Analysing Traffic Flow and Traffic Hotspots from Historic and Real-Time GPS Data

Christopher Bartolo and Thiago Matos Pinto

1. Software Engineering Cience and Technology Research Institute, University of Hertfordshire, Hatfield, Herts, AL109AB, UK

Dazzle Panel 53, Rudolph street, Sliema, Malta. SLM1275

2. Computational Neuroscience Science and Technology Research Institute, University of Hertfordshire, Hatfield, Herts, AL109AB, UK

Instituto Federal do Rio de Janeiro, Nilopolis, RJ, Brazil

Abstract: Traffic congestion is an increasing issue in many road networks. Considering the Maltese Islands, the traffic situation is in bad condition, despite the different attempts by the government to find a solution for this problem. In this report, we investigate how traffic may be analysed, which methods are available for this analysis, and how the traffic flow can be depicted. Many methods for this analysis are presented with a customized method attempted for finding traffic congestions and checking the traffic flow around the towns. The solution implemented in this work enables users to provide data through a data collecting Android app, which displays traffic flow on an interactive map, while allowing the users to view the traffic flow on a web-based interactive map at their chosen town, date and time. The method works by analysing GPS data from vehicles attained from different methods of data collection, namely data collection from Android devices, and historic GPS data from tracking devices installed in commercial fleets. The results of this work demonstrate a method for analysing traffic flow with accurate results, a usable and easy method of depicting traffic, with positive user feedback.

Key words: Road traffic analyses, road traffic flow, road traffic plotting, real-time traffic data.

1. Introduction

Traffic is a growing concern, especially in industrial areas. Although some towns and countries around the world have traffic management and control systems in place, it is still not properly handled in many others. For the purpose of this work, the main focus will be the Maltese Islands in Europe. The traffic crisis in Malta is continuously increasing, as can be seen in Fig. 1, which shows the statistics of how traffic increased over the past few years, with even non-peak hours having main roads across the islands congested. The worsening traffic situation on the islands is not only causing frustration, but also reflects an increasing cost for maintaining and attempting to control it.

1.1 Objectives

The study aims at investigating and analysing the traffic flow around the Maltese Islands through GPS data. This broad aim can be divided into the following:

• To gain insight of the traffic flow and traffic congestion on the Maltese islands;

• Developing an algorithm to analyse traffic from GPS data;

• Applying the developed algorithm and research to a real-world application;

• Improving upon existing solutions by gaining an understanding on why they are not wildly used around the Maltese Islands.

2. Results

A gap in existing traffic solutions such as Waze,

Corresponding author: Christopher Bartolo, M.Sc., research field: design of electronic systems.
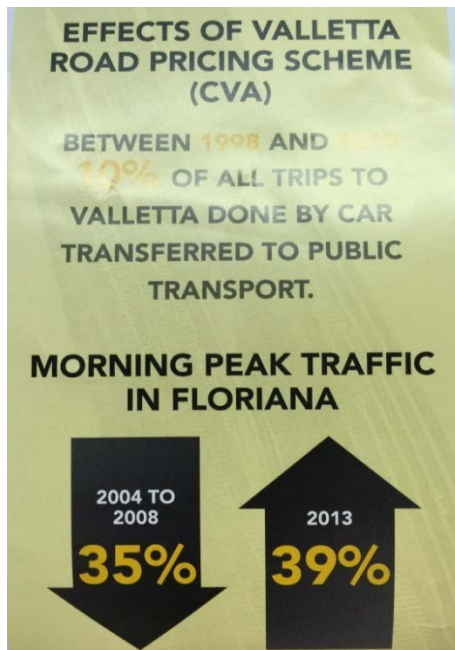
Fig. 1   Increasing traffic in Floriana, Malta [1].

Google Maps, and Nokia was found due to missing capabilities of providing traffic information for many countries such as Malta. Another issue is due to the fact that people disable location services on their phones not to be tracked by any other entity, and also disable their internet functions due to limited data packages. Without the ability of data logging from many users, there is a lack of data that may cause issues in trying to analyse and display correct traffic information.

In our proposed solution, tFlow, the main information is based on historic information, and commercial tracking information. The advantages of basing the algorithms on these are that data collection is constant and accurate. Having said this, it does come with its own disadvantages. A lot of priority has been given to historic data, so there should be roads under construction for a long period, and once the roads are re-open, the traffic data may be incorrect until the historic information is updated.

The best way to implement this solution to work efficiently and present correct and accurate data that cover most roads is to get as many tracking companies to provide their data. Ideally the data are provided in

real-time, as soon as it is logged, but recent historic data would be efficient enough to provide good information.

2.1 How Is Traffic Flow Analysed and How Can GPS Data Be Analysed for Depicting and Analysing Traffic Flow and Congestion?

