Floating car data

A technical & business revolution in traffic management

Alessandro Taddei
Infoblu S.p.A. – Italy
alessandro.taddei@infoblu.it / alessandro.taddei@autostrade.it

Abstract
Floating car data can be looked at as the first step of V2I environment where the most suitable return channel is TISA Tpeg™. A very number of markets were boosted since May 2000 when the U.S government discontinued its use of Selective Availability in order to make GPS more responsive to civil and commercial users. Even though more than a decade is already gone investigations on the impact of applications and services based on GPS analysis for traffic management is still ongoing. First of all the paper provides a background on gps and fixed sensor data then looks at gps from business perspective, focusing also quality and privacy issues of this big data. Then the paper analyzes gps based applications for traffic management. Development took place in two phases requesting a change in the business model too. In the end the paper depicts new research areas including an initiative currently on going in the EU R&D Mobinet project.

Keywords: Fvd/ gps/fcd, big data, traffic monitoring
INTRODUCTION (BUSINESS AND TECHNICAL BACKGROUND)

Since US president Mr. Bill Clinton opened the gps usage for civil and commercial purposes in May 2000, many market arenas started and became very important. Most of them are B2C markets like car navigation devices, personal navigation devices, traffic, runner devices etc. On the other end B2B and B2B2C markets were boosted too; examples of them are fleet management, usage base insurance, etc.

Floating car data can be looked at as a first step of V2I environment where the communication infrastructure is provided by mobile communications. The most suitable return data channel for end user is represented by TISA Tpeg™ Traffic Flow and Prediction that is able to describe speeds in segments than can be as small as tenths of meters. [1] [2] [3]

A different scenario enabled by the same data is represented by traffic management application. In fact data are received, processed and then delivered to traffic managers / traffic control centres.

This first step is also a real business case. According to a post on Linkedin ® by RnR Market Research the market for Traffic Management in 2014 is estimated to be 3.56 billion USD and is expected to reach 16.89 billion USD in 2019.[4]

At the same time the cost Total Cost of Ownership of old style technologies like loops, triple/infrared/.. sensors is still very high and is discouraging road management organizations to make new large investments.

Many new sensor technologies had been developed but they did not become a market standard. For instance usage of Bluetooth probes as sensors in different road sections were seen as very promising but latest news about scrambling MAC address to prevent from for privacy issues are lowering the interest about this technology. [5] [6]

FVDS ARE BIG DATA

Analyzing gps data is a challenge. First of all gps are big data, receiving and analyzing gps positions means to be able to deal with a continuous flow of data whose size is in the range of tenths of gigabytes a month.

As an additional challenge the data can’t be processed off line but has to be processed in few seconds from their reception.

Business and quality aspects

Floating Vehicle Data are produced by a very number of sources like taxi companies, fleet management companies, insurance companies etc.

In the beginning the cost of gps was mostly influenced by the transmission cost. Gps positions were sampled every now and then they were grouped and sent in a text. The cost was very high and the availability of gps positions was very low.

In recent years due to a very large diffusion of B2B or B2B2C applications based on gps
along with the constantly decreasing cost of mobile data communication (2G, 3G+ …) large bunch of data can be purchased at low cost.

Worth considering gps data are produced for a main B2B target (e.g. fleet management) and their cost is recovered by this main target, this implies that they can be sold at low price for a different target (i.e. traffic management application) because this side sell is a pure additional value for the company owning them.

It is well known that commercial gps have low quality due to intrinsic low precision of receivers and negative effects that can influence (e.g. canyoning) but the large quantity and the freshness of the data overcomes negative aspects.

While many years ago gps were loosely sampled in time and space and were also affected by a delay in transmission due to the grouping before sending, the new trend is to sample at regular time frequency (between 2 and 60 secs) and to send just after the sampling or to have very short buffers before sending.

Our internal statistics shows that more than 95% are received in less than 10 minutes from the sampling.

Privacy issue

The paper has focused onto gps produced in a B2B scenario. There are two main reasons why this is the preferred scenario:

- quantity
- privacy

It is crystal clear that negotiating gps provision with one B2B provider is easier than making the same negotiation with a very number of end users.

Anyhow there is an additional important reason: a large provider as a fleet management company or an insurance company has in place contracts and systems allowing to be compliant with privacy regulations in different countries where the company operates.

The provision of gps to a third party, for instance for traffic management applications, is anonymized. This is large saving for the third party that is thus, by definition, not impacted by privacy regulations.

Conversely the same reasons can demonstrate why the idea of selling position of single end user is currently not viable on an economic perspective.

**TRAFFIC MONITORING APPLICATIONS - A FIRST SET OF APPLICATIONS (AND ISSUES TO BE SOLVED)**

When we started developing applications in the traffic management arena and mainly devoted to traffic monitoring we had a set of issues to be solved:

- map matching gps data
- fusing data sources
- choosing the application to be developed and develop them
We decided to have a three layers architecture:

- map matcher
- fusion engine
- applications

In the beginning the map matching system was quite ‘rough’ and able to match single positions ‘alone in the world’.

