Information Fusion 12 (2011) 4–10

Contents lists available at ScienceDirect

Information Fusion

journal homepage: www.elsevier.com/locate/inffus

Data fusion in intelligent transportation systems: Progress and challenges – A survey

Nour-Eddin El Faouzi [a,b,c], Henry Leung [d], Ajeesh Kurian [d,*]

[a] Transport and Traffic Engineering Laboratory, INRETS, LICIT, Bron F-69675, France
[b] ENTPE, LICIT, Vaulx-en-Velin F-69518, France
[c] University of Lyon, Lyon F-69003, France
[d] Department of Electrical and Computer Engineering, University of Calgary, Calgary, Alberta, Canada T2N 1N4

ARTICLE INFO

Article history:
Received 9 June 2010
Accepted 9 June 2010
Available online 18 June 2010

Keywords:
Intelligent transportation systems
Information fusion
Advanced traveler information systems
Global positioning systems
Incident detection

ABSTRACT

In intelligent transportation systems (ITS), transportation infrastructure is complimented with information and communication technologies with the objectives of attaining improved passenger safety, reduced transportation time and fuel consumption and vehicle wear and tear. With the advent of modern communication and computational devices and inexpensive sensors it is possible to collect and process data from a number of sources. Data fusion (DF) is collection of techniques by which information from multiple sources are combined in order to reach a better inference. DF is an inevitable tool for ITS. This paper provides a survey of how DF is used in different areas of ITS.

1. Introduction

Providing accurate traffic information is becoming a major challenge for the public institutions and private companies leading to the rapid growth of intelligent transportation system (ITS) [1]. At the same time, the emergence of new information technologies and the transformation that has occurred in road traffic management has both increased a need for very accurate road traffic information. In order to provide an accurate and more comprehensive traffic state on a road network, the traffic sensors that are usually used to measure the prevailing traffic conditions are ineffective. Other sources of data (such as cameras, GPS, cell phone tracking, and probe vehicles) are increasingly used to supplement the information provided by those conventional measurement systems. In addition, authorities normally keep track of traffic activities and archive such information. This offline information, together with the measurements from other sensors is often found to be useful in predicting the traffic trend. Multiple sources may provide complementary data, and multi-source data fusion can produce a better understanding of the observed situation by decreasing the uncertainty related to the individual sources. The fusion of multiple sources is perceived, rightly, as a well-adapted answer to the operational needs of traffic management centers and traffic information operators, allowing them to achieve their goal more efficiently. The primary goal of this survey paper is to acquaint the reader with the most significant applications of data fusion (DF) techniques in intelligent transportation systems and to indicate the directions for future research in this area.

The paper is organized into five sections. Section 2 describes basic traffic engineering operations with emphasis on data sources available. DF applications to the traffic engineering area are presented in Section 3. Section 4 describes prospective research analysis with conclusions in Section 5.

2. Data fusion background

Data fusion is applied in diverse fields in civilian and military applications such as surveillance and reconnaissance, wildlife habitat monitoring, and detection of environment hazards [2–5]. Several methodologies have been proposed in the literature for the purpose of multi-sensor fusion and aggregation under heterogeneous data configurations. Due to the different types of sensors that are used and the heterogeneous nature of information that needs to be combined, different data fusion techniques are being developed to suit the applications and data. These techniques were drawn from a wide range of areas including artificial intelligence, pattern recognition, statistical estimation, and other areas. Traffic engineering field has naturally benefited from this abundant literature. For instance, independent of specific application a variety of techniques can be used for ranging from a sample arithmetic mean

* Corresponding author.
E-mail addresses: elfaouzi@inrets.fr (N.-E. El Faouzi), leungh@ucalgary.ca (H. Leung), akurian@ucalgary.ca, ajeesh@ieee.org (A. Kurian).

1566-2535/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.inffus.2010.06.001

to a more complex DF approach. More precisely, a three-way split could be suggested:

– Statistical approaches: weighted combination, multivariate statistical analysis and its most up-to-date form data mining engine [6]. Among statistical techniques, the arithmetic mean approach is the simplest which is used for information combination. This approach is not suitable when the information at hand is not exchangeable or when estimators/classifiers have dissimilar performances [7–9].
– Probabilistic approaches: for instance, Bayesian approach with Bayesian network and state-space models [10], maximum likelihood methods and Kalman filter based DF [11,12], possibility theory [13], evidential reasoning and more specifically evidence theory [14–16] are widely used for the multi-sensor data fusion. This later technique could be viewed as a generalization of Bayesian approach [15–17].
– Artificial intelligence: neural networks and artificial cognition including artificial intelligence, genetic algorithms and neural networks. In many applications, this later approach serves both as a tool to derive classifiers or estimators and as a fusion framework of classifiers/estimators [6,8].