Traffic flow is analysed based on both historic and real-time data which is logged from mobile phones that grant permission and access to the traffic solution provider. Throughout this paper, we present different algorithms, tools and solutions, however what can be clearly understood, is that no solution is actually capable of presenting information on small countries such as Malta. Commercial solutions are fully based on the buyer providing tracking information of their fleet. One of the largest transport companies in Malta, Supreme Travel Limited, have a fleet of around 90 vehicles; which surely is not enough to efficiently and accurately plot traffic flow around the Maltese islands. The Maltese public transport has been going through very tough challenges over the past few years, and are not a reliable source of tracking data, especially when considering that a lot of traffic and traffic accidents have been caused by themselves; also it must be considered that they have a low speed limit policy, which will badly affect the mean speed. Having a continuous flow of large data sets from multiple tracking companies will help to establish accurate traffic flow results both historic and real-time. Should there be large vehicles as part of the data set, the mean speed might be affected slightly, but not enough to badly affect the data. A data collection app will assist mainly in real-time data analyses to capture one-off incidents in minor roads where commercial vehicles may not visit.

With regards to how the GPS data can be analysed to depict traffic information, existing research analyse GPS data based on entire road segments, making use of either map matching of Viterbi decoding. In tFlow, the solution experiments with a different method and a

positive result were achieved. tFlow iterates through different points along roads, which are fetched from Open Street Maps. For each point, data are matched within a short radius, decreasing problems of inaccurate GPS positions, and filtering out invalid data. The implementation assumes that the speed within the current point and the next point is on average, the same, and therefore gives the road segment the mean speed calculated for that point.

3. Methods

An artefact was developed and implemented for this research. For fast development and a quick release from prototype to production, the Rappid Application Development methodology was used, which is explained below.

3.1 Rapid Application Development

RAD (rapid application development) is a method of designing and implementing software. The concepts behind it have been known and practiced for many years in the past, but were not supported by many automated tools.

RAD is mainly based on three principles, these being joint teams with trained personnel from both the "development and functional user organizations", computer aided software tools to understand requirements and to assist in designing the requirements information which can then be re-used in the software engineering process, and an "iterative process for demonstrating the software to users as it is developed" [2].

The term RAD was coined by James Martin in the early 1990s and it was used to "distinguish the methodology from the traditional waterfall model for systems development". The aim behind this methodology is to decrease development time and to produce higher quality code and results. With this methodology comes "successive iteration, refinement, and an accelerated movement from a prototype to a production system" [3].

3.2 Algorithms

Specific algorithms were developed and implemented for this research. While there are many different proposed algorithms which have already been researched such as that presented by Pankaj Vermaet. al., that presented by Daid Barth, Waze, Google, and Deepak Garget. al. [4-6], the idea behind the methods and iterations presented here are to identify if a different, altered method is applicable for the Maltese islands and possibly applicable also to other small countries; especially when considering that existing, commercial or free solutions do not cater for such small countries.

The first algorithm is used to fetch and construct a database of points forming the different roads around specific towns. An open source project that actively creates, maintains and distributes "free geographic data for the world" is Open Street Map [7]. By Nominatim API which is a "tool to search OSM data by name and address to generate synthetic addresses of OSM points" [8], we are able to retrieve the necessary data, process it and temporarily store it to be able to extract the information needed for this research. The data are stored in separate MySql tables, one table storing the name of the road, town, and its unique ID, another table storing the street's unique ID and the list of points forming that street. After analyzing this data, it was found that most points are shared by multiple streets, so by combining the data into only "unique points", improved speeds and results were able to be obtained for the next algorithm. This can be seen in Fig. 2.

The second algorithm was developed fully in MySql. The idea behind this algorithm was to iterate through every point obtained in the previous algorithm, and match it with the GPS data within a 20 meter radius. The matched data are divided into four segments by angle of direction, and the average speed for each segment is obtained. This can be seen in Figs. 3 and 4.

Create new waypoint data (function createWayData)

    Set limit to x

    Fetch list of points which do not have a street name

    For each point do

        Get possible matches for point from the Nominatim API

            For each match do

                Insert match into database with information of street name, town, and country

                Parse geoJson given by Nominatimand iterate through

                    Calculate angle and matching segment

                    Insert node into database

If count of points not reached, call again function createWayData

Fig. 2   Creating way point data.

```
SELECT
        AVG(assetSpeed) AS speed,
        assetHour AS hour,
        assetDay,
        assetMonth,
        assetYear,
        DAYOFWEEK(assetLastUpdate) AS dow
FROM
 (
SELECT
  *
FROM
 assethistory_main
WHERE
assetSegment = 1
AND assetSpeed > 5
AND assetPointLat BETWEEN $latitude-0.001 AND $latitude+0.001
AND assetPointLon BETWEEN $longitude-0.001 AND $longitude+0.001
AND
MBRCONTAINS
  (
  LINESTRING
     (
     POINT (
         $latitude + (20/1000) / 111.1,
         $longitude + (20/1000) / ( 111.1 / COS(RADIANS($latitude)))
         ),
     POINT (
         $latitude - (20/1000) / 111.1,
         $longitude - (20/1000) / ( 111.1 / COS(RADIANS($latitude)))
         )
     ),
  pt
  )
 ) t1
GROUP BY assetYear,
 assetMonth,
 assetDay,
 assetHour ;
```