The fusion engine was able to fuse data from the map matcher along with pre-elaborated data provided by legacy systems we had in the group our company belongs to.

As a final step we had two main applications: a map and a traffic heatmap [Figure 1 - Heatmap (old version)].

![Heatmap (old version)](image)

Both the applications were able to demonstrate the data we had and to show the traffic situation in any moment.

Both the applications were two separate islands ie they were not integrated and, in addition, there were a lot of business logic inside each of them.

This was a first major drawback both from a user end point and from maintenance perspective.

Traffic managers (that were users of our applications) found boring and time consuming to repeat the same operations any time they looked at the same traffic issue.

At the same time any time we needed to re-do any modifications in the basic bricks twice, once in any applications.
TRAFFIC MONITORING APPLICATIONS – A FIRST (LIMITED) INTEGRATION

After having delivered the first set of application we started thinking to integration of the two applications.

As a first limited integration we added hot points to the map and our users were able to move quickly from the map to the heatmap thus avoiding the need to select by the user interface the same location.

Unfortunately at that time our users were not able to do the reverse path.

BUSINESS MODEL (R)EVOLUTION

At the time of fixed sensors road authorities made investments to install and run fixed sensors on roads to be managed. Then they selected a (or more) software product that fitted their target(s) and purchased it (them). The fixed sensors had recurrent costs (e.g. electric power) and maintenance costs. The software products usually had license costs, tailoring costs and maintenance cost.

This was the most popular business model.

The usage of gps, in addition to lower the costs, also changed the business model.

Gps data are available as side effect of fleet management, usage base insurance etc. This means they are generated on a large set of roads and not only on the road of interest of a particular authority. In addition as described in previous section they are sold in bulk quantity. This means that organizations elaborating them do the job on a very large set of roads. Doing the task for a single customer at a time is not economically viable.

In the new context the fixed sensor business model does not fit very well.

Thus we started selling services instead of selling software licenses. Any applications had a yearly fee and costs depended on roads, road length too, etc.

Even though the Total Cost of Ownership was lower than in the previous model we faced some initial issues due to the fee management in the balance sheets by customers (at least in the country where we operate, licenses are accounted as CAPEX while services are considered OPEX). [7] [8] [9]

An additional issue was related to the Service Level. In the licenses based model the service level has not a great relevance unless the customer wants to negotiate bug fixing support (that is usually not common due to its high costs).

The service business model we applied was dramatically impacted by Service Level Agreements. Our customers were mostly Traffic Control Centers operating 24x7.

TRAFFIC MONITORING APPLICATIONS – THE CURRENT SCENARIO

The second big phase of our internal evolution faced different topics:

- new easy2use versions of existing applications
- new applications
Floating car data – A technical & business revolution in traffic management

- new approach to application integration
- new architecture

**New easy2use versions of existing applications**

The first activity we did was a revision of existing applications making them more easy to use.

Our end users needed to focus on their tasks and to be able to concentrate on them and not to think how to get informed, for instance, about a traffic jam.

One of the most appreciated features we added was the space and time panning in the traffic heatmap.

Using a simple drag&drop end users were enabled to move in space (up or down) and in time (right and left).

The play button on the right top allowed in any moment to get back to current time (Figure 2 - Heatmap where mouse cursor shows a hand indicating panning feature).

![Figure 2 - Heatmap where mouse cursor shows a hand indicating panning feature](image)

Other shortcuts were added to display data far in space and in time (the application allows to have a direct access to last 12 months of detailed data – any moment in time is displayed in less than 3 seconds from the request).

The map was upgraded with the same focus: the map can display different data like current traffic situation, historic data, differences from historic data etc. All these data were made immediately accessible via simple menu always available on the top right end of the map itself (Figure 3 – the new map with the menu on the top right corner).
The caption of the traffic map and of the traffic heatmap is the same and is always displayed on the screen, too.

![Traffic Map](image)

**Figure 3 – the new map with the menu on the top right corner**

**Brand new applications**

We designed and developed two brand new applications: the gps diagram and the alternative paths.

The gps diagram displays single vehicle tracks.
The tracks are gps positions that are previously map matched. Our application allows to display light or heavy vehicles (Figure 4 – the track diagram).

The alternative paths is an application that display the travel time of defined paths that are external and alternative to some motorways paths. Our customers operate in traffic control center of different motorways. Each of our customer organization, by legislation, has defined some paths external to the stretch of motorway they manage. These paths are defined in advance in agreement with police, local and regional authorities, etc. The traffic stream is redirected to one of them in case of severe accident on the motorway. Our application allows them to monitor the overall travel time of the alternative path in the last 3 hours (Figure 5 – the alternative paths application (main page))
New approach to application integration

Our first integration approach was very limited and was based on the capability of a first application (map) able to call an API of the called application (traffic heatmap), the way back was not implemented.

