Management of Traffic Situations for the Road Freight Transport

Ricardo Pinto¹, José Sousa², and Gabriel Pestana¹

¹ INESC/IST, 1000-029 Lisboa, Portugal
² ESRI-Porugal, 1600-131 Lisboa, Portugal

Abstract — The management of a freight’s daily activity is nowadays performed in a very competitive environment. In most cases the management is performed remotely, with the manager at the situation room monitoring ongoing occurrences and having to take operational decisions based on the information collected from or provided by field workers. This scenario allows carriers to enhance their business processes to a more efficient level with lower costs. Some of the strategies considered by market carriers as an efficient management operation consists of maximizing the goods being transported by a simple vehicle at one time. However, this is not a simple task, because traffic conditions change constantly thus causing delays. One way to improve business efficiency is through dynamic route planning for on-time route changes in case of any traffic occurrences that might compromise predefined service levels. Indeed, updated information about traffic conditions improves carrier business activities by providing them with the possibility to make real-time adjustments to transport routes and delivery plans. With this in consideration, the paper presents a methodology for a collaborative collection and distribution of road traffic data among fleet drivers using GPS and GPRS services. A Datex2 based on-road incident reporting scheme is presented. When confronted with traffic abnormalities, fleet drivers fill out an online Datex2 based incident form. Registered data is feed to a geographic data fusion, clustering and impact estimation unit that cross-references processed data with GPS-tracked vehicles and their respective routes. Drivers on course to traffic congestions are notified and route alternatives are suggested. In addition, drivers that do not change course receive an incident update request, so as to provide new insights on the development of the incident.

Index Terms — Fleet Collaboration Network, Datex2, Common Alerting Protocol, Data Fusion, Data Clustering, Incident Duration Estimation, Time-Series Analysis, Alternative Route Suggestion

I. INTRODUCTION

The road freight management has shown a regular increase over the last years. According to EUROSTAT, between 1996 and 2007 the quantity of transported goods has risen from 1.3 billion tkm (ton-kilometer) to 1.9 billion tkm, making the road transport the most used form of transportation, holding a 42% of the total of transported goods in Europe. For this to be possible, necessarily millions of transport vehicles roam the earth on a daily basis. It isn’t hard to picture that low transport fees and timely deliveries are fundamental in order to survive in this business. However, traffic conditions are somewhat volatile and unpredictable, sometimes compromising quality of service. This is a problem for freight management Organizations. Not only do they have to deal with unpredictability in their businesses, but must also account for profit losses due to delays, suboptimal vehicle usage and unnecessary fuel consumption. Thus said, current information about the state of traffic conditions improves carrier businesses by real time adjustments to transport routes and delivery plans. But there are two main problems with acquiring these data: (1) the first one is related to data sources. There are various sources that may have current traffic data related to a particular area. Nonetheless, most often these data do not follow any specific format at all. The ones that do follow come mainly from highway management organizations or other road management related organizations. Besides being proprietary data, these data are confined to a limited number of roads and extents. (2) The second problem is related to data quality and format. Provided traffic data may turn out not as useful as first thought of. Moreover, data integration issues are always a problem. Standards for describing road data are somewhat recent, being the Datex2 standard an up-to-date European reference for the interchange of road traffic data [1] and the UNETRANS a very thorough model for the organization of road related data, covering infrastructure, routes and dynamic data [2]. Consequently, various road management organizations have not yet adhered to a common traffic data model or communication protocol. However, over the last decade there has been given much attention to the study and design of Intelligent Transportation Systems, leading to the adoption of ITS in several countries, like the USA, Canada, Germany, Australia, United Kingdom and Japan [3]. We give special highlight to the United Kingdom and Japanese community for their continuous contributions to the field, as can be found in [4][5][6] among others. In this paper we focus solely on freight carriers and propose a non-third-party dependant methodology for a collaborative collection and distribution of road traffic data among fleet drivers. Concretely, we use a simplified subset of the Datex2 specification for the elaboration of an on-road real-time incident reporting scheme. When confronted with traffic abnormalities, fleet drivers fill out an online Datex2 based incident form. Transmitted data is feed to a geographic data fusion, clustering and impact estimation unit that cross-references processed data with GPS-tracked vehicles and respective routes. Drivers on course to traffic congestions receive a notice about the occurrence along with suggestions for route alternatives. In addition, drivers that do not change course for whatever reason, receive an incident update request, so as to provide new insights on the development of the incident. In these conditions, the more drivers join this collaboration, the more data about the current state of traffic on a particular country or region is gathered. Additionally, a
special time-series analysis is performed for the construction of a dynamic view of how traffic occurrences evolve over time. The remaining text is divided into the following sections: first we introduce the architecture of the system. Next we focus on the architectural components of the central server and describe how they work together. Finally we give special focus to the clustering component and the analysis component.

