traffic safety viewpoint, road network

Zoltán FAZEKAS*, Zsolt BIRÓ, Péter GÁSPÁR
Computer and Automation Research Institute (MTA SZTAKI)
Kende u. 13-17, H-1111 Budapest, Hungary
**Corresponding author*. E-mail: zoltan.fazekas@sztaki.hu

STATISTICAL ANALYSIS OF DATA DESCRIBING THE RELATIONSHIP BETWEEN DRIVER, TRUCK AND CHARACTERISTICS OF THE ROAD

Summary. Road vehicles equipped with appropriate measurement equipment, computing facilities, data storage and data communication capabilities can be considered valuable data sources for the description and quantitative characterisation of road traffic. The data obtained from these vehicles provide valuable direct and indirect information pertaining to traffic states and various aspects of traffic safety in respect of the analysed road network. In this study, trucks' abrupt braking events, detected by the trucks' on-board safety protection units, were analysed. The road locations of the detected abrupt braking events can be characterised by a number of features ranging from the specific traffic regulations (e.g., speed limits) in force to the socio-cultural environment of the location. The abrupt braking data evidence was used for identification and description of non-trivial interactions of drivers, trucks and roads. Some of the more interesting results and conclusions of the experiments are reported herein.

ANALIZA STATYSTYCZNA DANYCH OPISUJĄCYCH RELACJE MIĘDZY KIEROWCĄ, SAMOCHODEM CIĘŻAROWYM I CHARAKTERYSTYKĄ DROGI

Streszczenie. Pojazdy wyposażone w odpowiednie urządzenia pomiarowe, systemy przetwarzania numerycznego, pamięci do przechowywania danych oraz pokładowe układy do teletransmisji tych danych mogą być cennym źródłem dla konstruowania charakterystyk przemieszczania się pojazdów w specyficznych warunkach ruchu drogowego. Dane uzyskane z takich zapisów dostarczają nam cennej ilustracji pośredniej i bezpośredniej dla opisu stanów ruchu drogowego w aspekcie przepisów bezpieczeństwa ruchu drogowego; z uwzględnieniem specyfiki analizowanego odcinka sieci drogowej. W przedstawianych badaniach oceniano elementy bezpieczeństwa ruchu małych samochodów ciężarowych na podstawie danych z pokładowych układów rejestrujących, dla stanów gwałtownego hamowania, gdy aktywuje się system ABS. Miejsca rejestracji danych na szczególnie ważnych odcinkach drogi zostały odpowiednio odniesione do przepisów bezpieczeństwa ruchu drogowego wraz z elementami otoczenia tych odcinków dróg. Ewidencja ruchu zawiera szczegółowy opis charakterystyki jazdy na charakterystycznych odcinkach drogi z uwzględnieniem elementów ekonomiki eksploatacji pojazdów. Akcje gwałtownego hamowania pojazdów wykorzystano do identyfikacji i opisu szczególnej reakcji kierowców w odniesieniu do rodzaju pojazdu oraz charakterystyki jezdni. Wybrane elementy tych interesujących badań wraz z dyskusją zarejestrowanych danych zostały zaprezentowane w niniejszym raporcie.

1. INTRODUCTION

Vehicular emergencies, such as abrupt braking events, constitute a small, but important slice of road transport actions and events, a slice that deserves special attention from transport and safety engineers, and statisticians. Due to their size and weight and the damage they can cause, such an attention is particularly well founded for trucks.

A number of circumstances shape the spatiotemporal character of road traffic [11]; some of these contribute significantly to the probability of road accidents. The circumstances normally considered in this respect include season, time of the day, weather, weekday/weekend/ holiday, quality of road surface. Comprehensive research studies analyze the causes, the fatalities, and the locations of the road accidents, as well as, their perceptible trends, and estimate their economic consequences [10].