Although application of DF techniques to complex systems modelling is not new [18–20], there is a growing interest in their use in transportation systems. Road traffic could be considered as a field where benefits expected from the application of DF techniques are fruitful. However, the benefits come with challenges in assessing feasibility, effectiveness and usefulness of such approaches [21–23]. In traffic engineering literature, the interest for DF is quite new and it coincides with ITS advent. The first paper which mentions the DF was by Sumner in the early 1990s [24]. He acknowledges the importance of DF for effectiveness of ITS systems. Numerous papers exist regarding the application of DF in engineering [22,25,26].

Many of the data processing techniques originally developed by the US Department of Defence (DoD) to support the identification and tracking of military objects can be used today to aid traffic management on streets and highways [28–32]. The DoD data fusion model consists of a hierarchy of five processing levels. Level 0 deals with pre-processing of data from the contributing source. It may normalize, format, order, batch, and compress input data [2,30]. It may even identify sub-objects or features in the data that are used later in Level 1 processing. For traffic management, Level 1 processing concerns the gathering of data from all appropriate sources, including real-time point and wide-area traffic flow sensors, transit system operators, toll data, cellular telephone calls, emergency call box reports, probe vehicle and roving tow truck messages, commercial vehicle transmissions, and roadway-based weather sensors [26,27]. Level 2 processing identifies the probable situation causing the observed data and events by combining the results of the Level 1 processing with information from other sources and databases. These sources may include patrol reports and databases, roadway configuration drawings, local and national weather reports, anticipated traffic mix, time-of-day traffic patterns, construction schedules, and special event schedules. Level 3 processing assesses the traffic flow patterns and other data with respect to the likely occurrence of a traffic event (e.g., traffic congestion, incident, construction or other pre-planned special event, fire, or police action) that impacts traffic flow. Level 4 processing seeks to improve the entire data fusion process by continuously refining predictions and assessments, and evaluating the need for additional sources of information. Sometimes a sixth level is added to address issues concerned with enabling a human to interpret and apply the results of the fusion process. The DF process investigated in traffic literature involves basic functions such temporal or/and spatial alignment of input data, data association and data mining for knowledge extraction purpose. This later purpose is also one of potential objectives of a multi-source information fusion [33].

3. Opportunities and challenges of ITS data fusion

Technological advances in the area of road telematics (such as on-board electronic systems, vehicle localization mechanisms, telecommunications, and data processing) have favoured an improvement in existing means of traffic data collection. This includes the invention of sensors or new architectures: equipment on board vehicles, off-set on the side of roads, multiform data collection, etc. Basic traffic data is largely based on road sensors embedded in the pavement. Such sensors mostly use inductive loop detectors (ILDs) and are able to measure temporal traffic characteristics. They detect basic parameters needed by operational traffic engineer, such as traffic volume, occupancy and speed at a given point. Other fixed sensors such as optical detectors, ultrasonic detectors were developed and used for network surveillance. These fixed sensors are very useful, but they fail in measuring spatial behaviour of traffic. In addition they suffer from their limited reliability, with their prohibitive cost in attaining significant coverage of the roadway network. Other fixed sensors with spatial capabilities have been developed and used to supplement loop detector data: cameras, RMTS (Remote Traffic Microwave Sensor), etc.

With the recent deployment of ITS applications and needs for real-time and accurate data to fulfil a wide range of purposes (real-time traffic operations monitoring, incident detection and route guidance applications), have given rise to another complementary source of data allowing traffic parameter estimation. One of the actual trends is probe vehicle data collection technique, also known as Floating Car Data (FCD) and its extended version (xFCD). According to this technique, cars on the road shift from a passive attitude to an active one and act as moving sensors, continuously feeding information about traffic conditions to a Traffic Management Center (TMC). More recently, within cooperative systems, research was carried out where vehicles are connected via continuous wireless communication with the road infrastructure, exchanging data and information relevant for the specific road segment to increase overall road safety and enable cooperative traffic management. Different automatic vehicle identification (AVI) systems, based on different technologies, can be used as detection devices: automatic vehicle tags identification, automatic license plate matching techniques, and global positioning system (GPS). With the advances in wireless communications and the spread of cellular phones, technical developments in cellular positioning provides the opportunity to track cell phone equipped drivers as traffic probes. Many research studies have demonstrated the feasibility of using cell phones as traffic probes [34,35].

Therefore, a wide spectrum of data and heterogeneous sources of information are of potential use for a given traffic situation. As a result, many problems of traffic engineering become a typical data fusion problem. Indeed, some of the basic road traffic engineering used to meet operators needs have been gained through the conventional fixed detectors. However, in the context of traffic operations where highly accurate information are needed, ITS framework for instance, the information provided by fixed detectors data alone may be not sufficient except in some special situations (for some network configuration and/or with a high detector coverage). One of the reasons is the lack of full traffic information at the network wide scale is of importance. So, for traffic operations improvements, other sources of data are increasingly used in order to supplement the information provided by conventional measurement techniques. The purpose of DF is to produce an improved model or estimate of a system from a set of independent data sources. For traffic applications, the desired model is the state vector of the traffic phenomenon. These estimates may include statements about current or future vehicular speeds, mean speeds, travel time, vehicle classification and similar topics of interest to travelers and traffic operators.

