



Generalitat de Catalunya
Institut d'Estadística de Catalunya

Creació d'una infraestructura de càlcul de rutes per a l'anàlisi espacial de resultats estadístics

Un assaig per a l'estudi de la distribució dels desplaçaments obligats a Catalunya

Gabriel Garcies Ramon
Treball final de màster
Barcelona, juliol de 2016

Supervisió a càrrec de:
Dr Joan Nunes
Eduard Suñé

Sumari

- I. Introducció
- II. Creació del graf
- III. Càlcul de rutes
- IV. Anàlisi dels resultats
- V. Conclusions

I. Introducció

- Objectius:
 - Produir un graf que representi la xarxa de comunicacions real amb diferents mitjans de transport per a l'àrea d'estudi.
 - Establir un sistema de càlcul de rutes òptimes per a dades massives.
 - Analitzar els resultats estadístics obtinguts.

- **Àrea d'estudi: Àrea Metropolitana de Barcelona**
 - Entitat supramunicipal (36 municipis) de 636 km².
 - Més de 3 milions d'habitants.
 - Més de 10 milions de desplaçaments a l'any.
 - 3,9 desplaçaments per persona i dia (2014).

Mapa de localització de l'Àrea Metropolitana de Barcelona



- **Software i font de les dades:**

1. DADES SOBRE DESPLAÇAMENTS

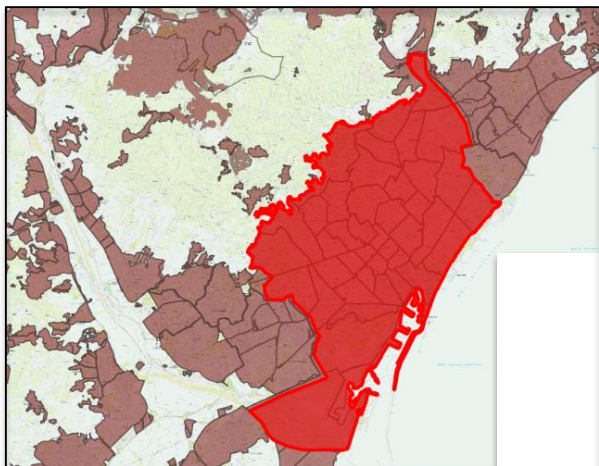


2. GRAF



3. CARTOGRAFIA DE BASE

Àrees urbanes (SIGPAC) + Codis postals (IGN)



4. SOFTWARE



II. Creació del graf

- **Modelització del transport a peu:**

1. Operacions i tractament de les dades

Assignar una tipologia de via.

Excloure segments invàlids/excepcions.

Càlculs necessaris per la funció de *Routing*:

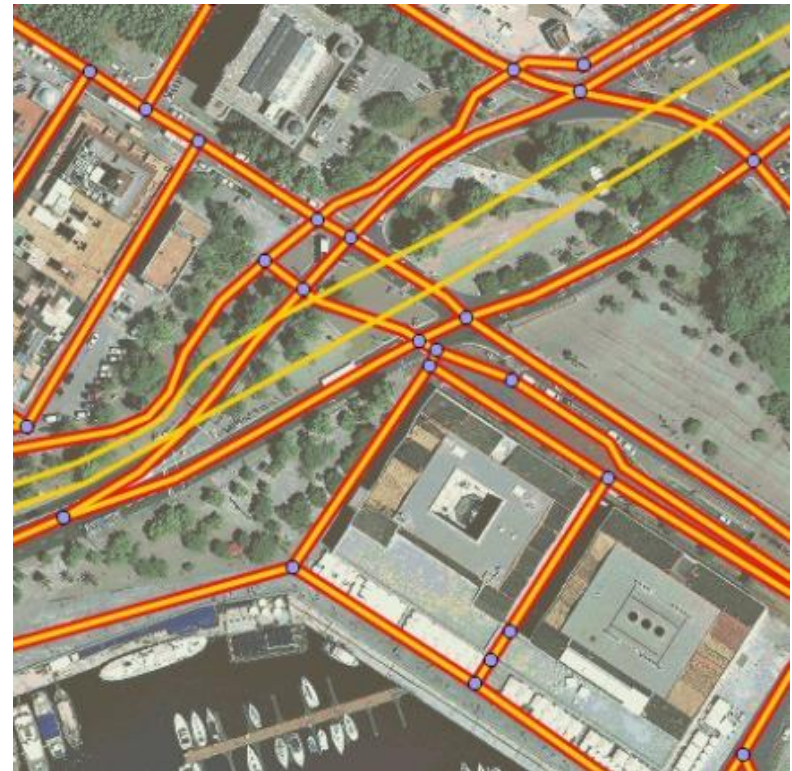
- Coordenades geogràfiques.
- Longitud dels segments.
- Velocitat mitjana segons el mode de transport.