Apart from analyzing actual crash data, traffic safety experts use various modelling techniques to find out how dangerous certain road sections and road facilities are. Traffic volume-based accident prediction formulae, for example, are often used for this purpose [8]. As a more traceable, easily adjustable and analyzable alternative, various surrogate safety measures computed from the microscopic traffic models were proposed [4] and has been used in road safety assessment ever since [5]. Such measures can be calculated either from measurement data, or from data generated in simulation runs. As abrupt braking events frequently co-occur with traffic incidents used as input by Surrogate Safety Assessment Model (SSAM) and methodology [4, 5], it seems to be worthwhile to work out practical ways for utilising abrupt braking data in this manner.

Trucks of a transportation fleet can be seen as probe-vehicles moving in road traffic. Data received from them provide valuable direct and indirect traffic information in respect of the covered area. Abrupt braking events initiated by the drivers, for example, can be considered as symptoms of traffic safety problems over the covered routes.

Some results achieved in one of the subprojects of the project acknowledged later on in the paper are presented herein. The subproject concerned was aiming at carrying out various statistical analyses on the trajectory; braking and inertial data collected via data communication systems from trucks' safety systems. The research question posed is whether abrupt braking events could form a solid basis of a surrogate safety assessment method.

2. THE TRAJECTORY AND BRAKING DATA SET

Trajectory, braking and other inertial measurement data was collected from a fleet of trucks for some weeks and a trajectory and braking database (TBDB) was set up for research purposes. The details of the data gathering and the structure of the database are given in [1].

The data records in TBDB represent measurement and status data gathered and/or produced by on-board vehicle safety systems in respect of critical truck manoeuvres and during respective corrective actions initiated by these systems. The signals available onboard are processed effectively and efficiently by a number of electronic data processing elements. On-board safety systems normally available on trucks include anti-lock braking system (ABS), electronic stability program ESP, and roll-over prevention mechanism (ROP). These safety systems can detect manoeuvres like abrupt braking events (i.e., braking events resulting in decelerations exceeding some predefined level), and sudden steering (i.e., steering resulting in higher than usual yaw angular acceleration) that are critical for the controllability and for the stability of the trucks.

The trajectory and braking data stored in TBDB was drawn on in an earlier paper to look into the statistical relationship of the vehicle velocity just before abrupt braking events, the deceleration intended and the type of the environment [1]. In another study, truck activities were looked at [2]. The spatiotemporal characteristics of the truck trajectories were used to locate and identify the low and close-to-zero speed manoeuvres, such as loading/unloading a truck, precise parking at a distribution hub and refuelling at a petrol station, manoeuvres which are critical from a shipment and vehicle security point of view.

Using the data stored in TBDB, some non-trivial interactions between drivers, trucks, roads and their surroundings can be identified and characterised. Several statistical analyses and methods can be considered for the purpose.

2.1. Surroundings of vehicular emergency locations

The vehicular emergencies, such as abrupt braking events, typically happen on roads and roads themselves are built within some environments. These environments can be characterised with features, some of the features being static, others changing over time.

In case of quasi-static features, the change is slow, while for dynamic features, the change is clearly noticeable. The road locations where vehicular emergencies occur can be characterised among others by the number and types of road facilities in the vicinity (static), the specific traffic regulations imposed on the surrounding area (quasi-static), the traffic patterns formed under these conditions (quasi-static), the socio-cultural and economic features of the area (quasi-static) and the actual traffic (dynamic).

2.2. Visualisation of trajectories, abrupt braking events and surroundings

With up-to-date publicly accessible geographical/navigational databases and visualisation programs, such as Google Earth and Google Maps, experimentation with trajectory data has become fairly straightforward. The data records comprising the actual geographical positions and braking measurements need to be first converted into a Keyhole Markup Language (KML) representation.

A KML file, or indeed several of those, can be opened by the Google Earth program. Trajectory represented in the file can be displayed. An example is shown in Fig. 1. Locations (e.g., assignment starting points and destinations, braking locations) are displayed as shown in Figs. 1 and 2.

Using Google Earth program, one can zoom onto the individual braking locations; furthermore, one can follow a given truck in a virtual manner. Another useful facility of the program is StreetView. Photos are generated for the particular location from Google's 360º view image database, e.g., for a braking location.