II. DATA COLLECTION AND ARCHITECTURE

The proposed system is composed of two main architectural components: a GPS and GPRS enabled equipment installed in each vehicle for sending and receiving traffic data, and a central server responsible for clustering, analysis and distribution of traffic information.

In this architecture, each driver is responsible for notifying traffic occurrences that are disturbing the normal flow of traffic. To accomplish this, they resort to an on-board installed equipment for a rapid, onsite report registration. Every report is sent to the central server, which in turn, is responsible for data processing and latter data distribution.

The server has the following main activities: (1) to group registered records related to the same traffic occurrence; (2) to determine how the impact of situations evolve over time; (3) to reference situations with location, time and seasonality; (4) to cross-reference traffic data with carrier vehicles for notification of drivers on course to active traffic situations; (5) and to send incident update requests to drivers on-route to occurrences.

As depicted in Fig. 1, the communication of data between vehicles and server flows both ways. Vehicles are only responsible for sending data about traffic situations, be it a previously described situation, or a new one. The server is responsible for two types of messages: incident notification messages and situation update requests.

For traffic data communication, each fleet vehicle has a GPS and GPRS enabled equipment with an interface for onsite incident reporting. The main purpose of this equipment is to facilitate a rapid report elaboration. With few clicks, a user is able to fully describe a traffic situation and send it to the server. In order to elaborate a traffic report schema as simple as possible we resorted to the Datex2 specification. Although being an enormous and very thorough specification, we were able to squeeze out a schema that consists of a simplification of a subset of the Datex2, shown here in Fig. 2. It is a rather simple model, and translates to a rapid report elaboration. The location of an occurrence is solely described by its geographic coordinates. This further simplifies the elaboration of incident reports but has its downfalls. We will elaborate more on this topic further ahead.

The central server is composed of five main components (view Fig. 3): (1) a GIS component responsible for reverse geocoding GPS coordinates and for best route calculation considering current post-processed traffic data; (2) a clustering component responsible for record grouping using an agglomerative hierarchical clustering algorithm and ontological record classification; (3) an analysis component responsible for incident duration estimation and time-series composition and analysis; (4) a business logic component that contains all processes related to driver communication and vehicle tracking; (5) and a database component where all data is stored.

![Figure 1 – General view of communication between carrier vehicles and central server.](image)

![Figure 2 – Schema of incident reports and update notices.](image)

![Figure 3 – Central Server internal architecture.](image)

This simplifies report elaboration because the user doesn’t need to describe the road he is in. Furthermore, it results in a
faster and more focused report elaboration, since the driver only needs to describe the actual occurrence. But as consequence, the transmitted data is insufficient for performing a road network aware data clustering algorithm. It would be preferable to know the address of the occurrence as well as the directionality along it. And when located on highways, even more data is needed. To say that a congestion has formed on a specific highway is at best ludicrous. The kilometric point of the incident is fundamental to determine its location. Hence a reverse geocoding process is needed in order to translate the geographic coordinates into network locations. Whereas geocoding consists of determining the geographic coordinate of a point along a road, reverse geocoding consists of the exact opposite. It involves determining the closest address to a given set of geographic coordinates. This works well for addresses, but fails in precision when the occurrences are located on highways. In these situations, reverse geocoding is not enough. One requires the kilometric point. But kilometric point calculation is not as straightforward as one might think. One must previously elaborate special geometric objects that result from a process called dynamic segmentation. These objects enable a runtime computation of routes and route properties. Each route computation comprises an iterative process that runs through a set of previously identified calibrated linear features, assembling preselected properties along the way. It is described by short contiguous geometric features geographically positioned in respect to a specific coordinate system or Datum, being each one associated with a set of attributes related to driving rules, such as directionality, impedance, turning, velocity limits and so forth [7][8].

