U.PORTO

LIACC

artificial intelligence and computer science laboratory

Mining Spatial Data from GPS Traces for Automated Map Generation

Fernanda Lins & Michel Ferreira

Faculdade de Ciências da Universidade do Porto
Departamento de Ciência de Computadores

OUTLINE

- **Introduction**
- **Data Set**
- **Automated Map Generation**
- **Road Classification**
- **Conclusions and Future Work**

INTRODUCTION

The main purpose of our work is to automated generate highly detailed and accurate vectorial road maps from GPS traces.

MOTIVATION:

- GPS receivers have become ubiquitous devices. They are present in PDAs, cell phones, jogging watches or even key holders.

- With the constant bandwidth increase in wireless networks, tracking companies are now transmitting position reports "in-raw", producing a huge amount of spatial information.

- Map generation, refinement and update are very expensive if done in the conventional way.

INTRODUCTION

Several projects have been developed in order to take advantage of spatial data produced by GPS devices:

- Wikimapia → Points of Interest
- OpenStreetMap → GPS Traces

Data aggregation is done in a manual way!

INTRODUCTION

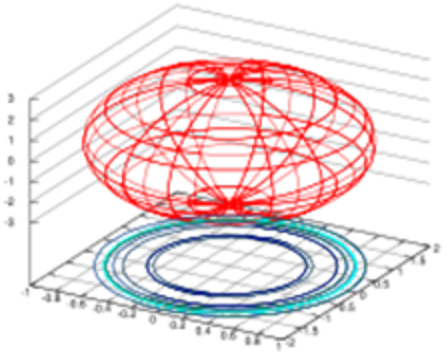
In order to accurately represent road information, a great number of GPS trace is required.

How to collect data?