Fig. 1. A short trajectory section and the location of an abrupt braking – marked by an ABS event – are displayed with Google Earth

Rys. 1. Trajektoria śladu gwałtownego hamowania oraz lokalizacja miejsca zadziałania mechanizmu ABS badanego pojazdu ciężarowego na mapie Google

2.3. Embedding TBDB into a wider data context

In order to fully exploit the traffic and traffic safety information incorporated in TBDB, it should be immersed in a wider business data context. In line with this approach, TBDB's can be and should be embedded in freight transportation information systems (FTIS), or even in supply chain management systems (SCMS). FTIS's are used from 1960's, while the more sophisticated multi-subsystem SCMS are more recent tools used by the leading freight companies. SCMS comprise subsystems ranging from demand forecasting subsystem to warehousing subsystem, and from requirements planning to customer services [6]. SCMS's are normally 'ab ovo' enhanced by geographical, road, infrastructure, business and community databases and state-of-the-art visualisation programs, or can be at least easily augmented by them.

Fig. 2. Destinations of truck assignments near Budapest, Hungary displayed with Google Earth
Rys. 2. Miejsca z zaznaczeniem pojazdów ciężarowych na mapie Google w okolicach Budapesztu

TBDB's embedded into SCMS's hold and can present a wealth of information – including traffic safety information – about the entities involved the freight transport activity of the company considered.

For example, relationships between the entities listed below are normally recorded in SCMS's: drivers (characterised by their name, age, gender, driving experience, etc.), vehicles (characterised by their make, dimensions, nominal and actual weight, speed, acceleration, heading direction, etc.), transportation assignments (characterised by start location, destination, planned route, transferred goods, etc.).

While entities listed below appear in the TBDB itself: vehicle trajectories (i.e., temporal sequences of geodesic positions), vehicular emergencies (characterised by the type of the actual emergency and the measured values of various internal signals, road location).

Freight transport entities including the ones shown below can be imported into the freight transportation information system from external databases and information systems; roads and road facilities (characterised by their location, geometry, number of lanes, usual quality of the road surface and its seasonal/weather induced changes, etc.), surroundings of road locations (including season and weather driven changes, socio-cultural character), weather conditions along the route of the vehicles.

For example, the abrupt braking event shown in Fig. 1 was detected by the truck's ABS. After a viewing of the location using Google Earth and StreetView, the braking location was classified as having occurred on a motorway in residential surroundings. The curvature of the road was classified as slightly bending. This event and its features appear as a dot in Fig. 3. The data shown there has been aggregated from several trucks.

Using the TBDB embedded into SCMS, one can answer queries like: what is the average actual cargo weight during a particular manoeuvre (e.g., lane changing, turning) that resulted in a particular vehicular emergency (e.g., a situation threatening with roll over of the vehicle) for individual trucks.

Analysing this query, one finds that the actual cargo weight is an attribute of the vehicle – i.e., a single data item – readily available in TBDB, while the manoeuvre needs some elaborate processing.

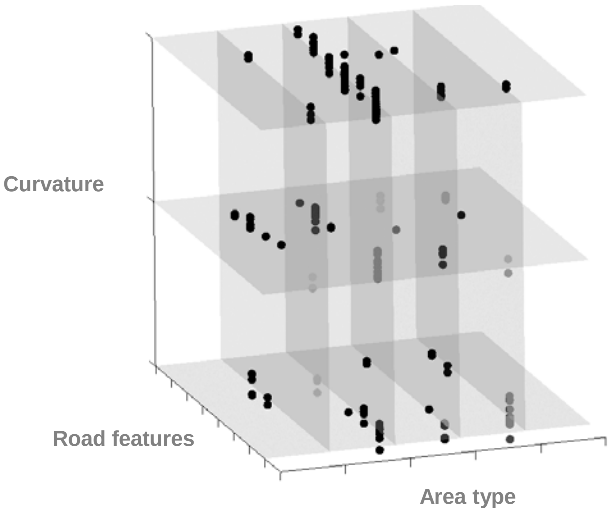

Fig. 3. Trucks' abrupt braking events in the feature space of area type (rural, business/industrial, residential, urban), curvature of truck trajectory (straight, bending, turning) and road location features (motorway, motorway entry/exit/junction, road entry/exit/junction, parking lot, parking lot nearby, junction nearby, test course)
Rys. 3. Awaryjne hamowanie ciężarówki w przestrzeni cech, typu: powierzchnia, krzywizna trajektorii jazdy i funkcje lokalizacyjne