In our case, we used the dynamic segmentation to create distance measures. The resulting object enables us to calculate the metric distance of any chosen point along the preprocessed route. One of the main advantages of dynamic segmentation is the abstraction it provides from the features it is composed of. Changes to the underlying network can be made at any time, being those immediately reflected on the route’s shape and properties. This hides the complexity of routes and encapsulates all route changes to one single place, making them innocuous to route users. With both reverse geocoding and dynamic segmentation it is possible to determine with a high degree of accuracy the network location of a geographic coordinate. Using these two tools together, the GIS component is capable of assigning a network related context to transmitted traffic reports, favoring subsequent clustering and analysis processes.

III. CLUSTERING COMPONENT

Considering multiple sources simultaneously feeding data to the system, leads us to reckon the possibility of more than one record being referent to the same one occurrence. In these conditions, clustering records makes sense, because it promotes data unity and additionally contributes to a better ascertainment of the location of the occurrences, their magnitude and also their duration.

The first challenge we face is how to handle a large amount of data distributed across a vast geographic terrain. We base our solution on the divide-and-conquer paradigm. First off, records are separated into different sets. Each set refers to one highway and specific driving direction along it, which results in pairs of sets for each individual highway. Then, for each set containing traffic data, a grouping methodology is applied. This methodology has three main stages: (1) on the first stage, we separate the records by applying a variant of an agglomerative hierarchical clustering algorithm; (2) for each resulting group, an ontological analysis is made so as to determine which records are really related to the same occurrence; (3) finally, for each subgroup resulting from the previous analysis, two calculations are made: first the center of mass of traffic records is calculated - where mass represents the reliability of a record - so as to estimate the real location of the occurrence; second, the linear traffic extent of the event is estimated, for impact analysis and duration estimation. Each one of these stages will be subsequently described.

A. Clustering Algorithm

The first stage consists of organizing the georeferenced records into separate groups based on their geographic proximity. An agglomerative hierarchical clustering algorithm is applied. When the algorithm starts, each record consists of a separate cluster. On each iteration, the two closest clusters are merged. The centroid of a cluster is calculated through the mean of the kilometric points of all records contained in the cluster. The algorithm stops when there is only one cluster left. Formally, a set of clustered records C, has a relative mean position \( P_c \) defined by:

\[
P_c = \frac{\sum_{i=0}^{n} p_i}{n}, \quad p_i \in C
\]

Equation 1

where \( p_i \) represents the kilometric point of the \( i \)th element contained in the set \( C \). Another way to represent the merged clusters is through a dendogram (view Fig. 4). Each leaf node represents an individual record, whereas the remaining nodes represent clustered records, being the root node the only cluster left after the final iteration. A dendogram is a typical result of an agglomerative hierarchical clustering algorithm. After the dendogram is created, it needs to be cut into subtrees, so that each subtree has only geographically close leafs. Being the adjective “close” a somewhat subjective description of proximity, it was determined that leafs geographically close should have a standard kilometric point deviation below a certain specific threshold. This means that each node in the dendogram stores the standard deviation of all leafs of the subtree it is root. In this sense we have:

\[
S = \sqrt{\frac{1}{n-1} \sum_{i=0}^{n} (p_i - P_c)^2}
\]

Equation 2

As follows, standard deviation tends to increase from bottom nodes to the root node of the dendogram. With a depth first
search starting from the root node, when reaching a node with a standard deviation below the previously specified threshold, the node is considered the root of a subtree with geographically close leafs, thus being the whole subtree cut from the dendogram (view Fig. 5) and it’s leafs stored in an individual group for subsequent analysis.