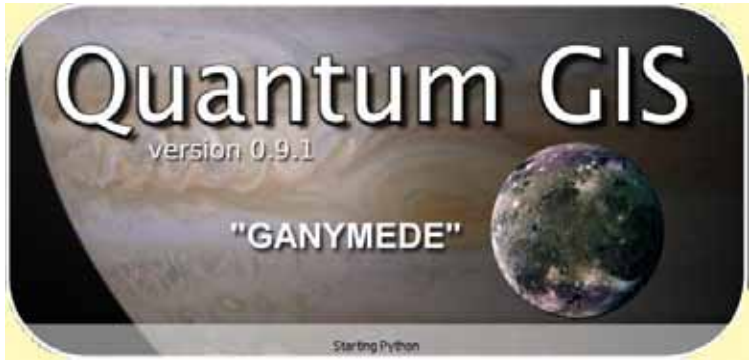
SOFTWARES

PostGIS

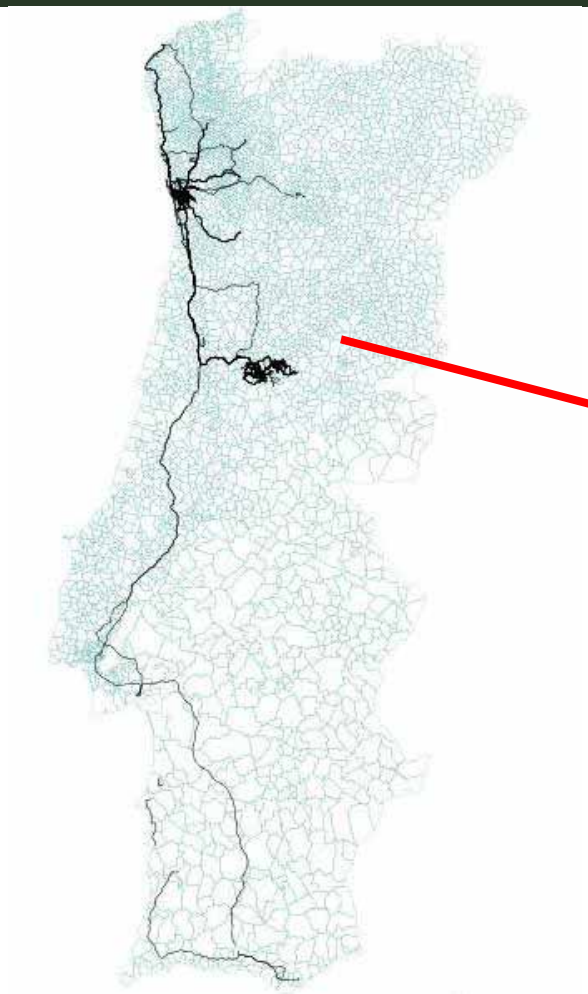
python™



GnuPlot

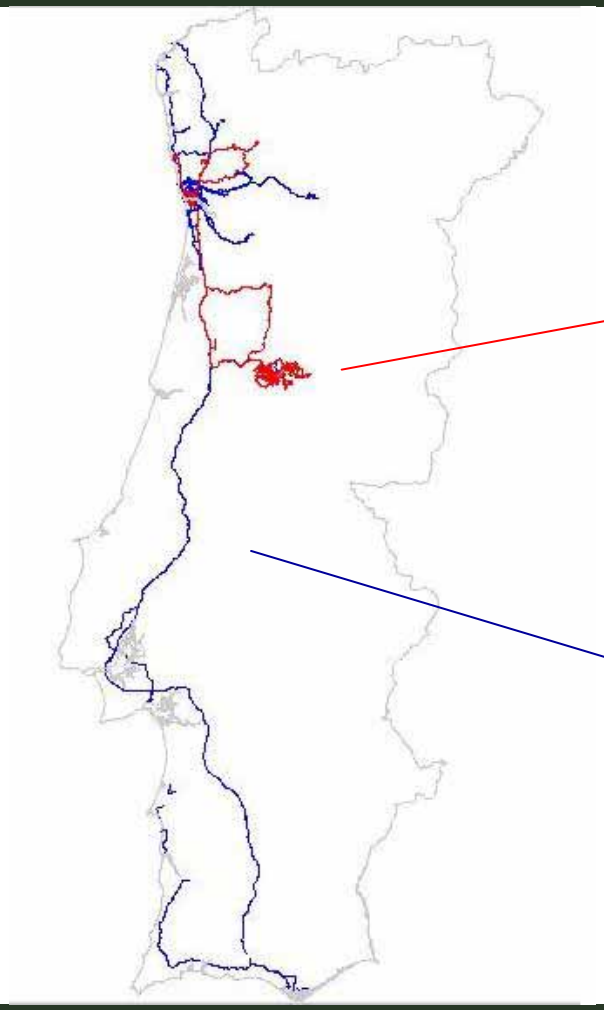


DATA SET



24.946 km

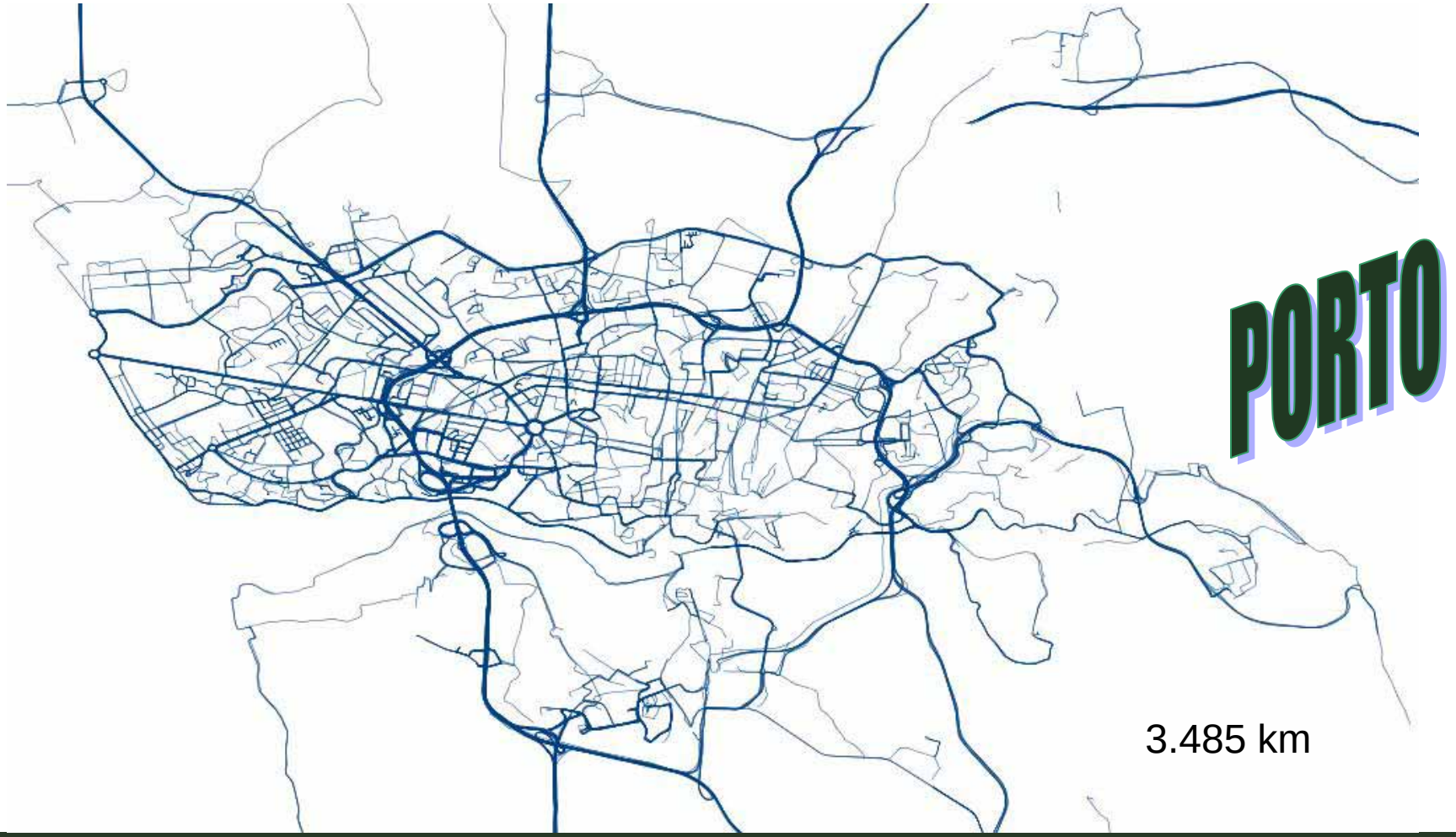
DATA SET



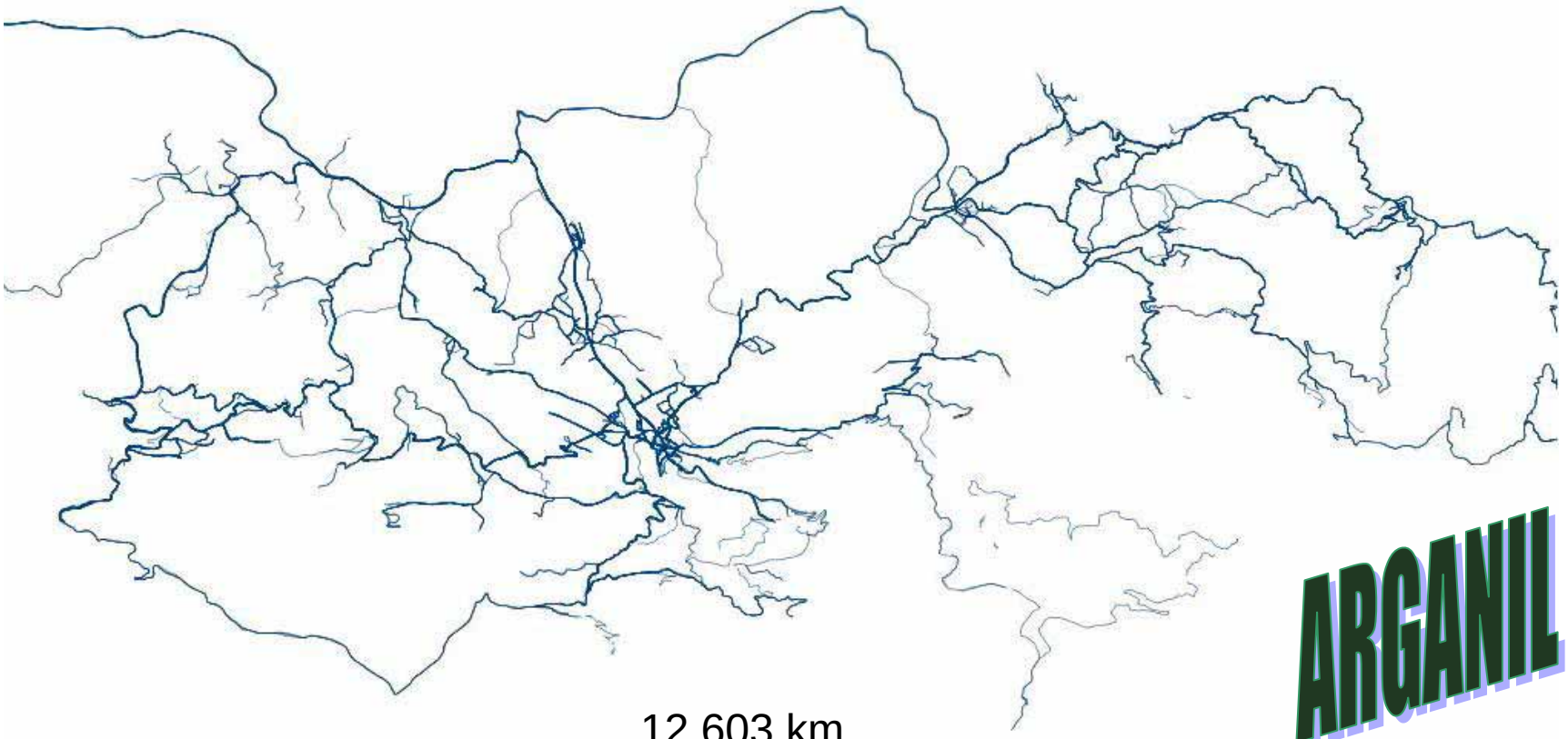
Compact real-time logs

Offline GPS logs

DATA SET



DATA SET



12.603 km

ARGANIL

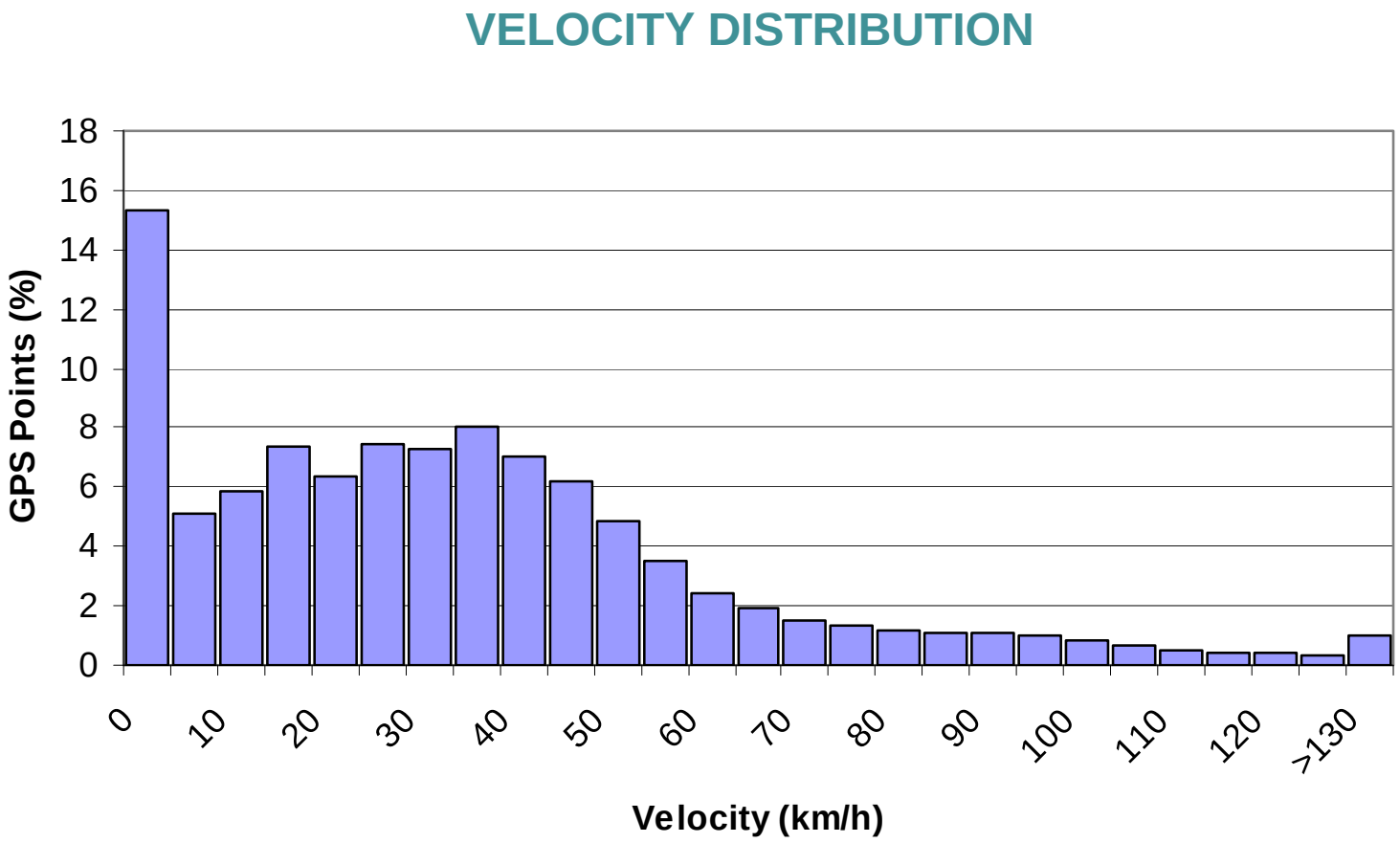
PRE-PROCESSING

In order to guarantee the integrity of our data set we first have to extract and filter the relevant data from GPS logs.

Our filters are based on:

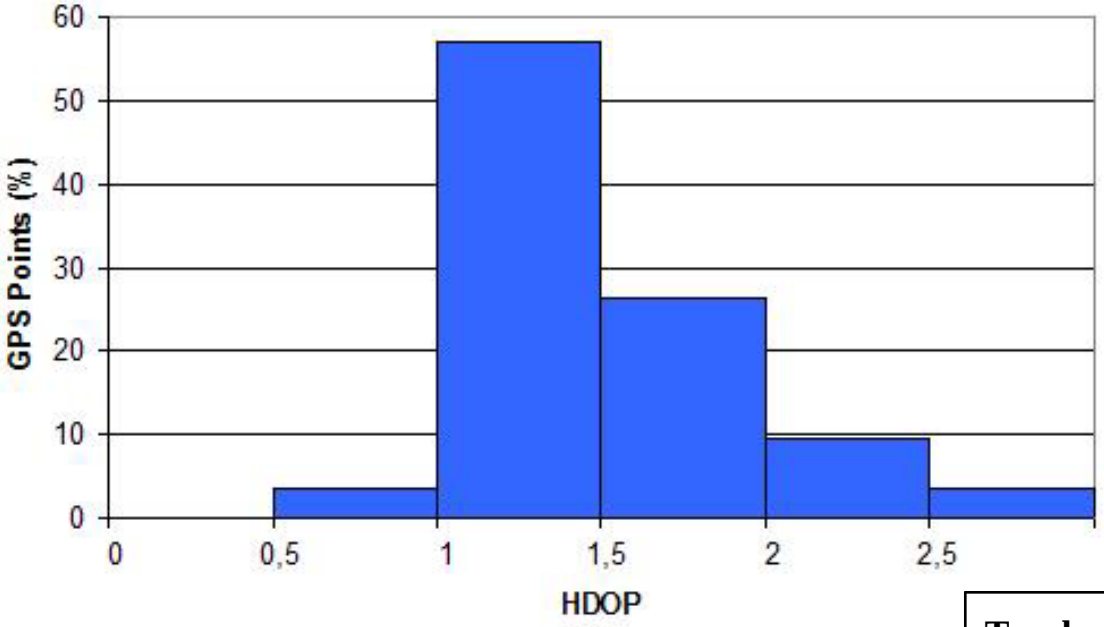
- Velocity;
- HDOP value;
- Number of tracked satellites.