Depending on the concrete datasets available and used for the purpose, this processing can be simple and straightforward, e.g., based solely on mechanical characterisation of the vehicle movement, that is, on some subset of positional, speed acceleration, angular, angular velocity and angular acceleration data, which results in a rather simplistic semantic view of manoeuvres, detecting manoeuvres neglecting the wider context of the road structure and surrounding traffic.

Alternatively, the processing of the query can be complex and more demanding, if data on road geometry, geographical location, socio-cultural environment, weather circumstances, aggregate, stratified, or even video-based local traffic information are considered. The importance of placing trajectory and braking data in a wider traffic context – e.g., augmenting it with omni-directional video sequence – was covered in [3]; the arrangement is sketched in Fig. 4.

Fig. 4. An omni-directional camera arrangement for traffic safety measurements
Rys. 4. Układ aparatu do wielokierunkowego pomiaru bezpieczeństwa ruchu drogowego

To provide traffic context for manoeuvres, firstly features corresponding to road markings, traffic signs and vehicles are extracted from the omni-directional image data. Based on these features, and considering the road surface model and the sensor model, the outliers are removed. Then a lane – compatible with lane model used – is fitted on the features (e.g., on blobs) detected. Then taking ego-vehicle dynamics into account, tracking is carried out with respect to the road and after proper filtering to other vehicles. Finally, the surrogate safety measures are calculated based on these tracking results.

A manoeuvre comprising several driving actions can be interpreted more profoundly if facts about the location are known. This aspect of the wider context is dealt with in more detail in [2]. Truck manoeuvres can be advantageously described using syntactic approaches [2, 9].

2.4. Statistical analysis and evaluation

The statistical methods used for analysing data from the TBDB range from the customary methods used for computing distributions and probability density functions to mathematically more demanding methods.

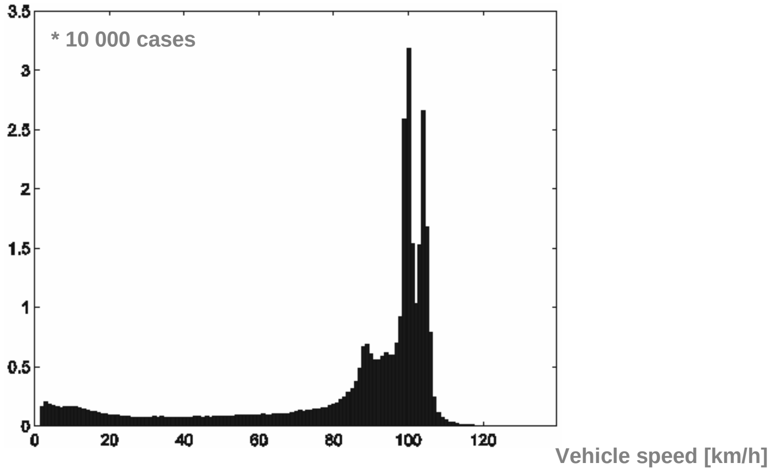

Fig. 5. Distribution of truck speed computed from non-emergency samples
Rys. 5. Histogram prędkości jazdy samochodów ciężarowych z próbek nie krytycznych

The distribution analyses were carried out to explore the how the truck drivers, who are the firsts to detect traffic emergencies and initiate driving actions meant to prevent them turning into accidents, whether these actions judged viable by the built-in vehicular safety equipments, or not, the trucks, which can be empty, partially or fully loaded, and the roads with their different surroundings and traffic affected each other at the time of data gathering; whether some of the tendencies and interactions found in the TBDB have more general relevance.