4. Its data fusion applications

El Faouzi and Lesort [7] and Sethi et al. [36] published the early papers in which a practical traffic and transportation problem was addressed as a data fusion problem. For the last 15 years, various authors have made significant contributions to the field of DF in transportation systems. ITS offers indisputably the most relevant framework for DF and also the most challenging, see e.g. [21–23]. Others conventional problems in transportation modelling have also being concerned with multi-source processing, namely: planning problems, demand estimation, traffic estimation, etc. [37]. A variety of functions are assigned to ITS to address the traffic congestion and safety problems. These functions were designed to achieve a specific task in order to cope with an operational problem. One can mention advanced traveler information systems (ATIS), automatic incident detection (AID), advanced driver assistance (ADAS), Network control, crash analysis and prevention, traffic demand estimation, traffic forecast and monitoring and accurate position estimation. Each of these sub-systems can make use of different information sources. DF techniques can then be used to combine them to yield better results.

4.1. Advanced traveler information systems

ATIS is one of the several ITS sub-systems that offer users integrated traveler information. In ATIS, different automatic data collection techniques are of potential use to comprehend traffic conditions and derive relevant indicators to assist in driver guidance [21]. One form in which the user's information is presented is the travel time, and a number of systems are based on its dissemination. In this context, it is used as a measure of impedance (or cost) for route choice strategies. Travel time is also used by network operators as an indicator of quality of service (QoS). This raises the problem of estimating travel times with an acceptable degree of accuracy which is a particularly difficult task in urban areas as a result of difficulties of a theoretical, technical and methodological nature. Thus, in order to find out the traffic conditions that prevail on an urban road, the traffic sensors that are usually used to measure traffic conditions are almost ineffective. New measurement devices proliferation (cameras, GPS or cell phone tracking, etc.) mean that other sources of data are increasingly used in order to supplement the information provided by conventional measurement techniques and improve the accuracy of travel time estimates. As a result, travel time estimation becomes a typical DF problem.

Many authors discussed the statement of requirements of data fusion for ADVANCE program [38–40]. ADVANCE was an in-vehicle ATIS providing route guidance in real time that operates in the northwestern portion and northwest suburbs of Chicago. It uses probe vehicles to generate dynamic travel time information about expressways arterials and local streets. ADVANCE uses a general framework for combining data from loop detectors and travel time reports of probe vehicles using inference rules. Evaluations of the proposed algorithms found that probe data greatly improves static (archival average) link travel time estimates by the time-of-day. Dailey et al. [41] reports a detailed description of a current data amalgamation (fusion) within ITS project and the presentation of a new quantitative data fusion algorithm to estimate speed from volume and occupancy measurements. Since then, many other contributions present various frameworks for the evaluation of ITS effectiveness based on data fusion. Data fusion approaches are presented for various types of measures of effectiveness and techniques for handling biases of various kinds are developed.

El Faouzi et al. [7,8,17] proposed an estimation framework for real-time traffic characterisation based on multi-source data. As an illustrative example, a multi-source travel time estimation was performed based on two data sources: data from conventional loop detectors which deliver Eulerian data and probe vehicles collecting Lagrangian data. Travel time measurements collected by license plate matching technique were considered as a reference and were used for validation purposes only. The first technique used is of statistical nature and can be viewed as a distributed estimation problem: each source derives an estimator of travel time and the individual estimates are then combined according to weighted mean strategy. The weights were derived from variance–covariance estimation errors. Results display propensity of proposed schemes for estimation accuracy improvement. More recently, the evidence theory was used to solve the same problem [17,23]. In these contributions, travel time was broken into classes and formulate the estimation problem as a classification one. Various strategies for classifiers fusion were proposed and their evaluation shown some improvements capabilities in terms of classification rate.

Abe [42] reported work on travel time forecasting where data from automatic vehicle identification devices were used for the correction. Dynamic route guidance systems (DRGS) are also an area where DF is of potential use. Kühne [43] has proposed a framework for fusing information from various sources within the DRGS. Once again, the objective is travel time estimation and prediction. The data consists of loop detectors, probe vehicle and QoS indicator with some exogenous information: information on road works and incidents. The proposed solution was based on a weighted mean scheme. The weights were derived according to the source reliability. Choi [44] and Choi and Chung [45] have tackled the problem of generating travel time from loop detectors, probe vehicles and video-camera sources. They proposed a fuzzy logic based approach with its evaluation on a theoretical example.