PRE-PROCESSING



PRE-PROCESSING

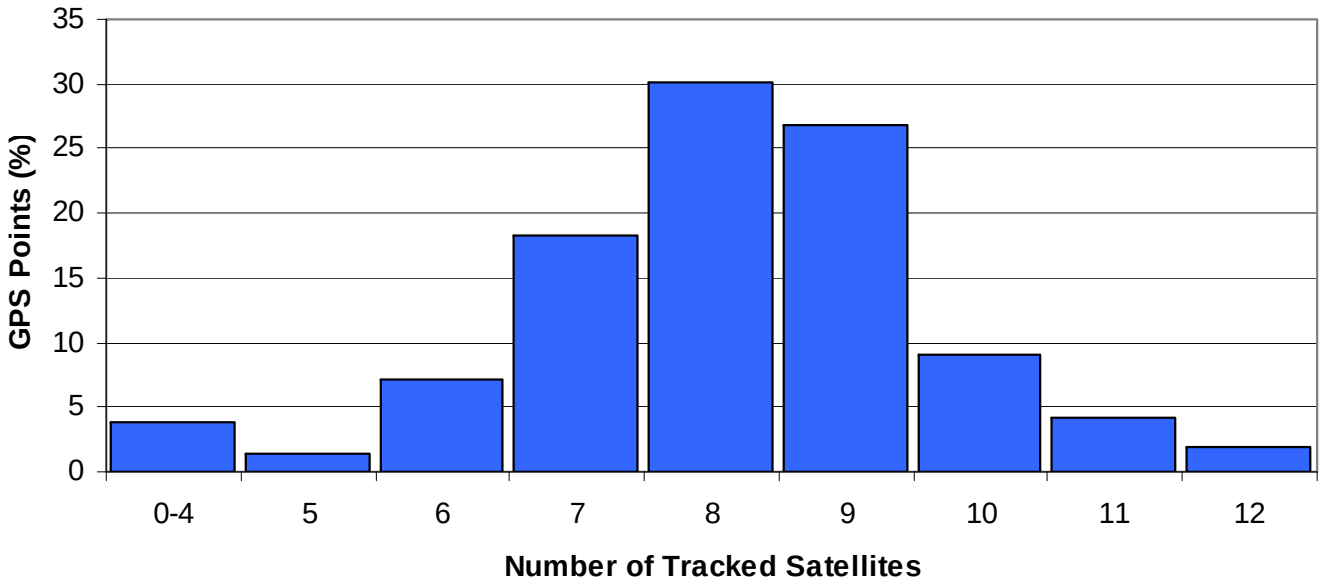
HDOP DISTRIBUTION



Tracked Area	HDOP Average	Discarded Points
Arganil	1.504	11,34 %
Porto	1.663	22,48 %

PRE-PROCESSING

TRACKED SATELLITES DISTRIBUTION



Tracked Area	Tracked Satellites Average	Discarded Points
Arganil	8	3,92 %
Porto	7	4.18 %

PRE-PROCESSING

Line Simplification

Douglas Peucker Algorithm

	GPS Traces	Simplified Traces	Eliminated Points
Total Number of Points	2.324.624	255.709	89%

	GPS Traces	Simplified Traces	Elimination error
Tracked Km	24.946	24.809	0,55%

AUTOMATED MAP GENERATION

The automated map generation is NOT a simple task:

Traffic changes could be dangerous;

GPS receivers are NOT precise;

A great number of GPS Traces is required;

There is no associated metadata.

AUTOMATED MAP GENERATION

ALGORITHM

The main purpose of our map generation algorithm is to construct a directed graph where edges represent road segments and nodes represent junctions.

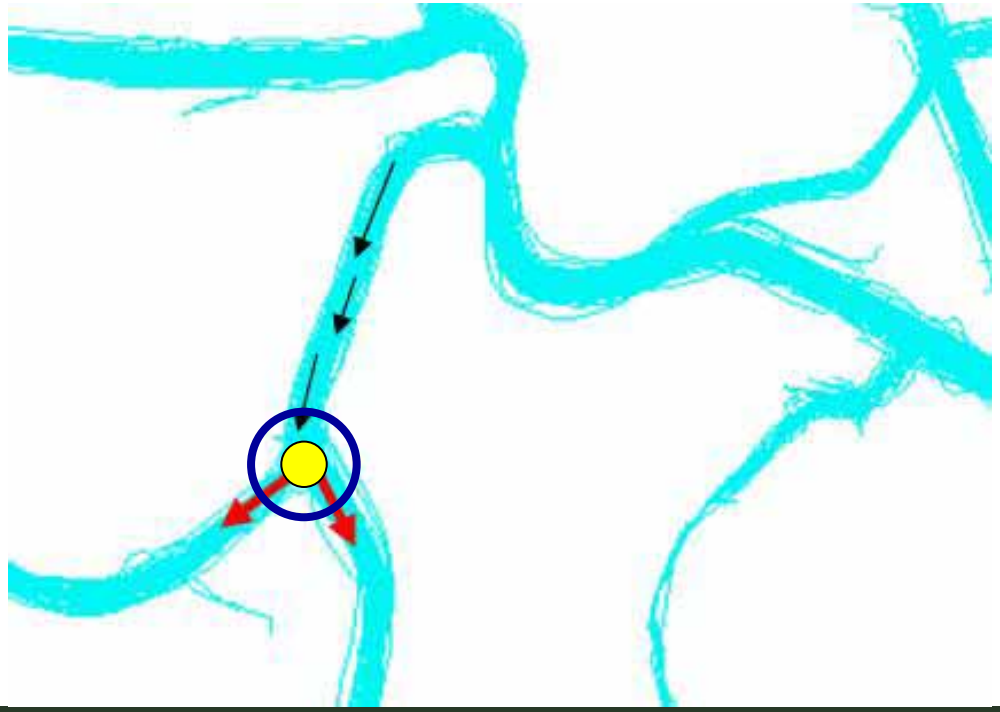
Each road of the tracked area must be covered by multiple GPS traces.

Our algorithm is implemented using spatial SQL queries to aggregate data from multiple traces to produce a weighted-mean geometry of road axles, diluting GPS errors.