Analysing TBDB data, it was observed, for example, that the activities concerning ESP or ROP increase when the driver is driving an unloaded vehicle in urban areas. Also, the expectation that the same road sections have different risks depending on the time of day or the weather conditions has been confirmed.

In Fig. 5, vehicular speed histogram for the time-triggered (i.e., not emergency) samples from TBDB is shown. As it is apparent from the diagram, the trucks involved in the data collection were cruising mostly at the speed of 80 km/h or faster. The histogram pertains to all the roads covered by the trucks including motorways, interstate and intercity main roads, urban roads, streets, rural roads, etc. However, speed values customary to the long-haul transport (i.e., the trucks' usual cruising speed on motorways and main roads) are more frequent. Even so, some of the trucks spent – and usually spend – considerable amount of time on the roads going through urban, residential, business and industrial areas characterised with lower speed limits and more frequent traffic necessitated stops (e.g., near junctions). This is indicated by the histogram's local maxima in the 0 – 20 km/h speed range.

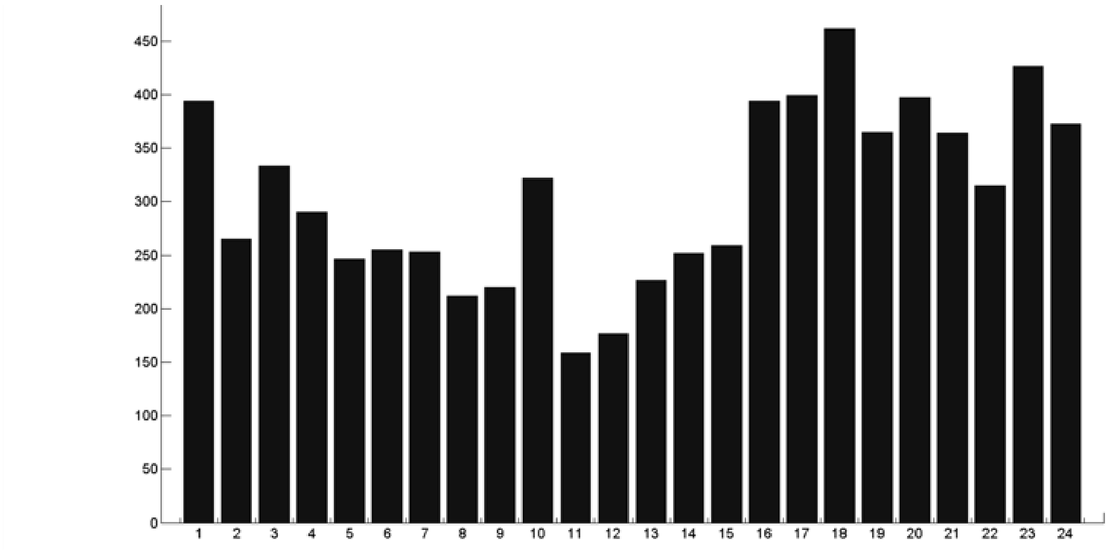

Fig. 6. Distribution of ABS events over the hours of the days
Rys. 6. Dzienny rozkład czasów działania ABS w godzinach

An interesting fact can be deduced from Fig. 6; the time distribution of ABS events is not uniform. The ABS events depend heavily on the external circumstances such as time of day. ABS occurs more frequently during night hours. The dependence of ABS events on a number of factors was further analyzed. Fig. 7 shows the distribution of ABS events as a function of vehicle speed and gross load percentage. According to this figure, ABS events mostly occur to unloaded vehicles. This fact can be explained by the somewhat 'macho' behaviour of certain truck drivers rather than being linked to some vehicular or road problems. In other words, drivers seem to take more care when they drive heavily-loaded vehicles. Another explanation of this observation could be that the same driver action results in faster changes in the movement of an empty truck than in the movement of a fully loaded truck and this may lead to emergency events being detected.