4.2. Automatic incident detection

Incident detection methods for automatic recognition of incidents, accidents and other road events requiring emergency responses have existed for more than tree decades. Most of the developed and implemented algorithms rely on loop detectors data. However, these algorithms work with mixed success. Recently, there has been renewed interest in incident detection algorithms partly because of the availability of new sensors and data sources. One of these sources is probe vehicles. Hence, AID belongs to the class of problems that can be solved by DF techniques. Applications of several data fusion techniques to traffic management to support incident detection have been reported in the literature, and the data fusion algorithms used includes Dempster–Shafer inference, Bayesian inference, and voting logic. Most of these applications have explored the use of probe vehicles data with the conventional traffic data for incident detection purposes. As an example of such work, Koppelman et al. [46], Ivan et al. [47,48] developed an AID system using surveillance data from two different sources: fixed detectors (e.g. inductive loop detectors) and probe vehicle specially equipped to report link travel time. The neural network approach was considered and two strategies were tested. The first one combined observed traffic directly to determine whether or not an incident is occurring. In the second, separate incident detection algorithms individually preprocess data from each source, reporting scores which are combined by neural network. Different neural network representations were studied in [48] and results found that probe and detector based incident detection on arterial networks shows considerable promise for improved performance and reliability. Dempster–Shafer inference or evidential reasoning was also used to perform an operational AID system [49].

Thomas [50] has investigated this problem from a multiple attributes decision making standpoint, with Bayesian scores. The author proposed an approach which utilizes the combinations of probe travel times, number of probe reports, detector occupancies and volume as the inputs. It is shown that models based solely on probe data lack in performance due to excessive overlaps in class distributions and models based on detector occupancies and vehicle counts by lane perform outstandingly. The probe data is shown to enhance the performance of detector data based models. More recently, Klein [51] studied the application of Dempster–Shafer inference to the traffic management to support incident detection and the identification of other events of concern to traffic managers. The application of the Dempster–Shafer inference algorithm to incident detection and verification is illustrated with an example consisting of three possible events, where data are supplied from three different types of sources. The available information is combined using Dempster's rule and the most probable event is identified. Incident detection algorithms fusion is another direction for classification accuracy improvement. This direction was investigated by Cohen [52]. In this paper three aggregation schemes were investigated: a logical aggregation, a neural network fusion and a veto procedure. From the validation step, which was carried on real-world data, results demonstrated that both logical aggregation and veto procedure outperform the single best algorithm.

4.3. Advanced driver assistance

Passenger safety is considered as of the important aspects of ITS. Driver assistance techniques are being developed from this point of view. In the past few decades, tremendous progress has been made with regards to vehicle safety and driver assistance. Early safety approaches emphasize precaution and focus on passive devices such as seat belts, air bags and lighting. In spite of crash-related injury severity rate reduction, drivers demand for greater improvements in vehicular transportation safety. Research trends show the use of active safety device which complements the traditional passive ones. ADAS and collision avoidance systems (CAS) are an illustration of such trends. The main objective assigned to these systems is to provide a more reliable description of the traffic scene surrounding the vehicle to vulnerable road users, in pre-crash situations and to systems like adaptive cruise control (ACC) and collision avoidance systems. Simultaneous localization and mapping is a technique used to obtain the static map of the environment and its position in the map [53]. Automated highways are another research topic where DF is an important area. On the other hand, autonomous vehicles are gaining importance due to their potential use in hazardous and unknown environments. In any case, the vehicle needs to sense its environment with an array of sensors and the sensory information needs to be used effectively to provide decision support. Challenges involved are the heterogeneous nature of the data and extracting relevant features from the measurements. Typically, sensors of different capabilities are used to gain complimentary information.

Simultaneous localization and tracking (SLAM) has been an active research area in robotics for the last ten years. SLAM consists of multiple parts; landmark extraction, data association, state estimation, state update and landmark update. Since individual steps can be achieved using a multitude of algorithms, there is no universally accepted algorithm for SLAM. Detection and tracking of moving objects form another set of techniques used to obtain information about the dynamic environment in which the vehicle is operating [54,55]. Recently, there have been many commercial products in the market capable of alerting drivers about lane changes. In such systems, artificial intelligence techniques are used with image processing tools to extract information from 2D and 3D cameras. Initially, edge based lane detection techniques were used. Since there is a good contrast between the road and lane markings, thresholding of the images was found to be useful and the next step would be a perceptual grouping of the edge points to detect the lane markers of interest [56–58]. In ARCADE [59], which uses slightly more advanced techniques than simple edge detection, one-dimensional edge detection is followed by a least median squares technique for determining the curvature and orientation of the road. Individual lane markers are then directly determined by a segmentation of the row-averaged image intensity values. Frequency domain techniques for lane extraction are detailed in Kreucher [60].