Figure 4 – Leaf nodes in (a) represent traffic occurrences. Nodes are grouped according to their geographic proximity. In (b), ellipses illustrate the grouped occurrences depicted in the dendogram.

Figure 5 – A subtree is cut when its root node has a standard deviation below a specific threshold. Further analysis can thus be confined to each resulting subgroup of records. Root nodes are depicted in blue in (a), (b) illustrates the grouped occurrences depicted in the dendogram.

B. Ontological Analysis

The second stage consists of applying an ontological analysis to each one of the identified groups so as to determine which records are really related to the same occurrence. A depiction of the ontological model is shown in Fig. 6. First off a situation is set as the most generic type of traffic occurrence that can take place. A situation is described by for main parts: (1) the first one describes the spatial and temporal location of the situation, its estimated duration, and impact on the road. It is a rather different part from the remaining three, because the data it contains may change with time, due to the appreciations made by the analysis component. Location is the least probable information to change, but it may nonetheless. The other three parts describe the causes, aggravations and consequences of the situation. (2) Causes can be various types of accidents, congestions and road activities. They are named as causes of congestions because they are considered as the triggers of the congestion. (3) Aggravations are composed of driving conditions and current climate situation. A driving condition can describe various aspects, like quality of the road, ice or water patches on the road, visibility, and current congestion. The climate situation describes the current climate conditions. (4) Consequences are, as the name says, the consequences of the causes and aggravations, described as various types of congestions, which influence traffic conditions, and give rise to the current observed situation. A consequence is backpropagated, through the road, by a given cause, they may or may not be created or worsened by a given aggravation.

The ontological classification algorithm has the following steps:
(a) First, divide the records into two separate groups based on the directionality of the events and store each individual record in a separate set.
(b) For both groups, determine which records represent generalizations of others. For those who are, determine if proximity and time span are below specified values. In case they are, merge both sets together.
(c) For each traffic flow record, verify, to a previously defined extent, if accident records exist further down the road. If records are found and all of them are on the same set, group them together. If not all of them belong to the same set, group the traffic flow record with the set which has the most recent records. For all traffic flow records that remained ungrouped to accident records, group them together.

C. Location Estimation

The third stage consists of determining, for each set resulting from the ontological analysis, the real location of the occurrence and also the estimated congestion extent along the road. There are three types of possible sets: sets with traffic congestion records only, sets with congestion and accident records, and sets solely with accident records. Each one of those is treated in a slightly different way.
For sets with congestion records only, the goal is to determine the real extent of the congestion, using the reported records as reference. Assuming that communicated records tend to have a higher frequency near the center of the traffic congestion, we presume that the distribution could be approximated to a normal distribution. Moreover, since the number of records should usually be rather small, the T-Student distribution presents itself as an even better approximation.

Based on this assumption, we use the T-Student Distribution to calculate the confidence intervals of 90%, 80% and 70%. In this case, confidence intervals are used to answer the following question: assuming a specific set of records, calculate a kilometric interval that determines, with a confidence of X, where other congestion records related to this event might fall. Each confidence interval gives us the left and right side confidence limits, which in this case, represent kilometric points. Formally we have:

\[
\text{ConfLim}_{1.2} = P_e \pm \frac{s}{\sqrt{n}}
\]

Equation 3

where \( P_e \) corresponds to the value from the T-Student table that is retrieved based on the provided confidence interval and number of degrees of freedom. Pairs of limits translate to the start and end of congestions, respectively (view Fig. 7). These extents - stacked bottom up from 70% to 90% - give us a visual and kilometric perception of three levels of congestion: the 90% confidence interval has the largest extent and represents a low level of congestion, as opposed to the 70% interval, holding the shortest extent and a high level of congestion.