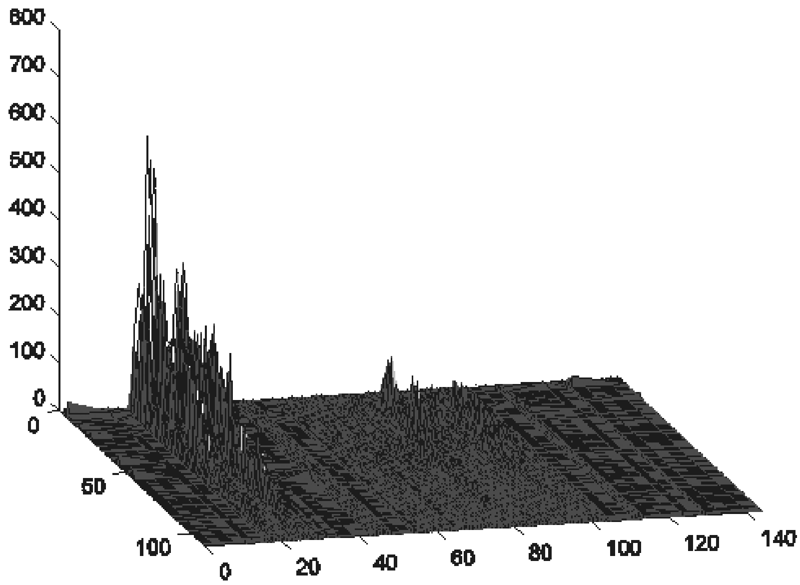

Fig. 7. Distribution of ABS events over vehicle speed (0 - 140 km/h) and gross load percentage (0 - 100 %)
Rys. 7. Rozkład działania ABS dla prędkości od 0 -140 km/h z obciążeniem brutto od 0 do 100%

According to the analysis of the trajectory and braking data stored in TBDB, high percentage of the abrupt braking events occur in urban traffic. It was also observed that a vast majority of ABS events involved vehicles that were unloaded. Another perhaps less surprising result of the mentioned analysis is that the rollover-critical situations for trucks occur primarily in urban areas, particularly at junctions and roundabouts and on winding roads.

In another experiment, trucks' speed histograms were calculated for different areas, namely urban, residential, business/industrial and rural areas. A given location belongs to only one of these surroundings classes. Two main methods of the multi-dimensional statistical analysis were used in conjunction with these histograms: the principal component analysis (PCA) and the hierarchical cluster analysis (HCA).

PCA is a mathematical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of uncorrelated variables called principal components. Thus, each principal component is a linear combination of the original variables and all the principal components are orthogonal to each other, so there is no redundant information.

The principal components may then be used as predictor or criterion variables in subsequent analyses. The number of measurement factors is expected to increase in future, e.g., weather condition, road surface, time of day or time span, external disturbances, longitudinal acceleration, and manoeuvre. In this case three factors are also expected to be sufficient in the PCA analysis.

Cluster analysis creates groups, or clusters, of data. Clusters are formed in such a way that objects in the same cluster are very similar and objects in different clusters are very distinct. A hierarchical cluster analysis with a cluster tree could be also useful in the given context. The tree has a multilevel hierarchy, where clusters at one level are joined as clusters at the next level. An important step in most clustering is to select a distance measure, which will determine how the similarity of two elements is calculated.

3. APPLICATION OF THE RESULTS

Based on statistical analysis of TBDB data many conclusions can be drawn. Some of these are quite obvious, some are more surprising. Information about the critical road sections or intersections can be built into a navigation device to support the truck and car drivers, or even motorists and cyclists. The companies which are concerned with traffic and road maintenance obtain verified information about the critical sections. They can modify the traffic by using signs of speed limit, signs of dangerous bends and suggesting alternative roads, etc. The variable-message sign usually provides traveller information about special events, such as traffic congestion, accidents or other incidents. Based on the result of the statistical analysis additional information may be provided depending on time of day or weather conditions.

For a given area, the interactions between road sections, such as multiple intersections, highway slip roads and roundabouts, can be analyzed. Based on the analysis alternative routes can be suggested taking advantage of possible parallel routes.