2. Segmentació de la xarxa.

Pgr_nodenetwork

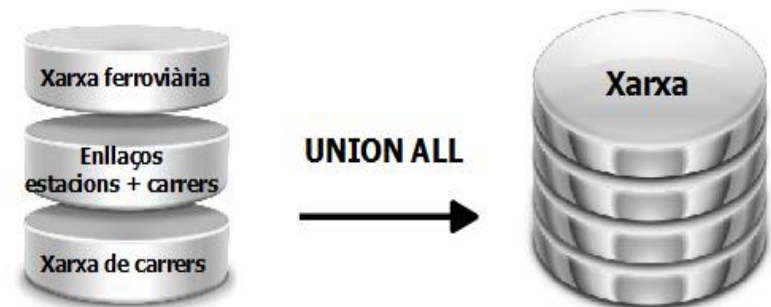
3. Creació de la topologia.

Pgr_createTopology



• Modelització del transport ferroviari:

1. Creació d'enllaços entre les línies de Metro, FGC, ADIF i Tram.
2. Creació d'enllaços entre les estacions i la xarxa de carrers.
3. Segmentació de la xarxa.
4. Creació de la topologia.

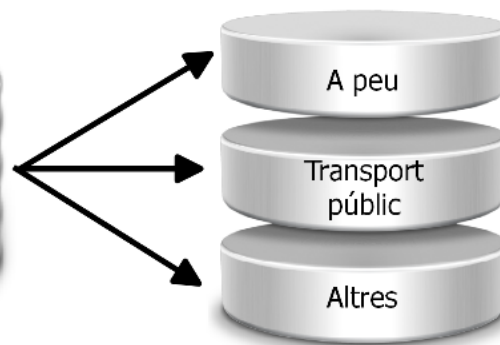


III. Càlcul de rutes

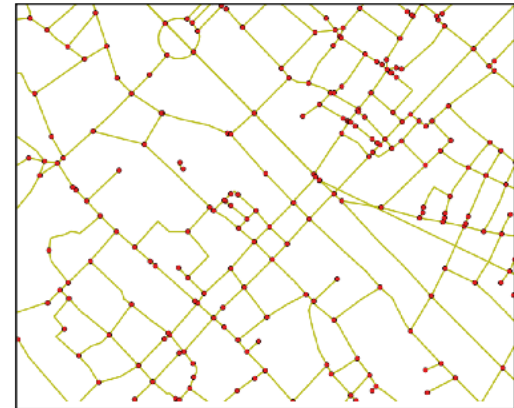
Dades sobre desplaçaments obligats a Catalunya
Font: IDESCAT



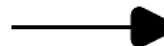
Segregació de les dades segon mitjà de transport