Operational systems are based on several sensor systems which are complementary and redundant and a DF process provides a fused description of the traffic scene. This fusion incorporates the data of the available sensors into a single description. The problems to solve here are: Data association: sensor data have to be associated with environment description, which require synchronization of the sensor data and associated object state. Whenever, there are multiple sensors used to sense multiple objects, there is a need to associate the measurements with the individual objects [2]. Once the sensor measurements are associated with appropriate objects, the next step is to remove the sensor bias. This procedure is called sensor registration. Finally, objects are tracked using fused sensor measurements. Again, Kalman filter, its variants and more recently particle filtering become an essential tool to perform this step [61]. Several papers report some results within this topic. For example, Murphy [62] discussed sensor fusion's role in-vehicle guidance and navigation, and proposed general methods for fusing data, and sensor-fusion activities within a robot architecture. In their work, Pei and Liou [63] proposed three-dimensional vehicle motion estimation by fusion of multi-source information. Image point and line features were considered for fusion. Langheim et al. [64] investigated DF systems for Automatic Cruise control with stop and go phenomenon, and Stiller et al. [65] reported a DF framework for obstacle detection and tracking.

4.4. Network control

Data fusion techniques were also applied in the road network control issues. In [66] and [67], the problem of constructing an adaptive online traffic control in urban or freeway road networks was investigated. Mueck work's [66] a model that determines queue length on the basis of vehicle counts from detectors located close to the stop line and on the basis of signal timings was derived. Wang and Papageorgiou [67] performed the freeway traffic control using extended Kalman filter. Along this direction, Friedrich et al. [68] introduced a new approach based on in queuing theory models for real time queue length determination. In this later method, Mueck's model serves as a quasi measurement with Kalman filtering technique.

4.5. Crash analysis and prevention

Although there has been a steady reduction in the number of accidents, they continue to sustain heavy losses in both human and economic terms. Reduction in the number of accidents could be due to multiple efforts: road infrastructure improvement, regulations on alcohol and speed, and an improvement in-vehicle safety. Many studies were carried out in order to explain the circumstances and the characteristics of traffic accidents. One way to conduct such explorative studies is to utilize retrospective data available such as traffic accident records. Along this direction, Sohn and Shin [69] employed both neural network and decision tree algorithms to find the classification model for road traffic accident severity (bodily injury or property damage) as a function of potentially related categorical factors. They noted that the classification accuracy of the individual algorithm was relatively low. Recognizing that, Sohn and Lee [70] use data fusion and ensemble algorithms to increase the accuracy. DF techniques are used to perform classifiers fusion using the evidence theory. Data ensemble combines various results obtained from a single classifier fitted repeatedly based on several bootstrap samples [71–74]. More precisely, they tried three different approaches: classifier fusion based on the Dempster–Shafer algorithm, the Bayesian procedure and logistic model; data ensemble fusion based on arcing and bagging; and clustering based on the k-means algorithm. Empirical results indicate that a clustering based classification algorithm works best for road traffic accident classification.

4.6. Traffic demand estimation

One of the most important problems in the field of transportation planning and control is the problem of origin–destination estimation from link counts. In order to decrease the cost of passenger surveys, traffic count are undertaken on certain links of the transportation network. An estimation of a most likely origin–destination matrix is then derived from counts. In the last two decades a greater number of models have been developed for Origin–destination (OD) estimation from link counts. Some of the proposed schemes derived the OD matrices by combining data from different sources. An illustration of this class of problems is the dynamic OD estimation initiated in [75,76]. Further developments along the same line were pursued later by many authors [10,77]. The Kalman filtering is of common practice in this class of problems. Ben-Akiva and Morikawa [78] have explored the OD estimation methods that combine different data sources (stated preference data and traffic measurements) and more recently, Lundgring et al. [79] described a method for adjusting time-dependent travel demand information with respect to link flow observations. They utilized the structure of the given OD-matrix, which is compounded from different sources, for making simple overall adjustments.

4.7. Traffic forecasting and traffic monitoring

Traffic flow forecasting has received increasing attention in the past years and different techniques have been developed mainly for traffic surveillance and control. Many prediction schemes of traffic flow were obtained by means of classic autoregressive models, especially time series techniques. Some authors have tackled this problem in the context of Bayesian framework [80]. Some others used Kalman filtering technique [81] or neural networks and system identification [82] and more recently a nonparametric paradigm was adopted via kernel techniques [83]. None of these proposals allow one to achieve highly accurate predictions except in some special situations (for some network configuration and/or with high detector coverage). This is induced to some extent by traffic dynamic which cannot be formalized by a single procedure. Therefore, in the context of traffic operations where highly accurate forecasts are needed, one can obtain different forecasts of the same quantity (the underlined assumption here is that different predictors are measures of the same quantity and/or various aspects of the same thing) by two or more different methods. The set of available methods may consist of alternative models, different forecasters, or a mixture of models and forecasters.