For sets with accident records only, the goal here is to estimate the actual location of the accident. Taking into account that each record has a specific reliability, we apply the center of mass theorem using the records reliabilities as their individual masses (view Fig. 8). Formally:

\[
P_r = \frac{\sum_{i=0}^{n} p_i r_i}{\sum_{i=0}^{n} r_i}
\]

Equation 4

where \( p_i \) represents the individual reliability of the ith record. The result is a kilometric point that best describes where the accident should be, based on the reliability weights of the reported records.

The last type of set is the one that contains congestion and accident records. The process applied here is similar to the ones above: first divide the set into two subsets, on for congestion records and one for traffic records. For the congestion records calculate the confidence intervals as described above, and for the accidents records the center of mass. Finally, the kilometric end limits of all confidence intervals are readjusted to end on top of the calculated center of mass of the accident records (view Fig. 9).

Figure 7 – Confidence Intervals are calculated for the kilometric points of three congestion reports. Depicted interval distances are only representative.

Figure 8 – The triangle in (a) represents the center of mass of the grouped records. The reliability of the new node is not yet defined. As illustrated in (b), the center of mass is closer to node 2 because of its higher reliability.

Figure 9 – End limits are readjusted to meet with the calculated center of mass of the accident records. Accident records are omitted for simplicity.

IV. DYNAMIC SEGMENTATION AND RECORD GEOREFERENTIATION

Taking into account the normalized traffic data format, transmitted data may come with three distinct description formats for their spatial locations: they may have only a textual description for their location; they may already be spatially referenced with an assigned geographic location; or they may have both. Each format needs to be treated in a specific manner. Highway traffic data with only a textual
description for its location needs to be geographically pinpointed, a process called geocoding or georeferentiation. Oppositely, traffic data accompanied only with geographic coordinates needs the inverse treatment, formally named as reverse geocoding. This process consists of determining the textual description that best describes the location of a geographic coordinate, typically by determining the closest address to the given coordinate, or in our case, by determining its kilometric point along the highway. The third format needs only to be checked for concordance between geographic coordinates and textual description. In order to accomplish any of these tasks, geographically located geometric and route data of the highways in question needs to be readily available. Geographically located geometric data of road networks is available for purchase by various vendors. These data consist typically of short contiguous geometric features geographically positioned in respect to a specific coordinate system or Datum, being each one associated with a set of attributes related to driving rules and conditions, such as directionality, impedance, turning, velocity limits and so forth [5][6]. Highways are described in a similar fashion too. However, these data only goes so far. In respect to determining a kilometric point of a specific location along a highway represented by geometric features, the available data, as it stands, does not suffice. The route that fully describes the highway must be determined, and its shape and location identified (one route for each driving direction along it). A process called dynamic segmentation solves this problem, and provides other advantages too.

Dynamic segmentation consists of a runtime computation of routes and route properties. Each route construction comprises an iterative process that runs through a set of previously identified calibrated linear features, assembling preselected properties along the way. In our case, distance measures were calculated (view Fig. 10). The resulting route enabled us to calculate the metric distance of any chosen point along a preprocessed route. One of the main advantages of dynamic segmentation consists of the abstraction it provides from the features it is composed of. Changes to the underlying network may be made at any time, being those immediately reflected on the route’s shape and properties. This hides the complexity of a route and encapsulates all route changes to one single place, making them innocuous to all route users.

V. ANALYSIS COMPONENT

The analysis component is responsible for the following tasks: (1) to assess the lifetime of traffic reports; (2) to adjust reports’ reliability as their lifetime shortens; (3) to discard outdated traffic reports from the system; (4) and to take periodic snapshots of the global state of traffic.