Fig. 3   MySql query to calculate average speed within a radius of a point.

```
Fetch all speed data
For each record do
        Fetch matching nodes from database
        Compress nodes into a single node based on highest occurance
        Generate code for similar nodes, and update database with new code
        Insert the average speed in km/hr into the database with the code, day of week, month, day and hour
```

Fig. 4   The code to store the resulting data from the average speed by segment by point.

4. Implementation

4.1 System Breakdown

The system has been developed having the following in place:

- Interactive Mobile Map;
- Activity Recognition;
- Location Handler;
- Traffic Painter;
- Interactive Web Map;
- Web map-traffic plotter;
- Analyser;
- Logger.

A personal Android phone was used for the development of tFlow. The mobile used was a Google Nexus 5, manufactured by LG, with an Android version of 5.1. The location and activity recognition services provided by Android will be used to achieve our needs. The 3G connection provided by the device will be used to connect to the server.

4.2 Interactive Mobile Map

The interactive map allows users to drag the map around to view their preferred locations. The traffic layer is also painted over the interactive map using the APIs provided by the map provider, with the most important function in this case being the polyline function.

4.3 Activity Recognition

This feature analyses the device accelerometer and GPS information to determine whether the user is walking, driving, or standby. This information is gathered by Android itself, and a listener is used in the

tFlow application that makes use of this information so that GPS location and speed is logged back to the server while the vehicle is driving.

Using the Android Activity Recognition API proved to be problematic, with the activity close to never being recognised correctly. This caused issues for acquiring proper information for the purpose of this project.

5. Traffic flow Display

Android—On receiving the town name the "traffic painter" compares the current plotted traffic location by town. Should the town name received be different, the application communicates with the server to retrieve the traffic analytics. On receiving the processed data from the server, the traffic painter iterates over the traffic records, and paints the lines with colours according to the average speed matched to the point.

Web Interface—The web interface takes user input with the town name, day of the week, and the hour. This is then used to filter the data in the database, and the road points are matched to the traffic information stored in the database. These points are then plotted on the interactive map as a layer on top of the map.

Plotting the road points and joining them together proved to be quite a difficult task. The process involved is described below:

- The previous received point, and the current received point are compared. If the distance between these two is greater than 200 meters, the points are not joined, but considered as two separate roads;
- When two points are considered to be on the same road, the speed of the previous point is taken as

the speed between the two points, and the two points are joined by a line filled with that colour;

• When the two points cannot be joined either because the current point has invalid data, missing data or unaccepted speed results, the first idea was to ignore the current point, and join the previous point with the next. However this may cause a "jump" to show, badly plotting the roads;

• If the current point is ignored due to Step 4, the fix for the "jump" is applied. This still sets the current point as the previous point, but colours it as transparent, so the "jump" is not visible, and the lines are plotted accurately over the roads.

5.1 Analyser

The analyser algorithm is the essence of the tFlow solution. It has four main functions, these being road point fetching, historic data processing, real-time data processing and traffic plotting.

5.2 Road Point Fetching

The first step of the analyser process is fetching all the roads around Malta, and storing them for matching. Each road is referred to as a way, and each way is made up of a set of nodes which are linked together in a specific order. This information is retrieved from Open Street Maps, which had a lot of missing information such as town boundaries. To fix this, the author took the time to contribute data to OSM by drawing and setting the town names and boundaries. After this process, the data were deleted and re-downloaded and processed. While processing the points, the angle between the nodes is calculated and stored. The angles are also put through a function which analyses in which segment the angle falls in.

5.3 Historic Data Processing

Three months of anonymous historic GPS position data were provided to the author by a tracking company which wishes to remain anonymous. This data consisted of GPS coordinates, speed, direction angle, ignition status, movement status, date and time. The historic data process is made up of the following steps:

• If and record contains invalid data, or speed is less than 5km/hr, delete it;

• The time and date are altered and moved to two separate fields as "day of week" and "hour";

• The direction angle is calculated and changed to show the segment number as one, two, three or four;

• For each road, the nodes are fetched in an ascending order;

• Each node is queried for the historic data within a short radius, facing the same direction/segment. The average speed is calculated for each "day of week" and "hour";

• Each unique node is given a unique ID, and all matching nodes in the same segment are given the same ID;

• The ID is stored in a table, with the average speed, day of week, and hour.