Often, the approach used is to find the single ‘best’ predictor in some sense (most accurate values, most appropriate models of the underlying process, most cost-effective, etc.) among the available forecasting methods. Another approach consists of combining these individual forecasts. The idea of combining estimators instead of selecting the single ‘best’ model has a long history and has generated intensive theoretical works since the seminal article of Granger [84]. In this work, it is showed that the linear combination of several predictors from a single data set can outperform the individual predictors, methodological and practical issues related to combining forecasts produced by different methods has been investigated extensively in various contexts with notable successes. In traffic forecasting under heterogeneous data sources configuration, El Faouzi provided a methodical framework to combine various forecasts of the same quantity [84,85]. He derived two predictors using nonparametric traffic flow using a kernel estimator and predicting scheme based on the propagation of a lagged upstream traffic flow. The proposed combination strategies exhibit very encouraging results. Data integration and data fusion were applied for other purposes. In [86], the integration problem of in-vehicle information and data provided by loop detectors was studied. The core of the integration step was the extended Kalman filtering. More recently, Sau et al. [87] investigated the traffic monitoring problem within the multi-source data. The particle filter is the estimation technique used in this context. Choi [88] examined the problem of missing data estimation and proposed a framework for missing data inference based on evidential reasoning.

4.8. Accurate position estimation

In modern transportation systems, information about the position and the orientation of the vehicle should be accurate. Inertial navigation systems (INS) are one of the earliest forms of navigation techniques. INS, which functions on the principle of dead-reckoning, has a potential problem of “integration drift” which is the accumulation of small errors in the measurement of acceleration and angular velocity into progressively larger errors in velocity, which are compounded into still greater errors in position. In the last few decades, GPS, which is initially developed as a military navigation aid, has gained a wide acceptance in civilian navigation systems. GPS is based on three major components: satellites orbiting the Earth; control and monitoring stations on Earth; and the GPS receivers owned by users. GPS satellites broadcast signals from space that are picked up and identified by GPS receivers. Each GPS receiver then provides three-dimensional location (latitude, longitude, and altitude) plus the time [89]. When the satellite signals are blocked by tall buildings or due to the other electromagnetic interference, GPS outage occurs. In such situations, due to the lack of reference signals, the estimation of position is impossible and the device ceases to work. DF can effectively be used to combat the drawbacks of both techniques. The benefits of using GPS with an INS are that the INS may be calibrated by the GPS signals and that the INS can provide position and angle updates at a quicker rate than GPS. For high dynamic vehicles such as missiles and aircraft, INS fills in the gaps between GPS positions. Additionally, GPS may lose its signal and the INS can continue to compute the position and angle during the period of lost GPS signal. The two systems are complementary and are often employed together.

One of the earliest approaches in GPS/INS integration is to use Kalman filter. For instance, in [90], a decentralized filtering strategy is developed for the GPS/INS integration. A Kalman smoother is used to integrate the differential range and phase measurements with the data from INS. Variants of Kalman filter are often employed to get a better integration. A constrained unscented Kalman filter algorithm has been proposed in [91] to fuse differential GPS, INS (gyro and accelerometer) and digital map to localize vehicles for ITS applications. For Kalman filter to work, accurate stochastic models of the sensors are required. Such requirements are often difficult to achieve and resulted in the application of artificial intelligence technique for GPS-INS integration. Different types of neural networks have been used to combine the GPS and INS information. For example, multi layer perception and radial basis function neural networks are successfully used of the GPS/INS integration [92,93]. GPS/INS integration is performed using adaptive neuro-fuzzy techniques are used in [94]. Such systems mimic the vehicle dynamics by training the AI modules during the availability of GPS signals.

5. Conclusion and future directions

Needs for DF has emerged from transportation applications for at least two decades and give rise to an emergent field which is somewhat in its infancy. This survey paper, although focuses exclusively in road traffic problems, has described the state of the art and practice of fusion of traffic data from various sources. For all the applications reported in this article, DF techniques seen promising. However, these encouraging results should not conceal the problem that still remains to be solved before any operational widespread deployment of DF in transportation field. These challenges include the accuracy necessary for the effective application, dynamic and real time aspects of the traffic and data quality, real time dimension. The assessment of the benefits of DF will be more readily performed with the increase number of successful practical applications of DF in transportation field. It is most definite, however, that there are real opportunities for greater DF application in road transportation systems. Prospects include the increased collection of usable data from different sources other than that to installed sensors for traffic surveillance. Wireless technologies, which offer (i) the potential of easier reporting and access to customized information (e.g. cooperative systems with vehicle-to-vehicle, vehicle-to-infrastructure and infrastructure-to-vehicle) and (ii) the new ability of tracking individual vehicles and information collected by FCD/xFCD will enrich the available information on traffic situation, will certainly accelerate needs for DF operational systems.