Graf viari amb estructura topològica

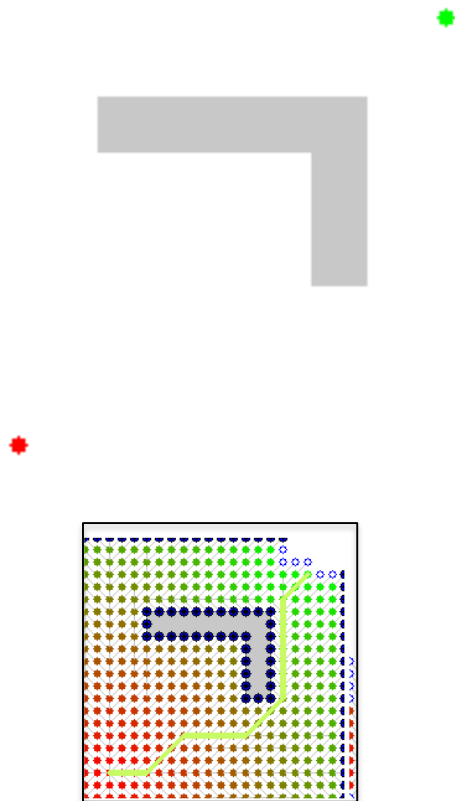


Càlcul de rutes amb:

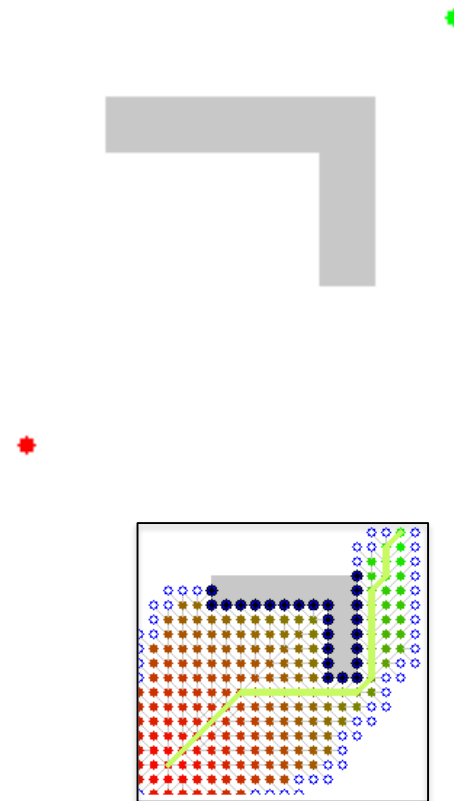


- **Funcions de *Routing*:**

Pgr_Dijkstra (1959)



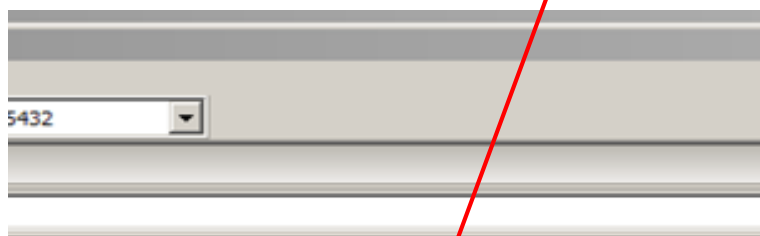
Pgr_Astar (1968)



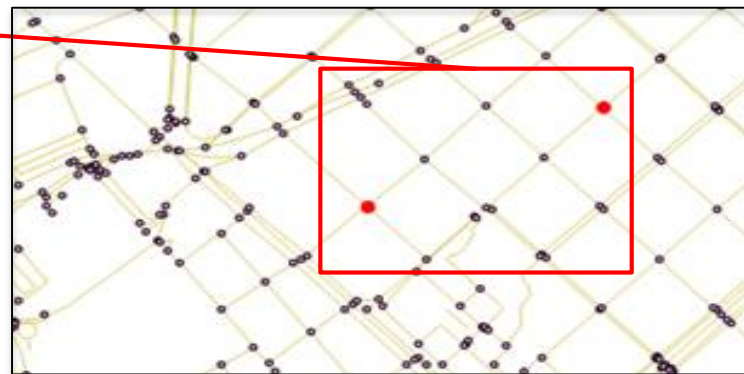
Font: GitHub.com

- **Funcionament inicial d'Astar:**

```
select * from pgr_AStar  
( 'SELECT id, CAST(source AS int4), cast(target as int4),  
cast(cost as float8), x1, y1, x2, y2 FROM  
routing_network', 62539, 62740, false, false)
```