This component runs parallel to the clustering component. While the latter is busy clustering records, the analysis component is responsible for removing outdated records from the system. In this manner registered traffic data is in a constant change as new records enter the system and old ones leave. Resulting clustered data also changes, leading to new clustered incident locations and congestion extents. In addition, the analysis component takes periodic snapshots of the current clustered traffic. As time passes and more data is gathered, time and location related patterns begin to appear. Specific road extents appear congested on rush hours. Others are more prone to accidents on busy Fridays. This is the main goal of the analysis component: to provide a geographic and time aware forecasting service for congestion prediction and best route calculation support.

In order to achieve this, the first step to take is to define how the lifetime of traffic reports is determined. As said before, reports have a specific duration, after which they are discarded from the system. This duration should be associated with the nature of the traffic reports in question. Here, we considered five categories of traffic reports: “Climate Situation”, “Incident – Accident/Obstruction”, “Activities”, “Driving Conditions” and “Traffic Status”. Each one of these should be related to different lifetimes. Moreover, the “Overall Impact” description should also be considered for record lifetime calculation. Geographic location of traffic reports is important too. Accidents or obstructions located in certain areas may prove to prevail longer due to lack of appropriate incident management organizations in the near surroundings, or poor local incident management plans and policies, among others. Finally, the time of day, weekday and month also influence incident probability as well as clearance time. Fridays for example, are especially hazardous days. According to EUROSTAT accidents tend to occur more on that specific weekday. Holidays may also account for less staff available on incident management organizations, leading to a slower service. Accidents occurring at night may take longer to be cleared, but may not influence traffic flow at all, because of the reduced traffic flow observed at night. Taking all these factors into account, we established that an incident’s report lifetime should be defined by the traffic situation described, the appended climate situation, the estimated overall impact, the location where the incident is found, and the time of day, weekday and month. Location takes into account the county where the incident took place, and differentiates between roads and highways. The time of
day is divided into the following sections: morning, noon, afternoon, evening and night. Notice that whereas the day is divided into four sections, the night has only one section. We chose these divisions based on an empirical assumption that traffic flow is more prone to fluctuate more during the day and between these sections than during the night, a period of substantial traffic flow reduction. Location aside, if we calculate the number of possible combinations of traffic reports, we get above fifteen hundred results. In order to make a statistical estimation of the duration of a specific type of occurrence at a specific location, we must first have a reasonably large set of reports related to that situation on that location, which, when put together, provide an insight on how long, on average, that type of incident lasts. This appears to be very impractical, considering the multitude of possible report combinations versus the projected low number of transmitted reports. A statistical analysis would only be possible after a very long period of time.

However, a better look at the situation shows that this is not that a problem at all. Our goal is not to determine, to the minute, how long different types of incident last. We are much more interested in finding out patterns of when and where congestions form, and accidents happen. Still, for this to be possible, we must have an idea of how long, more or less, an incident lasts.

Fig. 11 depicts two very simple examples of how congestions may form. The first one shows how an accident may create a bottleneck sufficient enough for congestion formation. The second shows a typical rush hour situation on which the maximum available network flow is surpassed by the traffic flow, thus leading to congestion.

This clearly is a simplification of how accidents and rush hours influence traffic flow. We show these two situations to point out that congestions will form only when traffic flow is higher than maximum available road flow. Congestion formation tends first to increase, until it reaches a maximum and then decreases again. It is only during these two stages that traffic reports will be sent. Finding out on which one of these stages each traffic report stands, contributes to a better assessment of how long an incident is supposed to last.

Based on this diagram, and bearing in mind how incidents and update notices are reported by drivers, and taking also into account how the clustering component groups records, we can layout the following timelines shown in Fig. 12. This diagram shows various timelines with clustered sets of traffic reports that were ordered chronologically. Each set shows an example of how increase and decrease stages may or not be inferred from ordered records. Assuming that increase and decrease stages are somewhat similar in duration we can estimate an incident lifetime solely knowing its increase stage or decrease stage. This is advantageous, because only in rare situations will there be enough data available for precise incident duration calculation.