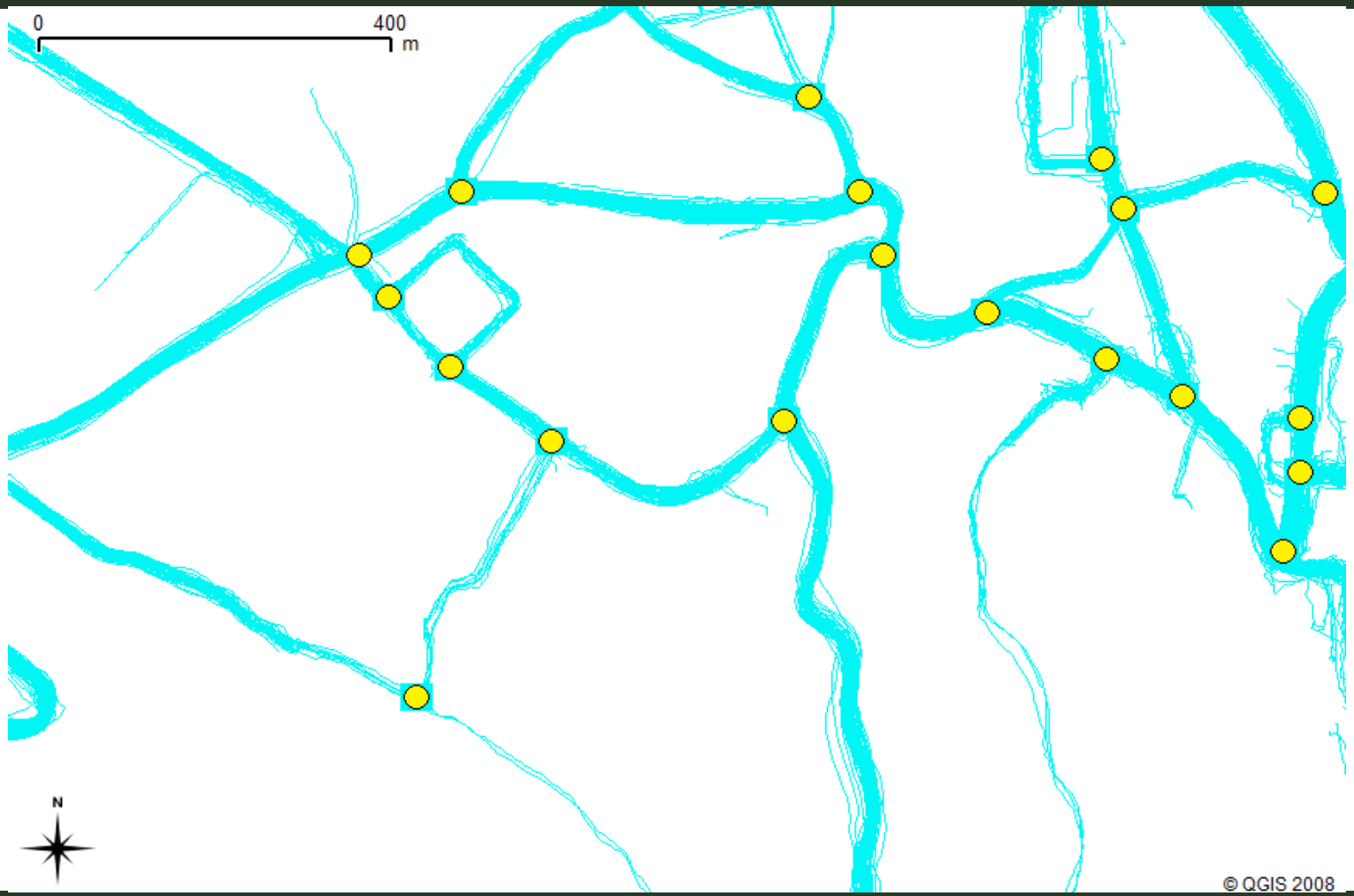
AUTOMATED MAP GENERATION

ALGORITHM

1. The first step of our algorithm is to identify **intersection areas** and to define intersection nodes;



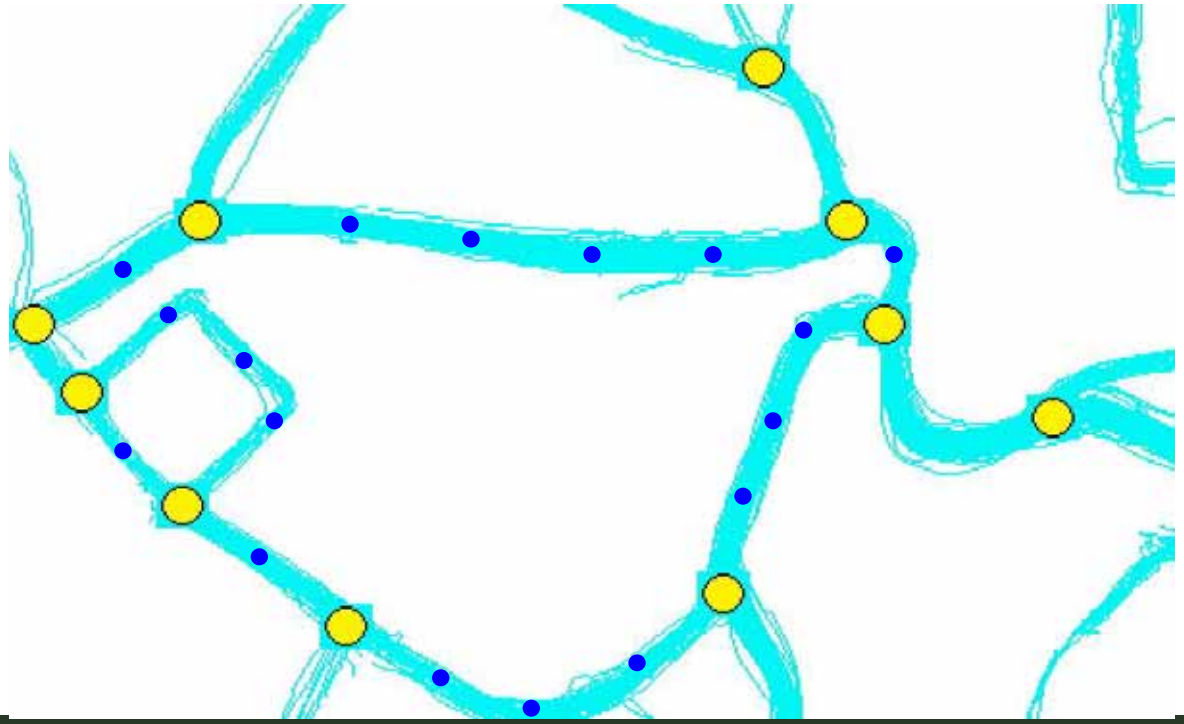
AUTOMATED MAP GENERATION



AUTOMATED MAP GENERATION

ALGORITHM

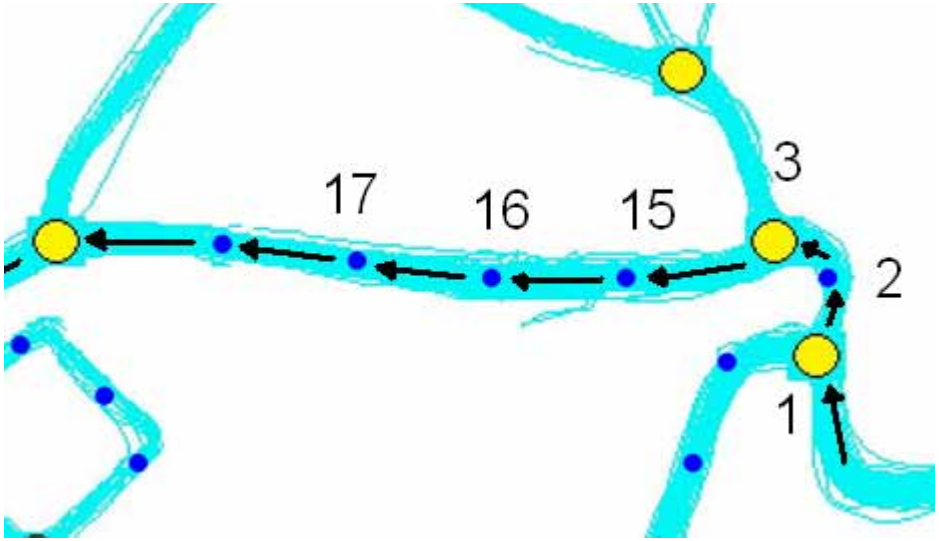
2. The second step of our algorithm is to apply a modified clustering algorithm for each set of GPS points situated between two intersection nodes.



AUTOMATED MAP GENERATION

ALGORITHM

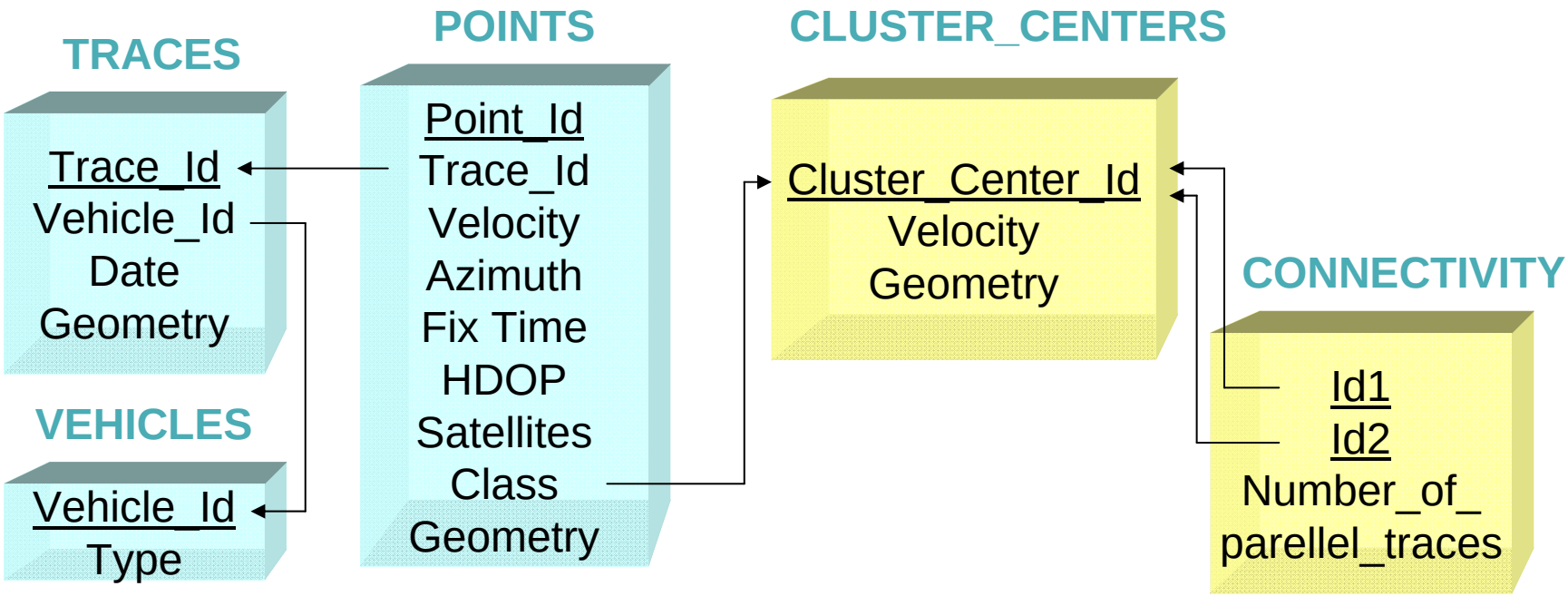
3. The last step of our algorithm is to establish the topological connectivity;