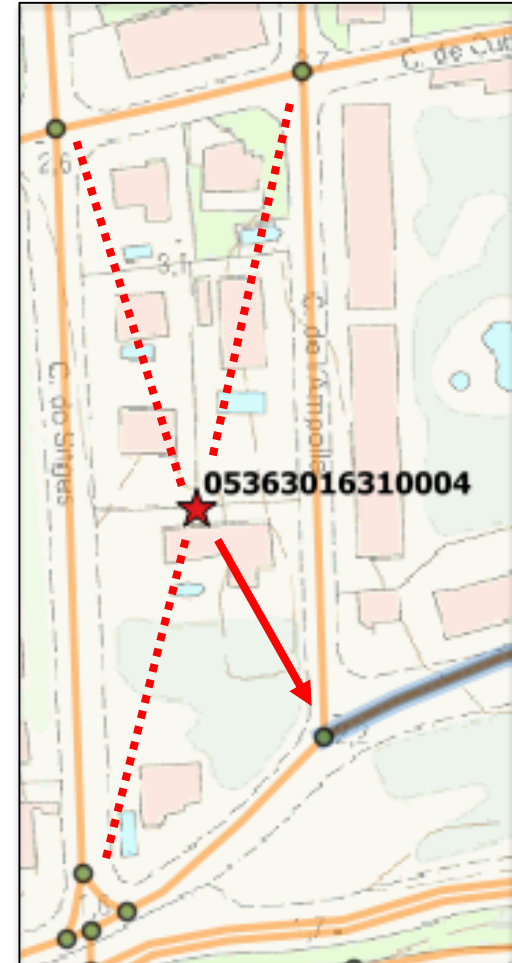
```
? FROM routing_network', 62539, 62740, false, false)
```



- **Automatització del càlcul de rutes:**

Funció Pgr_Astar_vertex_mes_proper:

```
select *  
pgr_astar('SELECT id,  
CAST (source AS int4),  
CAST (target as int4),  
CAST (longitud/velocitat as float8) as cost,  
x1, y1, x2, y2 FROM xarxa where  
tipologia=1',  
vertex_mes_proper(geom) ,  
vertex_mes_proper(desti), false, false)  
from orig_dest_utils where  
idpers='05363016310004'
```



- **Resultat d'executar la funció *Astar*:**

Registres que retorna la funció:

	idpers text	pgr_astar pgr_costresult
1396	04021012040004	(15, 8098, 97921, 5.12761442007374)
1397	04021012040004	(16, 10332, 117043, 2.5261965579403)
1398	04021012040004	(17, 32504, 102109, 3.10368963836257)
1399	04021012040004	(18, 14111, 53181, 1.57588048180671)
1400	04021012040004	(19, 66673, 53180, 4.684446428975)
1401	04021012040004	(20, 66672, 53179, 0.247542092130638)
1402	04021012040004	(21, 26937, 38276, 9.54329616686832)
1403	04021012040004	(22, 8099, 53182, 6.2676406278112)
1404	04021012040004	(23, 26976, 58880, 5.77326697892641)
1405	04021012040004	(24, 8123, 4923, 13.1467300556498)
1406	04021012040004	(25, 8124, 4926, 2.10468879608372)
1407	04021012040004	(26, 8125, 4924, 1.36872096174351)
1408	04021012040004	(27, 8126, 77237, 25.4668441930015)
1409	04021012040004	(28, 28972, 17578, 17.7132511843602)
1410	04021012040004	(29, 14146, 92106, 26.4536596778386)
1411	04021012040004	(30, 10353, 12836, 17.4736925315081)
1412	04021012040004	(31, 21476, -1, 0)
1413	04023000990003	(0, 32483, 79475, 2.79649697714169)
1414	04023000990003	(1, 32484, 19970, 4.59761010327461)
1415	04023000990003	(2, 32485, 109260, 20.292813016054)
1416	04023000990003	(3, 78411, 73742, 2.00982368769803)
1417	04023000990003	(4, 64069, 64677, 1.3279771446485)
1418	04023000990003	(5, 55792, 39870, 2.77252495641437)
1419	04023000990003	(6, 28933, 17551, 19.7063204350738)
1420	04023000990003	(7, 28934, 43872, 1.52665599599018)
1421	04023000990003	(8, 10258, 6089, 2.98730219736639)
1422	04023000990003	(9, 10259, 98115, 2.37223478161325)
1423	04023000990003	(10, 55366, 39455, 1.78080957029848)
1424	04023000990003	(11, 55365, 94409, 4.50309670215946)
1425	04023000990003	(12, 73891, 99031, 12.6349738442095)
1426	04023000990003	(13, 74623, 65914, 3.27030690449895)
1427	04023000990003	(14, 67521, 54341, 0.998876151217889)