References

[1] Federal Highway Administration, ITS system architecture and standards, FHWA ruling on architecture standards and conformance of projects to ITS architecture standards, August, 2001.
[2] D.L. Hall, Mathematical Techniques in Multisensor Data Fusion, Artech House, Norwood, MA, 1992.
[3] J. Mnyika, H. Durrant-Whyte, Data Fusion and Sensor Management: A Decentralized Information Theoretic Approach, Ellis Horwood, London, 1994.
[4] C. Harris, A. Bailey, T. Dodd, Multisensor data fusion in defense and aerospace, Journal of Royal Aerospace Society 162–1015 (1998) 229–244.
[5] B.V. Dasarathy, Sensor fusion, potential exploitation: innovative architectures and illustrative applications, Proceedings of IEEE 85 (1997) 24–38.
[6] J. Han, J. Kember, Data Mining Concepts and Techniques, Morgan Kaufmann Publishers, USA, 2000.
[7] N.-E. El Faouzi, J.-B. Lesort, Travel time estimation on urban networks from traffic data and on-board trip characteristics, in: Proceedings of the Second World Congress on ITS, Yokohama, Japan, 1995, pp. 88–93.
[8] S. Hashem, Optimal linear combinations of neural networks, Neural Networks 10 (1997) 599–614.
[9] N.-E. El Faouzi, Heterogeneous data source fusion for impedance indicators, in: IFAC Symposium on Transportation Systems, Chania, Greece, vol. 3, 1997, pp. 1375–1380.
[10] I. Okutani, The Kalman filtering approaches in some transportation and road traffic problems. in: Gartner, Wilson (Eds.), Proceedings of the 10th ISTTT, 1987, pp. 397–416.
[11] D. Huang, H. Leung, EM-IMM based land-vehicle navigation with GPS/INS, in: IEEE International Conference on Intelligent Transportation Systems, Washington, DC, October 2004.
[12] D. Huang, H. Leung, An expectation maximization based interactive multiple model approach for collaborative driving, IEEE Transactions on Intelligent Transportation Systems 6 (2005) 206–228.
[13] D. Dubois, H. Prade, Possibility Theory, Plenum Press, New York, 1988.
[14] A.P. Dempster, Upper and lower probabilities induced by multivalued mapping, Annals of Mathematical Statistics 38 (1967) 325–339.
[15] A.P. Dempster, A generalization of Bayesian inference, Journal of the Royal Statistical Society Series, B 30 (1968) 205–247.
[16] G. Shafer, Mathematical Theory of Evidence, Princeton 2702, Princeton University Press, 1976.
[17] N.-E. El Faouzi, L.A. Klein, O. De Mouzon, Improving travel time estimates from inductive loop and toll collection data with Dempster–Shafer data fusion, Transportation Research Record: Journal of the Transportation Research Board 2129 (2009) 73–80.
[18] H. Durrant-Whyte, Sensor models and multisensor data fusion, Journal of Robotics Research 6 (1988) 97–113.
[19] E. Waltz, J. Llinas, Multisensor Data Fusion, Artech House, Boston, 1990.
[20] Y. Zhou, H. Leung, P.C. Yip, An exact maximum likelihood registration algorithm for data fusion, IEEE Transactions on Signal Processing 45 (1997) 1560–1572.
[21] Federal Highway Administration. Data fusion for delivering advanced traveller information services, US DoT and ITS joint Program, 2003, pp. 247–254.
[22] N.-E. El Faouzi, Data Fusion: Concepts and Methods, Monograph: INRETS, 2000 (INRETS Edition).
[23] N.-E. El Faouzi, Travel time estimation via evidential data fusion, Recherche Transports Sécurité 68 (2000) 15–30.
[24] R. Sumner, Data fusion in PathFinder and TravTek, in: Proceedings of the Vehicle Navigation and Information Systems Conference: VNIS'91, 1991, pp 71–75.
[25] M. Abidi, R. Gonzalez (Eds.), Data Fusion in Robotics and Machine Intelligence, Academic Press, 1992.
[26] L.A. Klein, Sensor Technologies and Data Requirements for ITS, Artech House, Boston, 2001.
[27] N.-E. El Faouzi, Multiform traffic data collection and data fusion in road traffic engineering, in: N.-E. El Faouzi (Ed.), Proceedings of the Multisource Data Fusion in Traffic Engineering Workshop, vol. 87, INRETS publishing, 2003.
[28] F.E. White Jr., Joint directors of laboratories data fusion subpanel report: SIGINT session, in: Tech. Proc. Joint Service Data Fusion Symposium I, 1990, pp. 469–484.
[29] Data Fusion Development Strategy Panel, Functional description of the data fusion process, Office of Naval Technology, 1991.
[30] A.N. Steinberg, C.L. Bowman, F.E. White Jr., Revisions to the JDL data fusion model, Proc. SPIE 3719 (1999) 430–441.
[31] H. Leung, J. Wu, Bayesian and Dempster–Shafer target identification for radar surveillance, IEEE Transactions on Aerospace and Navigational Electronics 36 (2000) 432–447.
[32] E. Lefebvre, Advances and challenges in multisensor data and information, NATO Security through Science Series: Information and Communication Security, vol. 8, 2007.
[33] B.V. Dasarathy, Information fusion, data mining, and knowledge discovery, Information Fusion 4 (2003) 1.
[34] J.L. Ygnace, Travel time/speed estimates on the French Rhone corridor network using cellular phones as probes: STRIP project, Final Report, European Community Research Program SERTI, December 2001.
[35] Y. Youngbin, R. Cayford, Investigation of vehicles as probes using global positioning system and cellular phone tracking, PATH Working Paper, 2000.
[36] V. Sethi, N. Bhandari, F. Koopelman, J. Schofer, Arterial incident detection using fixed detector and probe vehicle data, Transportation Research 2 (1995) 99–112.
[37] J. Sussman, Introduction to Transportation Problems, Artech House, Norwood, Massachusetts, 2000.
[38] A. Tarko, N. Rouphail, Travel time fusion in ADVANCE, in: Pacific Rim TransTech Conference: A Ride into the Future, July 25–28, 1993, Seattle, Washington, USA.
[39] P. Thakuriah, A. Sen. Data fusion for travel time prediction: a statement of requirements, Advance Working Paper Series, vol. 22, 1993, Task TRF-TT-07, Urban Transportation Center, University of Illinois, Chicago.
[40] S. Berka, X. Tian, A. Tarko, Data fusion algorithm for ADVANCE, Advance Working Paper 48, 1995.
[41] D.J. Dailey, P. Harn, P.-J. Linet, ITS data fusion, Research project T9903, Task 9: ATIS/ATMS Regional IVHS Demonstration, Report# WA-RD 410.1, Washington State, WADOT, 1996.
[42] A. Abe, The correction of the forecasting travel time by using AVI data, in: Proceedings of the I.V.H.S., 1997.
[43] R.D. Kühne, Data fusion for DRG systems, in: IFAC Symposium on Transportation Systems, Chania, Greece, vol. 3, 1997.
[44] K. Choi, Data fusion for generating the link travel time with insufficient data sources, in: Proceedings of the I.V.H.S., 1997.
[45] K.Y. Choi, Y. Chung, Travel time estimation algorithm using GPS probe and loop detector data fusion, in: Proceedings of the 80th TRB Annual Meeting (CD-Rom), Paper No. 01-0374. Washington, DC, Transportation Research Board, 2001.
[46] F. Koppelman, V. Sethi, J. Ivan. Calibration of data fusion algorithm parameters with simulated data, ADVANCE Project, Technical Report, TRF-ID-152, 1994.
[47] J.N. Ivan, J.L. Schofer, F.S. Coppelman, L.L.E. Massone, Real-time data fusion for arterial street incident detection using neural networks, Transportation Research Record 1497 (1995) 27–35.
[48] J.N. Ivan, Neural network representations for arterial street incident detection data fusion, Transportation Research C (1997) 247–254.
[49] S.C. Byun, D.B. Choi, B.H. Ahn, H. Ko, Traffic incident detection using evidential reasoning-based data fusion, in: Proceedings of the 6th World Congress on ITS, Toronto, Canada, November 1999.