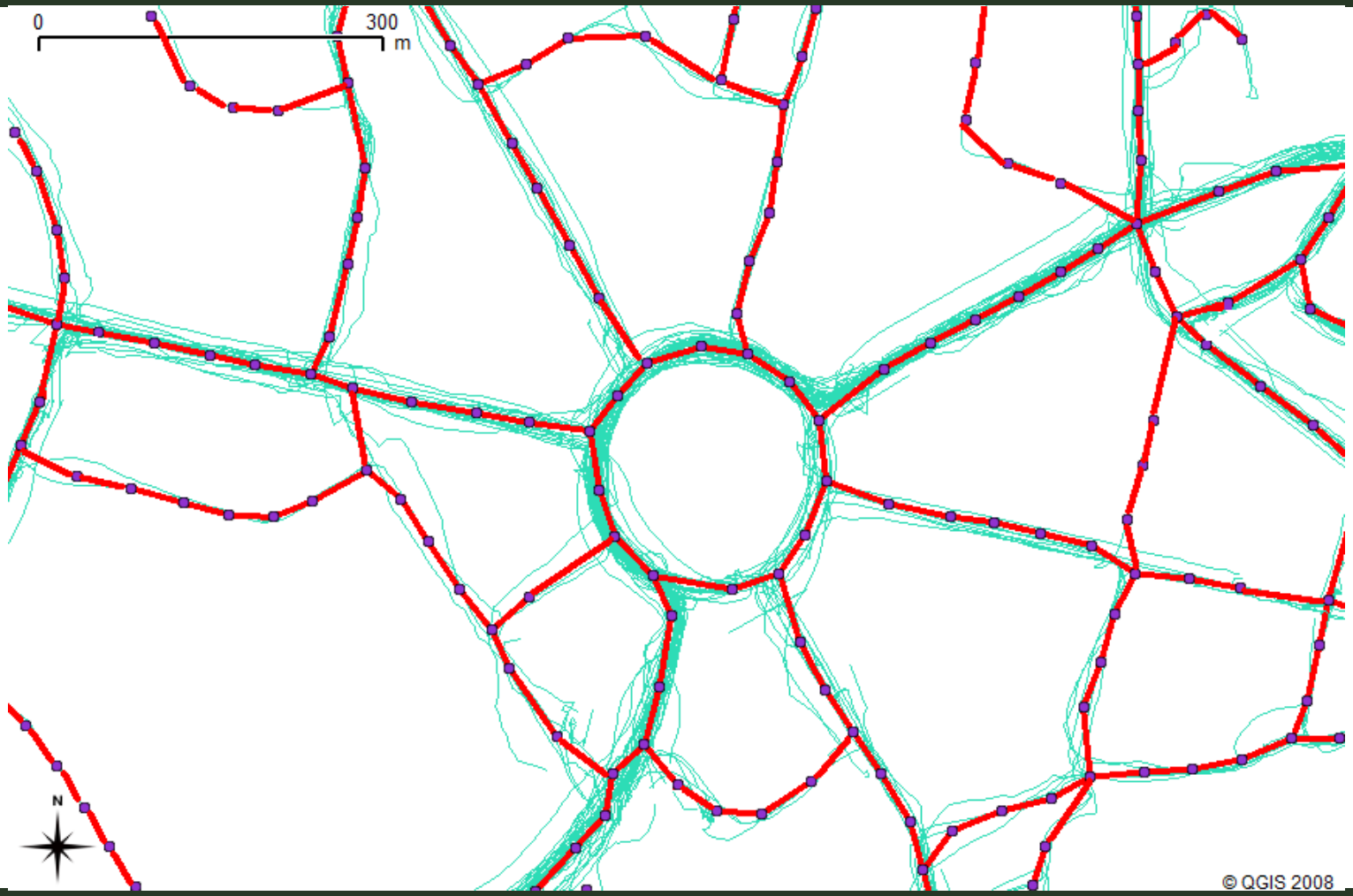
CONNECTIVITY		
Id1	Id2	Number_of_parallel_traces
1	2	54
2	3	54
3	15	47
15	16	47
16	17	47