Temps de desplaçament

ID vèrtexs

Resultats agrupats per individu:

	idpers text	temps double precision	distancia numeric
1	03775000830001	350.649353575581	385.7142889331391132
2	03775002200003	591.896805775868	651.0864863534553132
3	03775002690001	596.050534857287	655.6555883430154532
4	03775002690002	596.050534857287	655.6555883430154532
5	03775002690003	5902.60201185809	35381.51008132646182
6	03775003360003	475.931113743654	523.5242251180197132
7	03775003870001	472.721064078395	519.9931704862343932
8	03775003870002	155.686577989891	519.9931704862343932
9	03775003870003	472.721064078395	519.9931704862343932
10	03775005150001	850.730677686386	935.8037454550241132
11	03775005500002	522.491759985332	574.7409359838648932
12	03775006210003	628.519752704209	691.3717279746295132

```
create table
desplacaments_persona as
SELECT
idpers,
sum((pgr_astar).cost) as
temps,
sum(xarxa.longitud) as
distancia
FROM desplaçaments
JOIN xarxa ON (pgr_astar).id2
= xarxa.id group by idpers
```

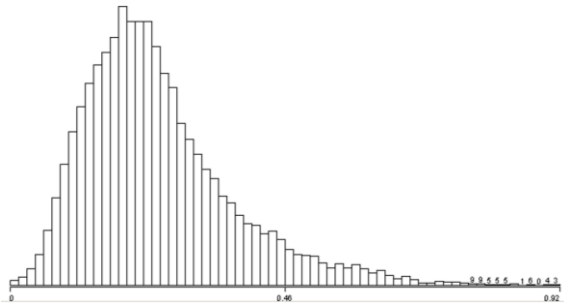
IV. Anàlisi dels resultats

- **Metodologia d'anàlisi:**

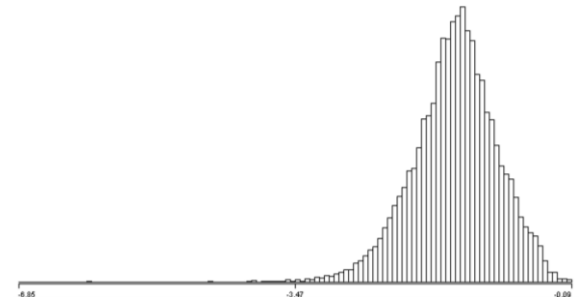
$$\text{Càlcul de l'error relatiu: } \varepsilon = \frac{|d_e - d'|}{d'}$$

d_e distància euclidiana entre els punts d'origen i destinació

d' distància *Astar* corregida entre els punts d'origen i destinació



Distribució de l'error relatiu (ε)