4. CONCLUSION

A database of trajectory and vehicular emergencies was built and used for multiple statistical analyses and evaluations. The vehicular emergencies, such as abrupt braking and steering, were observed and collected by on-board vehicular safety systems (e.g., ABS, ESP, ROP). With statistical analyses – ranging from simple to computationally more demanding – performed on this database, hidden, not obvious relationships between drivers, trucks and roads can be explored and determined.

Based on statistical analysis a large number of conclusions can be drawn; some obvious ones, some that are not so obvious, and some which are actually quite surprising. The statistical results can be utilized in many ways, for example, interactions between the road sections, such as multiple intersections, highway slip roads and roundabouts in a given area can be assessed from a traffic safety viewpoint.

The geographical areas near the truck trajectories were classified as urban, residential, commercial/ industrial and rural surroundings based on morphological and other visual characteristics.

The results concerning these surroundings showed that the driving style of the truck driver is significantly affected by the different surroundings and obviously by the corresponding restrictions and speed limits. The TBDB data, which comes from a particular country, are concrete and particular, however, to some extent they can be generalised, especially for countries of similar socio-economic and cultural development.

Acknowledgement

The financial support provided by the Hungarian National Office for Research and Technology in the frame of grant TECH_08_2/2-2008-0088 is gratefully acknowledged by the authors.

The authors wish to express their gratitude to Prof. J. Piecha for his useful comments, which have improved their paper.

Bibliography

1. Fazekas Z., Gáspár P., Kovács R.: *Visualisation and safety evaluation of truck route and braking data.* In: The Automobile and the Environment: International Congress of Automotive and Transport Engineering, A. Chiru (ed.), Cambridge Scholars Publishing, Newcastle-upon-Tyne, UK, 2011, pp. 485-492.
2. Fazekas Z., Gáspár P., Kovács R.: *Determining truck activity from recorded trajectory data.* Procedia Social and Behavioral Sciences, vol. 20, 2011, p. 796-805.
3. Gáspár P., Fazekas Z., Soumelidis A.: *Notes on vehicle-based visual traffic surveillance for crash-prediction.* Science & Military, 2009, pp. 26-32.
4. Gettman D., Head L.: *Surrogate safety measures from traffic simulation models.* Transportation Research Record: Journal of the Transportation Research Board, vol. 1840, 2003, pp. 104-115.
5. Gettman D., Lili P., Sayed T., Shelby S.: *Surrogate safety assessment model and validation: Final Report.* 2008, FHWA-HRt-08-051, Federal Highway Administration, Washington, D.C., USA.
6. Hesse M., Rodrigue J.-P.: *The transport geography of logistics and freight distribution,* Journal of Transport Geography, vol. 12, 2004, pp. 171-184.
7. Montella A., Colantuoni L., Lamberti R.: *Crash prediction models for rural motorways.* Journal Transportation Research Record: Journal of the Transportation Research Board, vol. 2083, 2008, pp. 180-189.
8. Onoe Y., Yamazawa K., Takemura H., Yokoya N.: *Telepresence by real-time view-dependent image generation from omnidirectional video streams.* Computer Vision and Image Understanding, vol. 71, 1998, pp. 154-165.
9. Piecha J., Staniek M.: *Vehicles trajectories movement description by means of syntactic method.* Transport Problems, vol. 4, 2009, pp. 53-60.
10. Sesztakov V., Török, Á.: *Monetary value of road accidents in Hungary*. Proceedings of International Scientific Conference on Modern Safety Technologies in Transportation, Kosice, Slovakia, 2007, pp. 250-253.
11. Vazirgiannis M., Wolfson O.: *A spatiotemporal model and language for moving objects on road networks.* In: Advances in Spatial and Temporal Databases, Jensen C. S. et al. (eds.), Lecture Notes in Computer Science, Springer-Verlag, vol. 2121, 2001, pp. 20-35.

Received 02.03.2011; accepted in revised form 19.07.2012