In addition, the “Overall Impact” described in each report is correlated with the location of the report along the timeline as well as with the total estimated duration of the incident. Collecting this information proves usable for further inference about traffic records. With these data we are able to better estimate on which stage traffic records are standing when they reach the system. Having established an approach for record lifetime calculation, we can pass to the second step, which consists of reducing the reliability of a record as time passes. Here, reliability can be seen as the correlation between the record and the actual reality.

![Traffic Flow Diagram](image)

**Figure 11 – Simple description of how changes to traffic flow and maximum road flow may create traffic congestions**

We assume that each record has full correlation when it enters the system. As previously said, the reliability of clustered records consist of the total sum of the reliabilities of its elements. But as time passes, reliability of individual records should decrease as their lifetime shortens. As follows, old records, that are almost ready to be removed from the system, should have almost no reliability.

Changing the reliability of a record has consequences on the clustering results. This is useful for maintaining a similarity between traffic data in the system and the actual reality they intend to describe. Hence, reliability, or weight of records, decreases over time. It remains only to determine how it should actually decrease. The idea is to map the lifetime of the function to the domain, and the reliability to the range, obtaining:

\[
  f: [0, L] \rightarrow [0, 1]
\]

where \( L = \text{"Record Lifetime"} \)

**Equation 5**

The simplest way to do this is using a linear function which traverses the points \( X_1 = (0, 1) \) and \( X_2 = (L, 0) \). In our case we decided to base our solution on the logistic function, the most common sigmoid curve. This function models an S-curve, starting with an exponential growth until saturation begins, and growth slows down, eventually stopping. It fits our needs well, because we want records to maintain high reliability when they enter the system, exhibiting thereafter a linear decline of reliability until reaching the final stage with very low reliability.

But for determining the logistic function needed for our specific purpose we need to start with the Generalized Logistic Function, also known as the Richards’ Curve, which enables the modeling of different types of logistic functions:

\[
  Y(t) = A + \frac{H - A}{1 + Te^{-B(t-M)^2}}^p
\]

**Equation 6**
where $Y =$ “Weight”, $A =$ “Lower Assymptote”, $H =$ “Upper Assymptote”, $B =$ “Growth Rate”, $T =$ “Center Dislocation” and $M =$ “Maximum Growth Center”.

Furthermore, using the timeframes associated to each record it is possible to selectively choose all previously registered records associated to a specific timeframe, so as to provide a view of what tends to happen on that selected frame. This is useful for the prediction of what is most probable to happen and hence what should be taken into account when choosing a route.

Fig. 13 shows the general behavior of the analysis component. It contains two internal components called ‘Identifiers’ that are executed periodically. The first one, called ‘Identifier of Ended Situations’, identifies all sets of grouped records that contain no longer active records, being these sets considered as terminated, and copied to the star model in the OLAP cube. The second identifier is occupied with the active situations, gathering them periodically, and passing them to the data analysis pipeline identified in the figure. In the pipeline, the records are ordered chronologically, then congestion fases are identified, followed by similar situations querying, estimation of duration and impact, and finally the readjustment of deviations.

![Figure 12](image12.png)

**Figure 12** – Various depictions of how ordering notices may bring insight on the duration of incidents

This general equation allows us to create various logistic functions. After some attempts we got to this final equation:

$$Y(t) = \frac{1}{1 + e^{\frac{X^2 (-L+1)}{2}}}$$

where $A = 0, H = 1, T = 1, B = -\pi^2, M = \frac{L + H}{2}$

and $L =$ “Record Lifetime”

Equation 7

The analysis component uses this function to periodically update the respective weights of each record. In turn, this influences clustering results, because the clustering component uses the weights of records to decide how records are clustered. Finally, the analysis component is responsible for taking periodic snapshots of current clustered data. These data will provide a dynamic view of how incidents evolve over time.