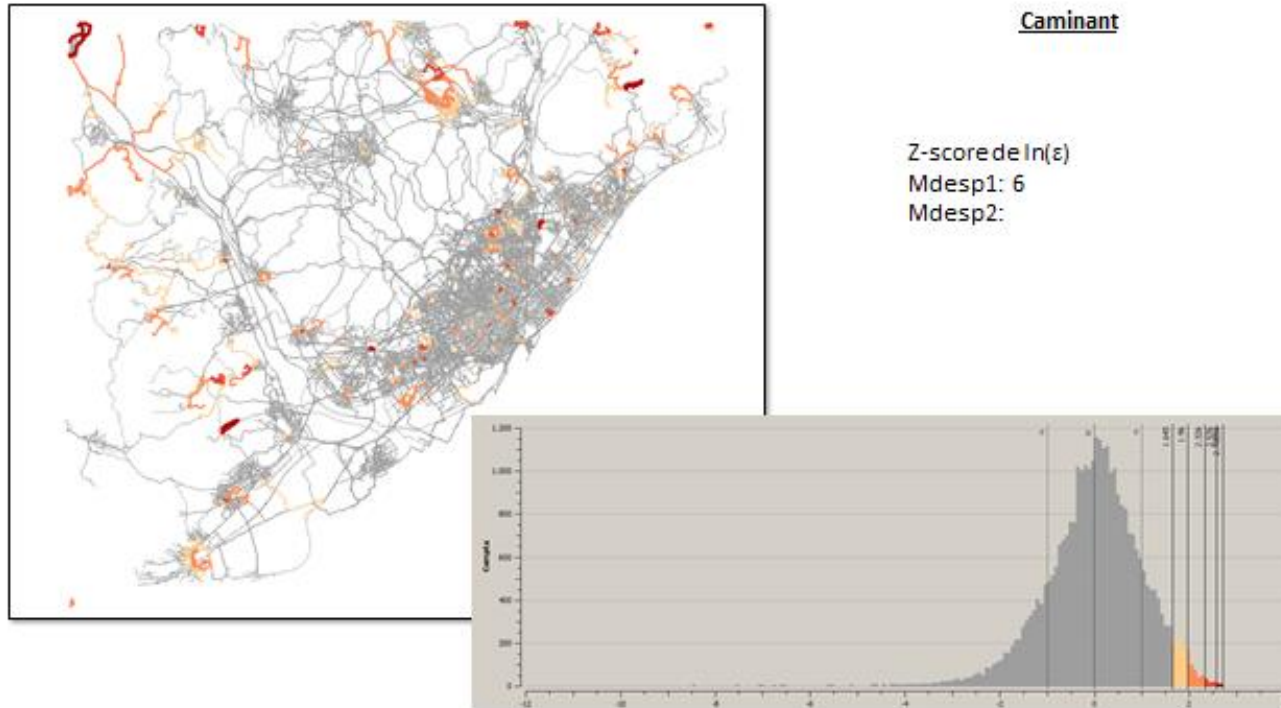


Distribució del logaritme de l'error relatiu $\ln(\varepsilon)$

- La distribució de ε és molt semblant a una **log** normal.
- La correlació entre d_e i d' és molt forta (coeficient de correlació > 0.9).
- S'ha calculat el $\ln(\varepsilon)$ i la seva distribució és molt semblant a una normal.

• Diagnosi de la xarxa:

- S'ha calculat el **Z-score** del $\ln(\epsilon)$ i s'han localitzat els **outliers** de la distribució.
- Permet la identificació d'errors en la construcció del graf.
- Permet avaluar l'eficiència de la xarxa de transport existent.



V. Conclusions

- El graf ha estat creat i testejat satisfactòriament.
- El càlcul de rutes massiu i l'automatització del procés són l'element diferencial del treball.
- La metodologia estadística per a l'anàlisi de l'eficiència del graf és un plus qualitatiu destacat.
- Durant el desenvolupament del treball han aparegut problemes i incògnites que s'han sabut resoldre.
- Es demostra el potencial del software lliure i es remarca la importància de la qualitat de les dades.

Gràcies per la seva atenció