This approach (direct call) was extremely complex when increasing the number of applications and thus the number of needed combinations. In addition we were requested by our customers to be able to synchronize all the applications in space & time upon end user...
Floating car data – A technical & business revolution in traffic management

request in order to look at the same traffic issue in all the applications. Thus we approached the issue by a new concept: an additional application called toolbar allowing to supervise all the other applications and making them possible to communicate. All the communications were from/to an application to/from the toolbar (Figure 6 – the toolbar application on the top of the track diagram). All the applications were also available as single applications. This approach allowed us also to have a more modular commercial approach.

![Figure 6 – the toolbar application on the top of the track diagram](image)

**New architecture**

To support all the changes described above we needed to develop and deploy a new, more structured architecture (Figure 7 – the current logical architecture).

![Figure 7 – the current logical architecture](image)

First of all we added two modules: a check-in and a check out modules to separate data
elaboration from data providers (check-in module) e applications (check-out module).

Then we upgraded the map matching module to a route matching module: the new system
was able to follow the route of each gps track and use this information to perform the
matching.

Last we moved some common business logic from applications to internal modules
(Post-fusion modules). This approached allowed for a more common logic (one
implementation for all instead of similar implementations distributed in different applications)
and for a better maintenance too.

PROJECT MANAGEMENT APPROACH

We distilled concepts from iterative methodology and from agile too. We concentrated on
continuous delivering of consistent steps. We allowed our end users to work with any single
delivery, any delivery had improved features and was consistent. We did not make a single
giant delivery of the full project.

Our approach implied some reworking but it was very positively accepted but our customers
that did not need to wait for a long time before being able to have new features. The
continuous usage of improved applications allowed also having frequent feedbacks too.
Worth adding that delivering a service and not a piece of software to be installed in customer
computers greatly facilitated this approach.

NEW ACTIVITIES AND RESEARCHES

Mobinet project

One of the most important research activity related to the work done is in Mobinet Project
[10]. Mobinet is an Eu R&D co-funded Integrated project in FP7. The project is implementing
a global multi-vendor business-to-business E-Marketplace and offering a uniform middleware
environment for the end-user, and an easy deployment and operation of services.
The WG is about selling and purchasing GPS data, evaluating their quality and profiling the
data to be sold/purchased. Infoblu is developing the data profiling service.
This service will receive in input a dataset of some days of gps data and will be able to extract
indicators such as

- number of positions per day and hour
- number of travels per hour and day
- distribution of travel durations per hour and day
- etc

Above indicators will be available both on the full dataset and in areas that are given in input
to the service. Areas are user-defined and can represent towns, regions etc. Area definition are
provided by service user along with the gps data set.

**TIM Big Data Challenge 2015**

Infoblu is partner of a Telecom Italia initiative named TIM Big Data Challenge [11]. TIM makes the largest heterogeneous Big Data dataset available to all those taking part, to allow them to use their imagination and skills to find the most interesting and innovative way by which to use it. The contest is open to individual professionals, academic or industrial entities and businesses. The most promising ideas will be rewarded at the end of the contest in an event to be held in Rome in the month of September, the Big Data Jam.

Big Data Challenge participants can choose either of the two tracks proposed to develop projects around the theme of Big Data for Competitiveness Boost.

The tracks are:

- **Academic Track**: dedicated to researchers and university students developing highly innovative solutions, albeit in an experimental state
- **Industrial Track**: for industrial professionals proposing more consolidated solutions, suitable for transfer to the market in the short- or medium-term.

Participants have two months within which to prepare their ideas. Of all ideas sent, a short list of the ten most promising (five per track), will be chosen. Of these, a selected panel will then reward the best idea for each track.

The data supplied as part of the Big Data Challenge dataset will be geo-referenced, anonymous, related to national territory, heterogeneous and regard the period March - April 2015.

The edition of the challenge was in 2014: 1,100 people took part from more than 20 different countries, submitting over a 100 design ideas.

Infoblu provided gps covering 7 Italian towns. The size of dataset for the challenge was some 70 GigaBytes.

**CONCLUSIONS**

FVDs/gps are available in large quantity in the last few years. The availability is increasing as long as the freshness while the price is decreasing.

FVDs had been used to create/verify traffic information and this is still a large market. Anyhow they can be used for building traffic monitoring application as well. The paper has presented some of them that were developed and used in an operational traffic control center in Italy with very positive feedbacks.

In addition to further developments in this context, we are exploring other segments using the same data. One of them is in the EU R&D Mobinet Project while another is represented by TIM Big Data Challenge.
References

[3] Tpeg™ TFP  
https://www.linkedin.com/groupItem?view=&item=594950779967853
[5] Are Bluetooth Probe Detectors Good?  
https://www.linkedin.com/groupItem?view=&item=594813711367650