5.4 Traffic Plotting

Analysing a large amount of historic data (amounting to over 31million records) required a lot of processing power, and was taking several hours to complete. A lot of time was spent creating the steps presented above, while also designing and optimising the MySQL tables and MySQL queries to be able to achieve a processing time of less than 2 hours.

5.5 Real-Time Data Processing

Real-time data processing goes through the same steps as the historic data processing, but with a smaller data set.

5.6 Traffic Plotting

The traffic plotter provides traffic flow information by giving a list of points, with a colour assigned to each point. The steps are below:

• The user inputs the town, day of week, and hour;

• The roads belonging to the select town are

fetched;
- The road list is iterated;
- For each road, all the nodes are fetched;
- For each node, a MySQL query is executed. This query fetches the speed assigned to the node by segment, hour and day of week;
- The node is assigned a colour based on the average speed, using the following legend:

(1) If average speed is 0 (possibly invalid, or heavy traffic congestion), colour is black;

(2) If average speed is between 1 and 14km/hr, colour is red;

(3) If average speed is between 15 and 24km/hr, colour is orange;

(4) If average speed is between 25 and 34km/hr, colour is yellow;

(5) If average speed is between 35 and 49km/hr, colour is blue;

(6) If average speed is between 50 and 64km/hr, colour is green;

(7) If average speed is greater than 64km/hr, colour is bright green.

6. Conclusions

Throughout this research the main focus was placed on understanding and analysing traffic based on GPS data by using road points rather than road segments, how traffic may be plotted with Malta at the base of the concept. Traffic is a major growing concern, and the Maltese government seems not to be close to finding a sustainable solution. The first point obvious point which comes to mind, is that the government continuously starts new road demolition and construction projects without first finishing off those which are still pending. An example of this is a few months ago, when one of the major roads of Malta, the coast road, was demolished, while another major project concerning a major road and roundabout, which was quite close, in Msida, was also started. At the same time, a bypass from Birkirkara to Mosta was also under maintenance, and a road which links Bugibba to the coast road, was and still is under maintenance. While a traffic flow analyses will provide an idea to both personal users and commercial fleet operators to how they should plan their trips and routes, it is still necessary for the core causes of traffic to be tackled.

Another major issue which is proving to be very difficult for the government, is the terrible and failed change-over of the public transport during the last few years. Although the operator has been once again changed this year, 2015, people still do not trust the public transport, and the use of personal vehicles is increasing while the use of public transport has decreased greatly. Car pooling appears to also be a rare situation in Malta, and close to no initiative is being given or placed for it, neither for car sharing, cycling, or walking for short distances.

Traffic accidents are very common, and constantly increasing, and most of the blame can be placed on road rage and driving frustration. With the use of proper traffic management and analyses, preparations and safeguards can be put in place. With the use of tFlow, and the possible future enhancements, traffic can be better controlled and understood. Claims caused by badly maintained roads can be reduced, and one-off traffic congestions may be found and taken care of faster. Apart from this, making use of the historic traffic analyses performed by tFlow, traffic hotspots can be found, urging the proper authorities to investigate and find the causes of such congestions.

References

[1] Msida. 2015. "I Am Traffic: Impact and Cost of Road Traffic In Malta." *Institute for Climate Change and Sustainable Development*.

[2] Harris, J. B., Eichorn, F. L., Goodwin, D. L., and Henson, J. L. 1997. "Use of Rapid Application Development Techniques: Designing that Staff Resource Tracking Tool." *Logistics Management Institute*.

[3] Agarwal, R., Prasad, J., Mohan, T., and Lynch, J. 2000. "Risks of Rapid Application Development." Presented at the Communications of the ACM.

[4] Barth, D. "The Bright Side of Sitting in Traffic: Crowdsourcing Road Congestion Data." [Online].

Available:
http://googleblog.blogspot.com/2009/08/bright-side-of-sit
ting-in-traffic.html.

[5] Garg, D., and Tripathi, J. 2010. "Algorithm for Detection of Hot Spots of Traffic through Analyses of GPS Data." Patiala.

[6] Waze, 2015. "Coutnries and Languages." Waze, 4. [Online]. Available: https://wiki.waze.com/wiki/Countries_and_Languages.

[Accessed 19 10 2015].

[7] OpenStreetMap. 2014. "Wiki: Open Street Map." Open Street Map, 10 7. [Online]. Available: https://wiki.openstreetmap.org/wiki/Main_Page. [Accessed 19 10 2015].

[8] Twain. "Wiki: Nominatim." Open Street Map, 2 10 2015. [Online]. Available: http://wiki.openstreetmap.org/wiki/Nominatim. [Accessed 19 10 2015].