AUTOMATED MAP GENERATION

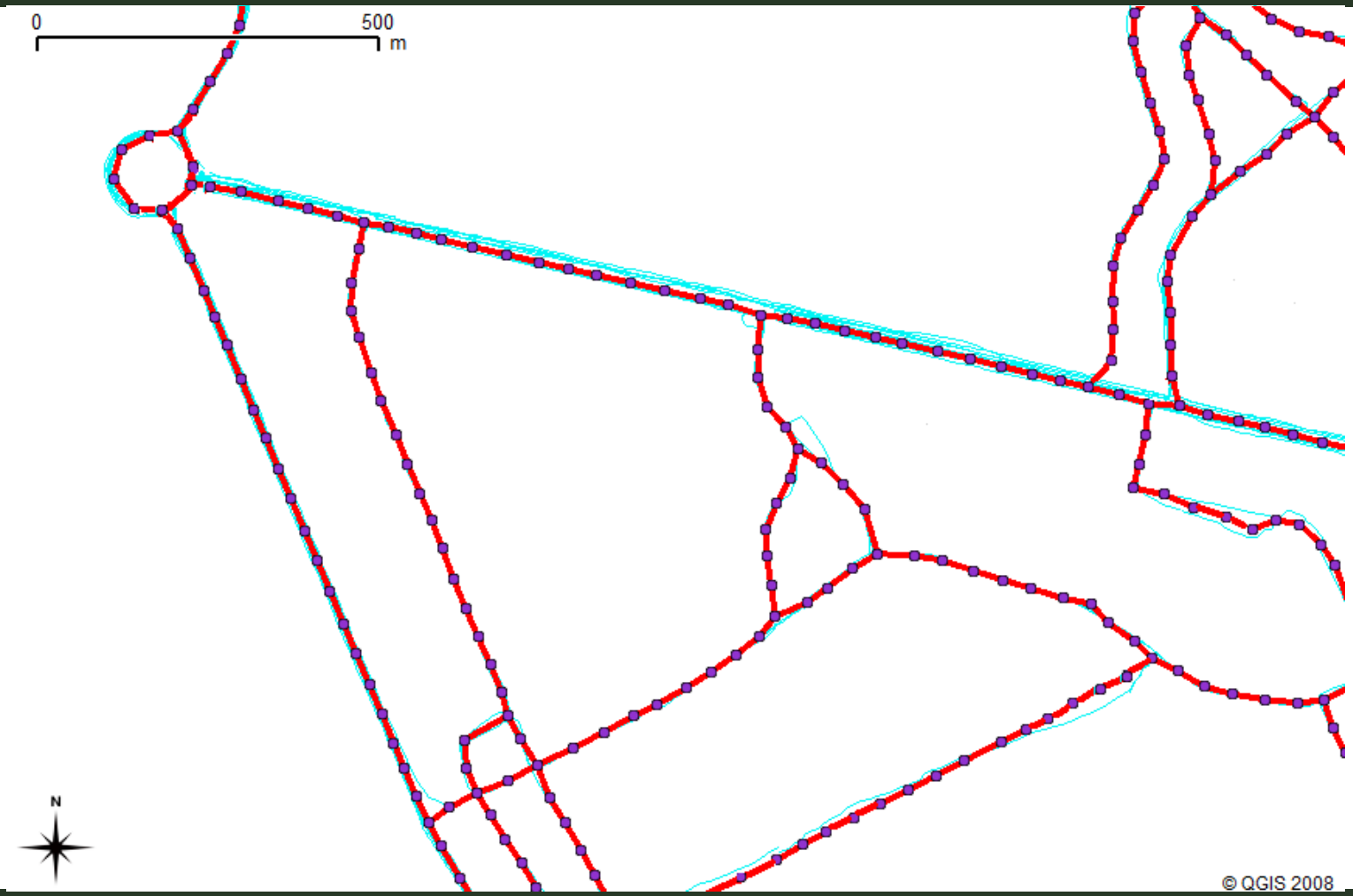
Spatial Database Structure



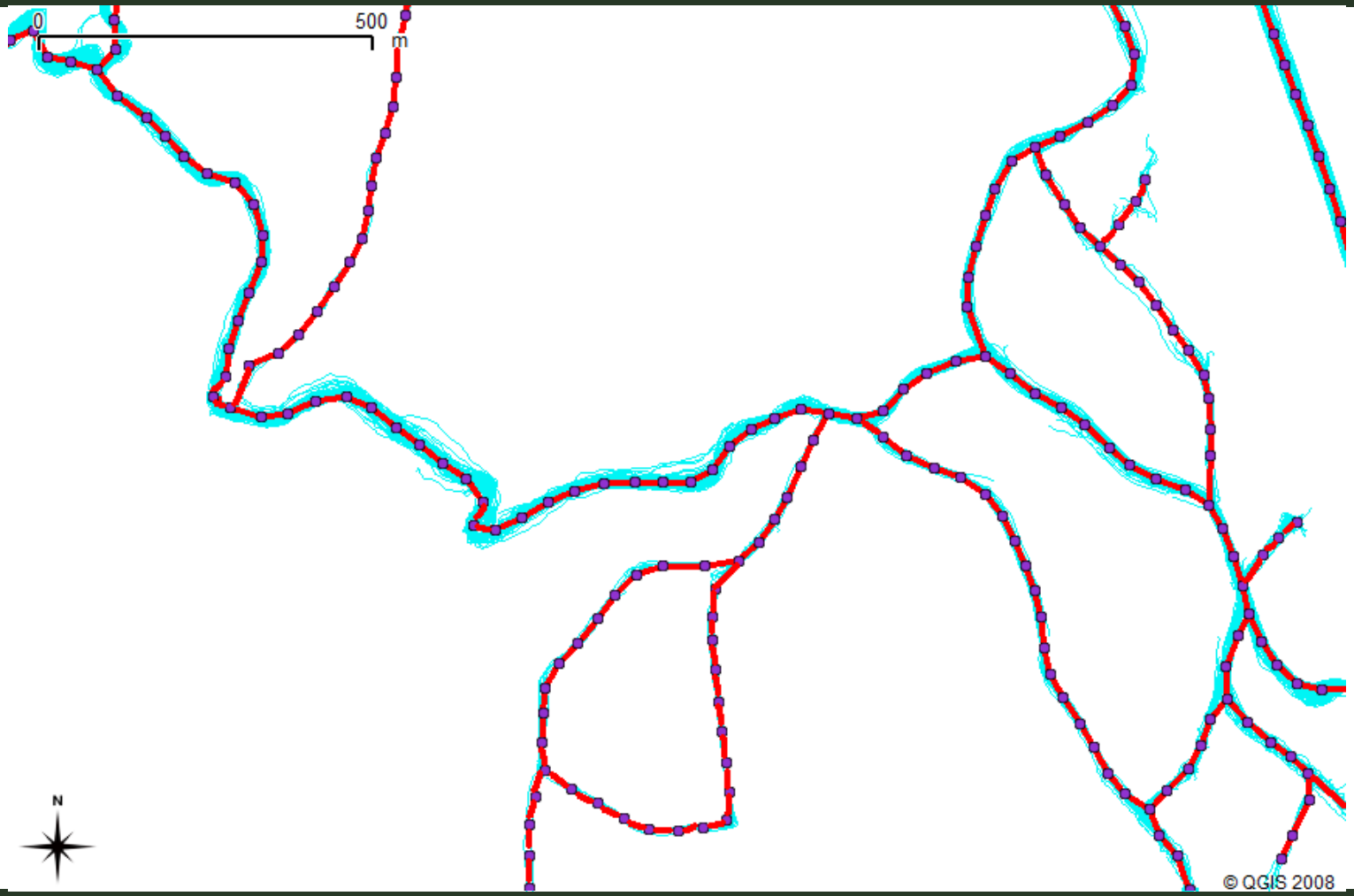
AUTOMATED MAP GENERATION



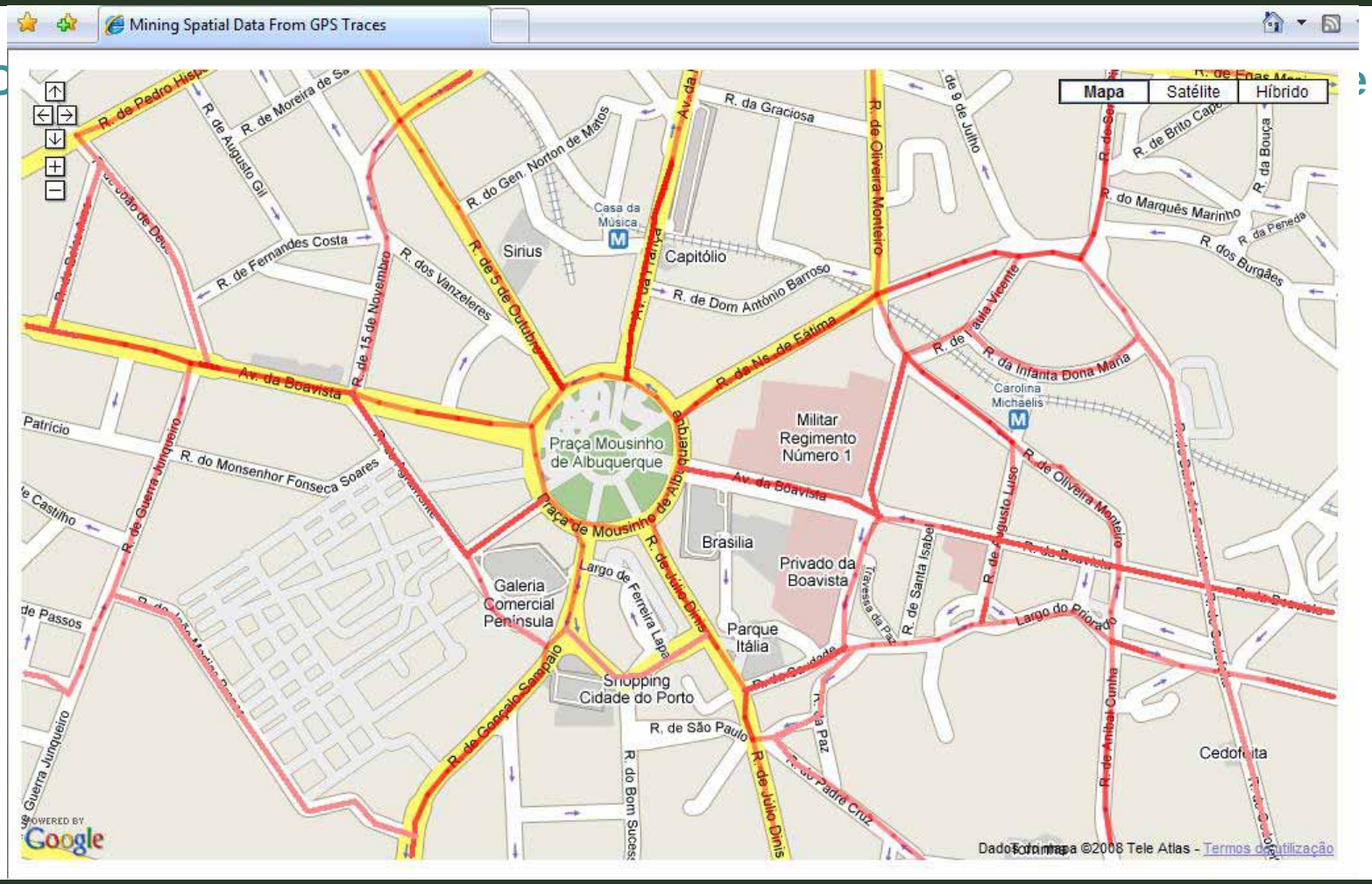
AUTOMATED MAP GENERATION



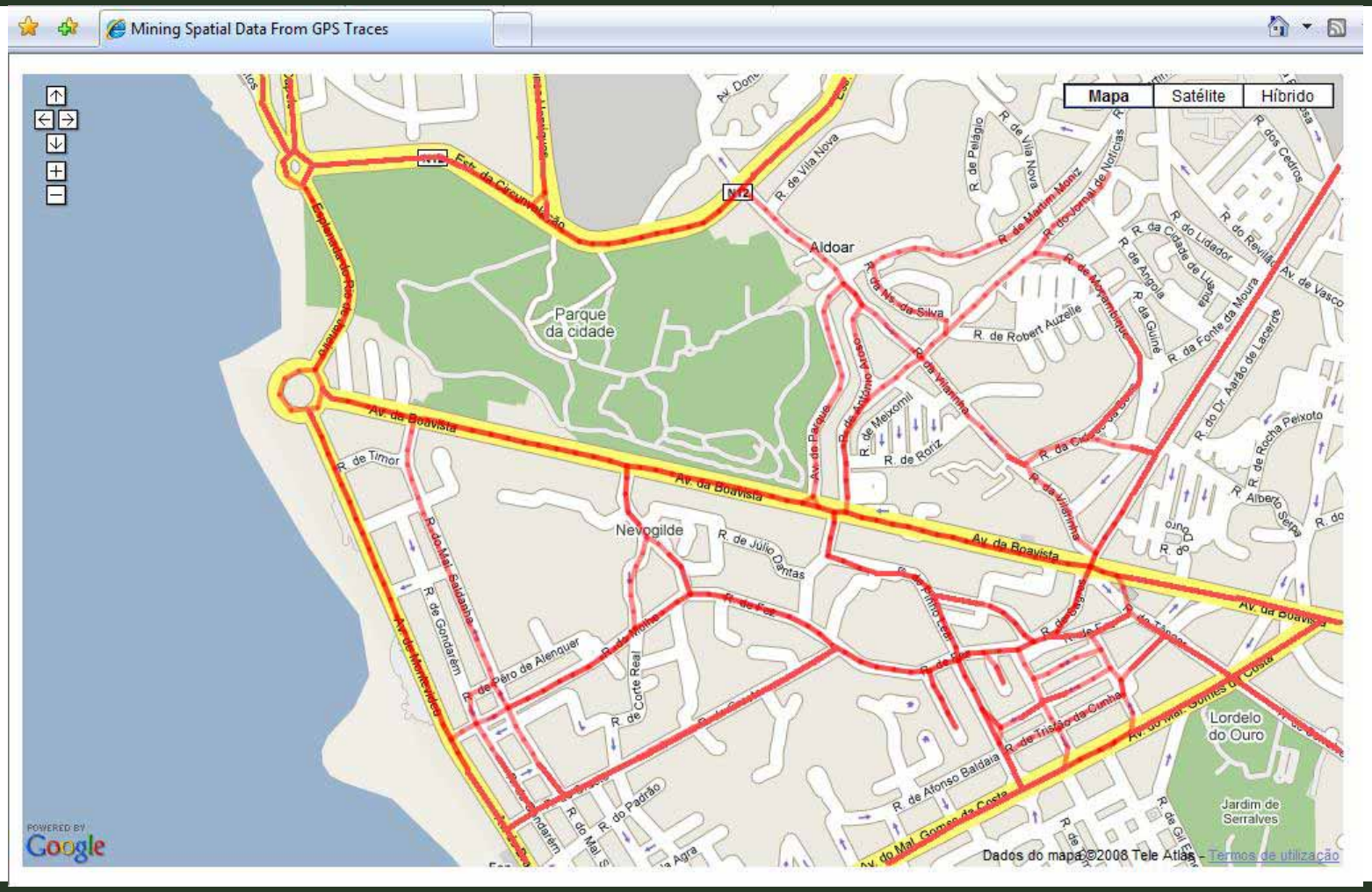
AUTOMATED MAP GENERATION



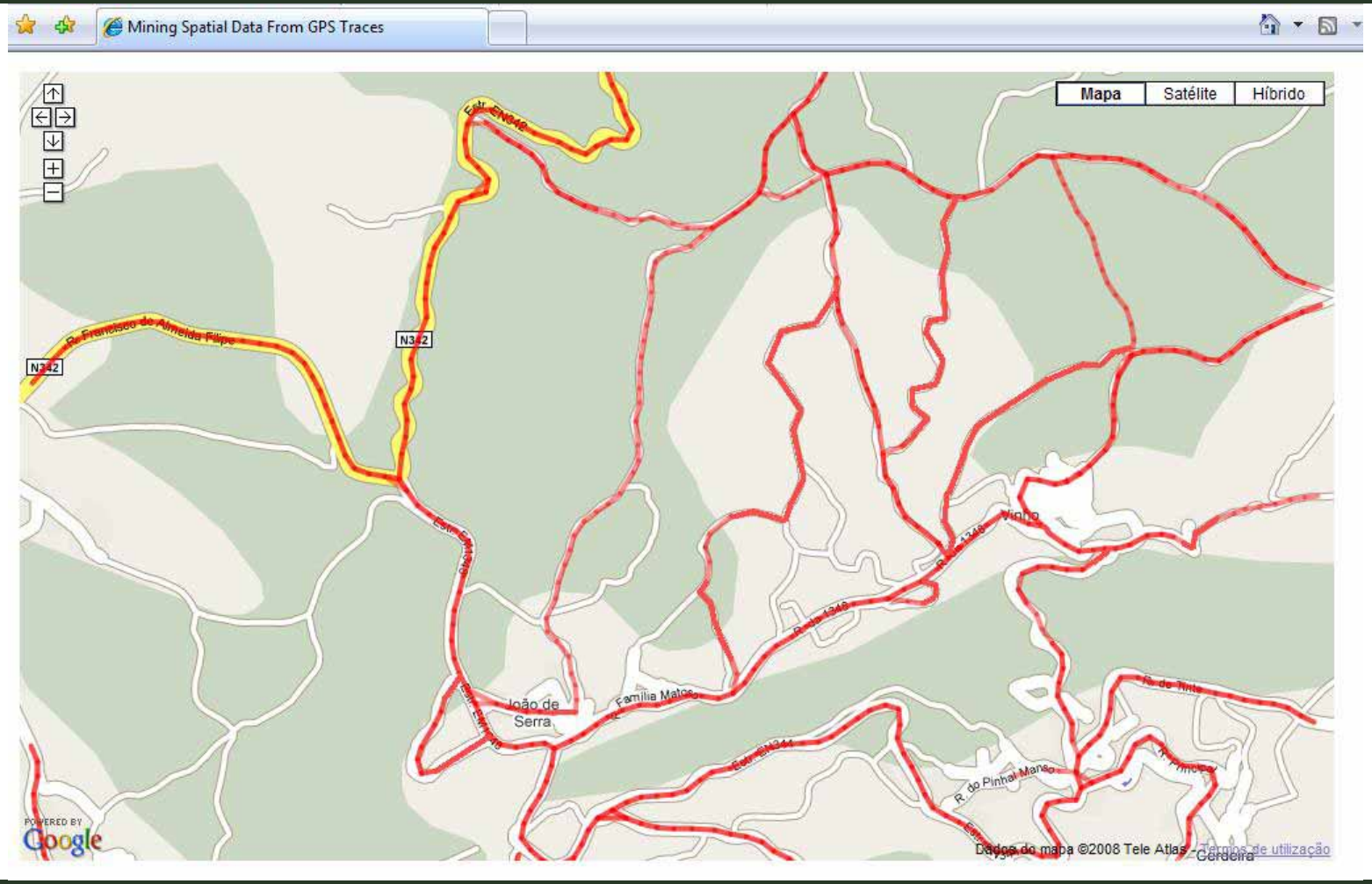
AUTOMATED MAP GENERATION



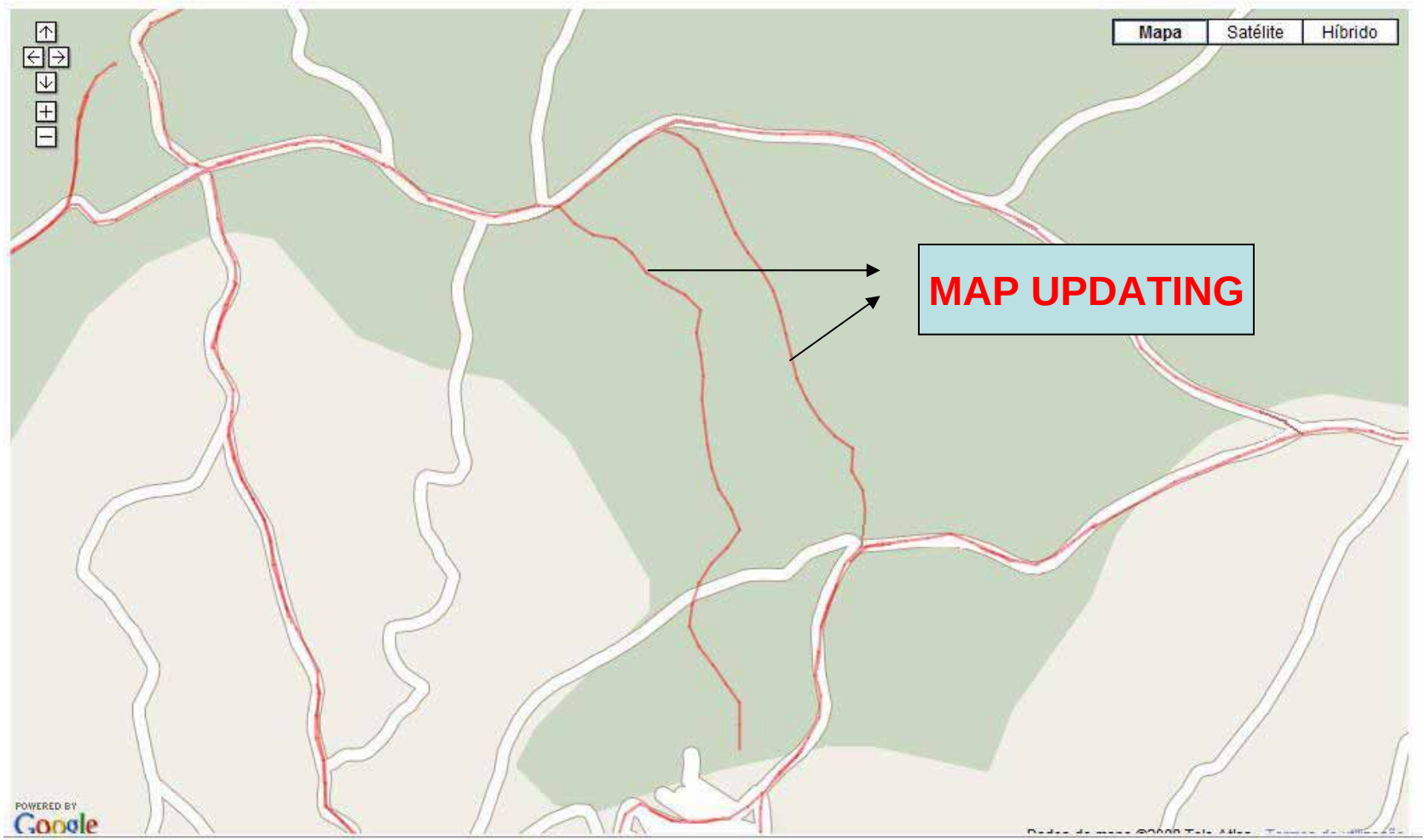
AUTOMATED MAP GENERATION



