}$	
	Equação 29	0,82246	0,88262	0,88887	0,9090	0,21	

Tabela 37: Resultados do PSO para diferentes funções *fitness*, para Curitiba

Métrica	Função objetivo	Mínimo	Média	Mediana	Máximo	Desvio padrão	p-value ANOVA
GEH	Equação 27	2,11188	2,69738	2,6456	3,39812	0,354	0,000
	Equação 28	2,0017	2,02524	2,01876	2,09212	0,020	
	Equação 29	2,10559	2,4113	2,40728	2,80228	0,187	
pGEH5	Equação 27	0,80282	0,84319	0,84507	0,88732	0,023	0,000
	Equação 28	0,88732	0,89202	0,88732	0,90141	0,006	
	Equação 29	0,80282	0,84178	0,84507	0,90141	0,024	
R^2 Arcos	Equação 27	0,88309	0,92709	0,92924	0,95214	0,017	0,000
	Equação 28	0,96283	0,96339	0,96349	0,96362	$2,2 \cdot 10^{-4}$	
	Equação 29	0,88017	0,92435	0,9282	0,94543	0,013	

Influência de função *fitness*

Resultados

- Mesma função *fitness* escolhida (**Equação 28**), independente de métrica (GEH, pGEH5 ou R^2 Arcos) ou algoritmo (GA ou PSO);
- Função *fitness* (**Equação 28**) que foi a segunda opção para o estudo de caso do Porto.

Comparação das técnicas de ODME

Tabela 38: Comparação das técnicas de estimação da matriz OD para Curitiba

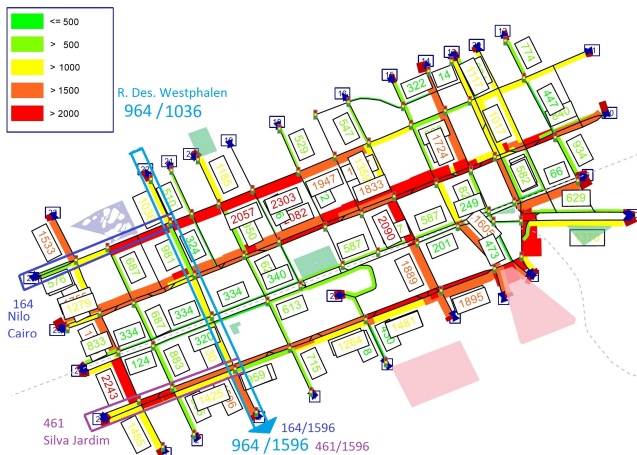
Métrica	Técnica	Mínimo	Média	Mediana	Máximo	Desvio padrão	p-value ANOVA
GEH	PSO	2,0017	2,02524	2,01876	2,09212	0,020	0,000
	GA	2,31036	2,40437	2,39622	2,481	0,051	
	LS	7,14832	7,14832	7,14832	7,14832	0	
	MILP	1,83757	1,83757	1,83757	1,83757	0	
	VISUM	1,861	1,861	1,861	1,861	0	
pGEH5	PSO	0,88732	0,89202	0,88732	0,90141	0,006	0,000
	GA	0,85915	0,88357	0,88732	0,90141	0,011	
	LS	0,42254	0,42254	0,42254	0,42254	0	
	MILP	0,87324	0,87324	0,87324	0,87324	0	
	VISUM	0,73239	0,73239	0,73239	0,73239	0	
R^2 Arcos	PSO	0,96283	0,96339	0,96349	0,96362	$2,2 \cdot 10^{-4}$	0,000
	GA	0,9572	0,95914	0,95947	0,96017	$8,0 \cdot 10^{-5}$	
	LS	0,79595	0,79595	0,79595	0,79595	0	
	MILP	0,95345	0,95345	0,95345	0,95345	0	
	VISUM	0,91047	0,91047	0,91047	0,91047	0	

Comparação das técnicas de ODME

Resultados

- MILP superior na métrica GEH
 - **PSO** superior nas métricas R^2 Arcos e pGEH5
-
- PSO superior em 2 de 3 métricas, para arcos
 - Não é possível aplicar métricas ODM
 - Somente PSO, GA e MILP obtiveram bons valores pGEH5

Fluxo de veículos na rede viária (PSO)



Conclusão

Conclusão

- **MILP** com melhor desempenho para **Porto**, nas métricas para matriz OD. Empate para métricas de arcos com PSO.
- **PSO** com melhor desempenho para **Curitiba**;
- Função *fitness* afeta o desempenho do GA e PSO, mas está mais relacionada às métricas e ao problema do que à decisão entre GA e PSO;
- Para Porto, mesmo **critério de escolha de sensores** independente da técnica utilizada. **Número de sensores** afeta as técnicas em intensidades diferentes;
- Necessidade de se utilizar **várias métricas** para compreender a estimação da matriz OD. As métricas são complementares.

Conclusão

Trabalhos futuros

- **Inclusão das técnicas** de inferência bayesiana e máxima verossimilhança na comparação;
- Uso de dados de **veículos sonda** e *map matching* para obter uma **matriz OD semente**, bem como **probabilidades de uso de arco** por par OD, para estimação da matriz OD em Curitiba;
- Uso de dados de **veículos sonda** e *map matching* como um auxílio para o **NSLP** em redes viárias reais, como de Curitiba.

Agradecimentos

- Um agradecimento ao Engenheiro Francisco Caron Malluceli, que disponibilizou dados para o estudo de caso de Curitiba, representando o IPPUC.
- Um agradecimento ao projeto Vinnova/KTH, graças ao qual ocorreu o Summer School em Estocolmo/Suécia.

Obrigado!

Contato:

luciano.pando@ifpr.edu.br

luders@utfpr.edu.br