![Figure 13](image13.png)

**Figure 13** – Main internal components and behavior of the analysis component.

VI. CONTROL ROOM-SERVER-VEHICLE COMMUNICATION

Vehicle tracking is fundamental. Each vehicle sends a periodic message to the server containing the identification of the vehicle and current GPS coordinates. The server in turn, has stored the delivery route and plan of every vehicle. Periodically, the server cross-references the route of every vehicle with the current clustered traffic data. For each vehicle on-route to a traffic incident, the server estimates the drive-time from the current vehicle position to the incident. If the drive-time is shorter than the lifetime of the incident, the server sends a notice to the operators in the control room of the freight carrier, informing them about the incoming traffic situation, a prediction of how the situation will be when the vehicle reaches the occurrence, and a list of route alternatives from which the operators can choose. The operator responsible for vehicle coordination may then choose to change route, selecting the route in the touch interface, which in sequence sends a message to the server telling it to update that vehicle’s route. The protocol used for the elaboration of notices sent to the operators is based on the Common Alerting Protocol [9].

Furthermore, when the central server determines that a vehicle
is close and on-route to an incident, it requests the driver for a situation update. If the driver chooses to comply, it selects so from the interface and fills out an incident update form. In this manner, the system promotes a situation update strategy.

VII. CASE STUDY

The study performed consists of an elaboration of a hypothetical scenario on one of the main Portuguese highways. It connects the two most important Portuguese cities, Lisbon and Porto. This scenario contemplates various incidents along the highway, and a high number of incident report notices. For the preparation of the scenario, first control data of incidents and congestions was set up. Then, various incident reports were distributed over the control data. The distribution of records contemplated a certain degree of error, where reliability of a record was correlated with notice similarity and proximity to control data. But this also had some error introduced. This means that some notices with high reliability were just wrong. The simulated data was communicated to the system and the clustering component was activated. Control, simulation and result data are displayed in Fig. 14. It is noticeable a similarity between clustering results and control data, despite the lack of complete incident description and further error introduction. Fig. 15 shows how the analysis component may provide further insight on traffic behaviour. Yellow circles highlight locations where previous incident reports were transmitted, and blue circles highlight areas with previous congestion reports. This information may prove useful for carrier control centers and drivers, because it can be used as a prediction of locations where incidents and congestions are most probable to happen, and thus should be avoided.

Routing capabilities were easily integrated in the system. Each service call requires the specification of at least two stop points (start and finish), and an optional list of road barriers. The result of the service call consists of a route that traverses through all the specified points, avoiding the specified barriers, if located on the best route path. The provided system service enables best route estimation, between two points, taking into account the collected traffic events. For each service request, first off, a list of all clustered traffic events is compiled. Then, for each clustered traffic event, its distance to the route start point is measured. The resulting list consists of all traffic events that are located close enough to be reachable from the route start point. This list is recalculated before every route request, and sent along as parameter to the remote routing service call. A routing example is provided in Fig. 16.

All data was integrated through SQL Server 2005 and ESRI ArcGIS Desktop 9.3.1. The resulting data was published through ArcGIS Server 9.3.1, and consumed by a Silverlight web application. Feature identification and routing tasks were added. Routing took into account clustered traffic data for best route estimation and used ESRI European routing services for best route calculation.

VIII. CONCLUSIONS

An approach for a collaborative collection and distribution of road traffic data among fleet drivers using GPS and GPRS services was presented. It consists of an on-road and on-site incident reporting strategy that uses the drivers of fleet carriers as sensors for road traffic data. Each fleet driver contributes to the common knowledge of current traffic state. The Datex2 protocol was used as reference for the elaboration of the traffic report schema. Collected data is sent to a central server that
clusters, analyses and distributes the data. Records were grouped based on an agglomerative hierarchical clustering algorithm, followed by an ontological classification algorithm. Incident location used reverse geocoding and dynamic segmentation processes for location calculation, followed by center of mass theorem and confidence intervals for incident centroid and impact extent estimation. Finally, a location and time aware incident composition and analysis process was proposed for incident classification support and time-series analysis.

ACKNOWLEDGMENT
R. Pinto thanks ESRI-Portugal for their active support on the elaboration of this paper and thanks his Master’s Supervisor, G. Pestana, for his very insightful reviews and comments that helped shape this document.

REFERENCES