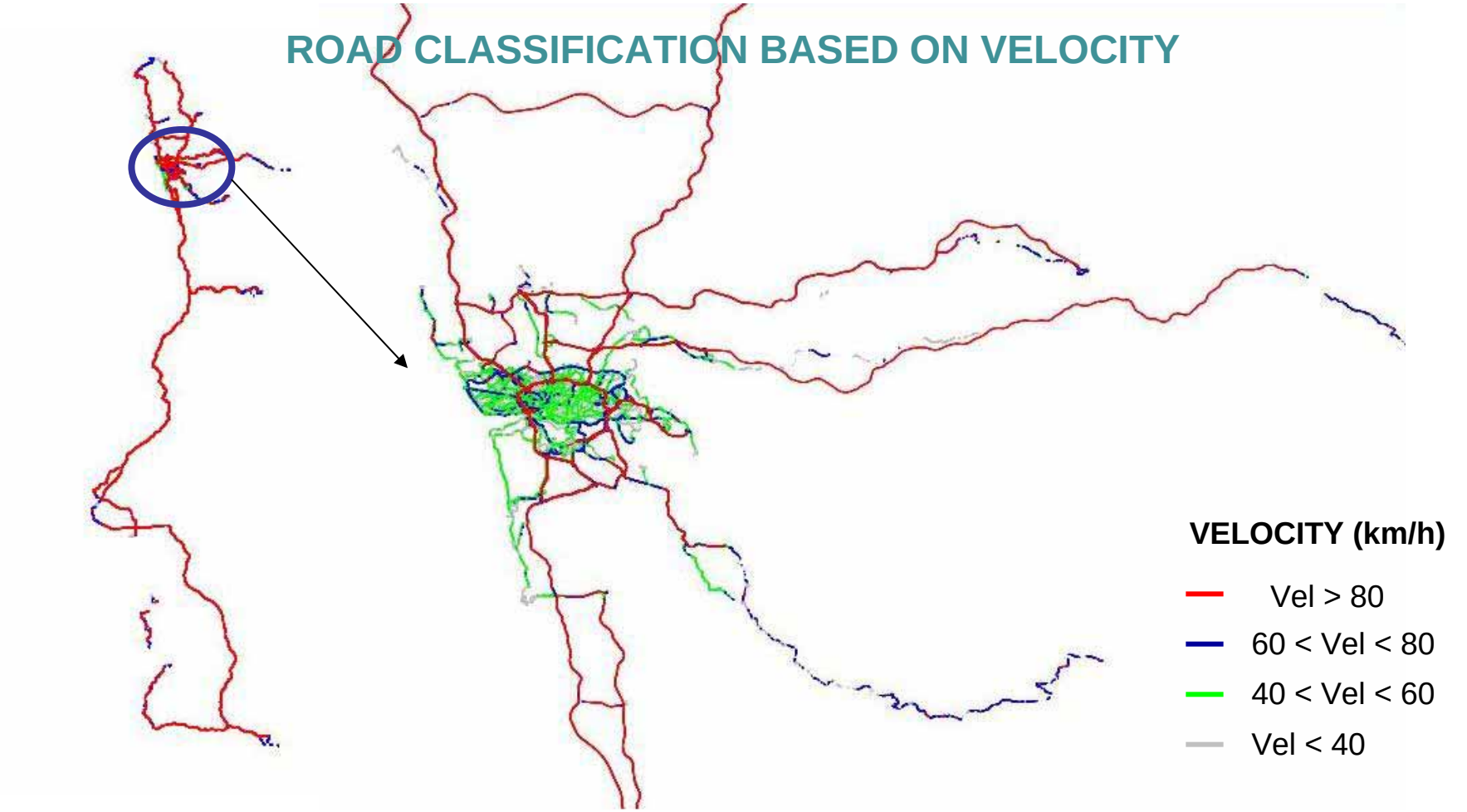
AUTOMATED MAP GENERATION



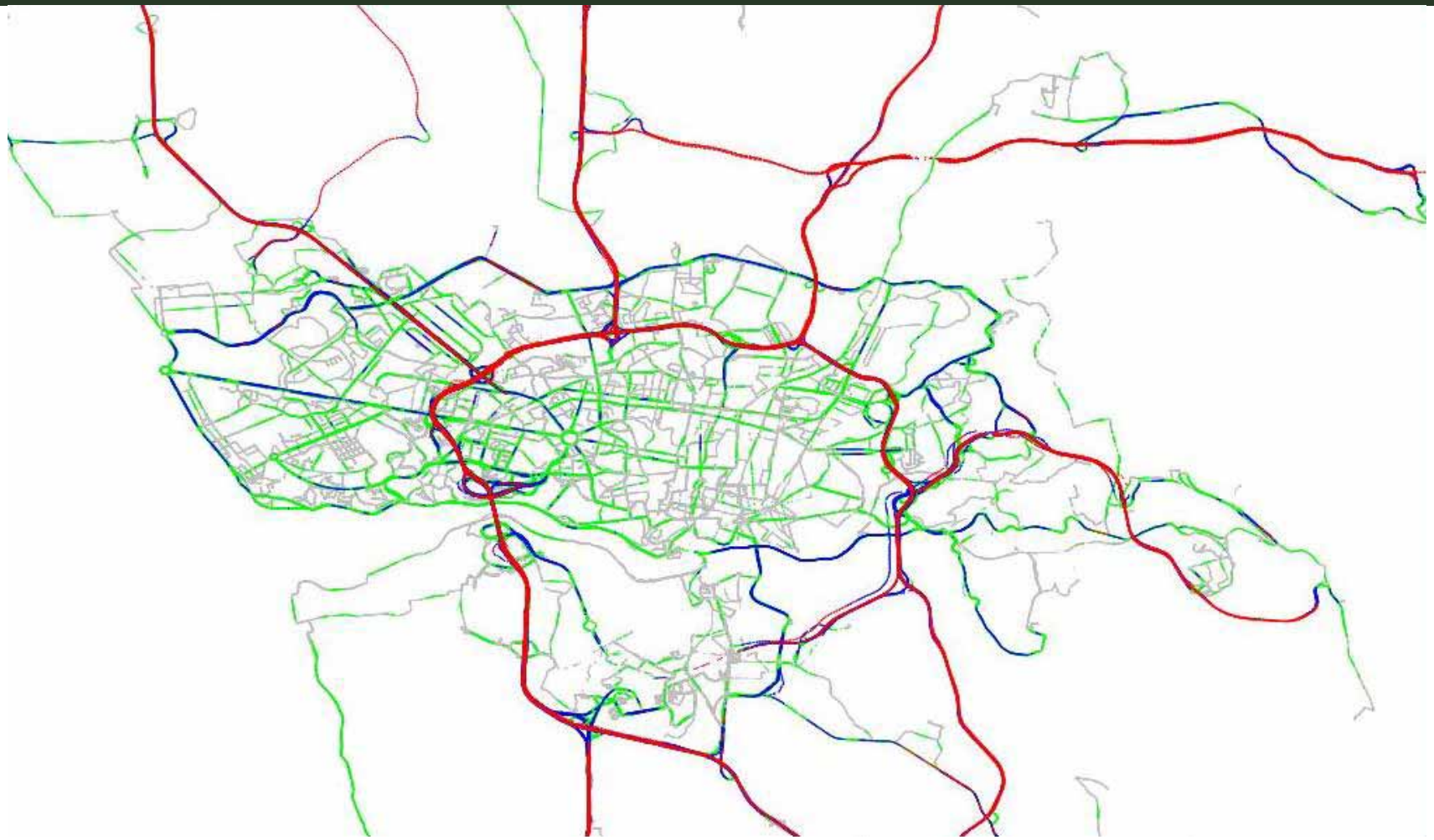
AUTOMATED MAP GENERATION



ROAD CLASSIFICATION

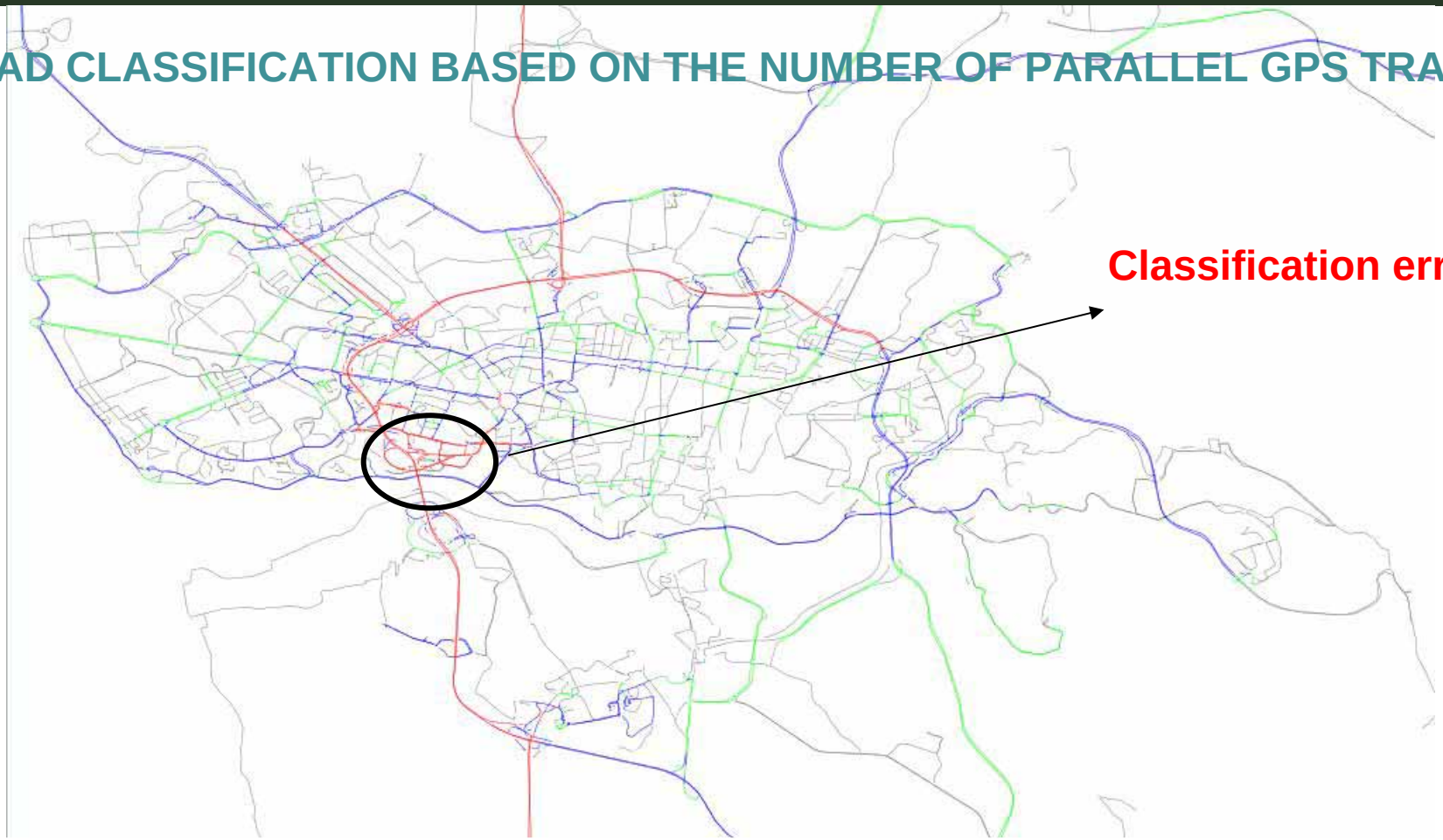


ROAD CLASSIFICATION

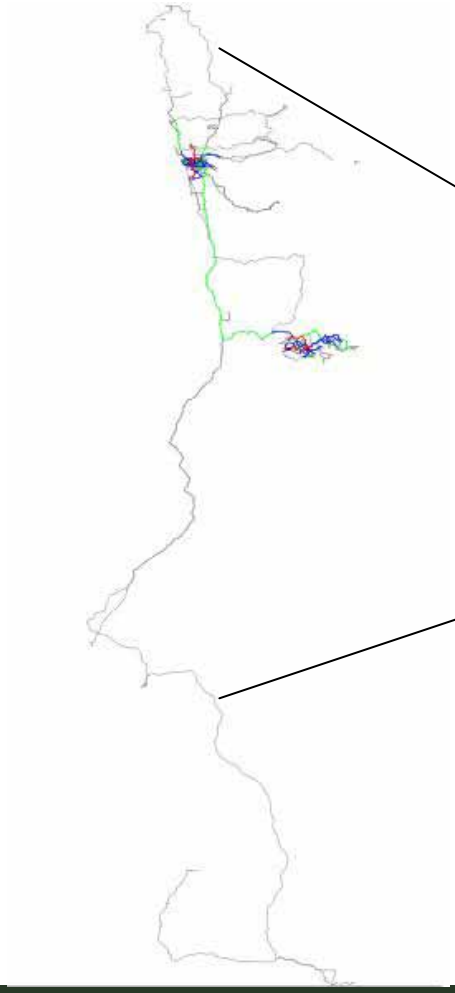


ROAD CLASSIFICATION

ROAD CLASSIFICATION BASED ON THE NUMBER OF PARALLEL GPS TRACES



ROAD CLASSIFICATION



Classification error

CONCLUSIONS AND FUTURE WORK

Advantages of our approach:

- no initial base map is needed,
- we can infer maps even in unknown terrains,
- new road segments can be integrated to an already generated map.

This study focused on the map generation and road classification processes, but data mining over position traces can yield more types of knowledge, such as traffic-lights location, parking information and points of interest of a particular vehicle.