[50] N.E. Thomas, Multi-sensor multivariate and multiclass incident detection system for arterial streets, in: Lyons, J.-B. Lesort (Eds.), Proceedings of the 16th International Symposium on Transportation and Traffic theory (ISTTT), Pergamon, Elsevier, 1996, pp. 315–340.
[51] L. Klein, Dempster–Shafer data fusion at the traffic management centre, Transportation Research Board, in: 79th Annual Meeting, Paper No. 00-1211, Washington, DC, Transportation Research Board, 2000.
[52] S. Cohen, Fusion of incident detection algorithms, in: N.-E. El Faouzi (Ed.), Proceedings of the Multisource Data Fusion in Traffic Engineering Workshop, vol. 87, 2003, pp. 109–117.
[53] H. Durrant-Whyte, T. Bailey, Simultaneous localisation and mapping (SLAM): Part I the essential algorithms, Robotics and Automation Magazine 13 (2006) 99–110.
[54] Z. Hu, H. Leung, M. Blanchette, Evaluation of data association techniques in a real multitarget radar tracking environment, Proceedings of SPIE 2561 (1995) 509–518.
[55] C. Wang, C. Thorpe, A. Suppe, Ladar-based detection and tracking of moving objects from a ground vehicle at high speeds, in: IEEE Intelligent Vehicles Symposium (IV2003), June 2003.
[56] J. Tsao, R.W. Hall, S.E. Shladover, Design options for operating automated highway systems, in: Proc. IEEE-IEE Vehicle Navigation Inform. Syst. Conf., 1993, pp. 494–500.
[57] S.K. Kenue, LANELOK: detection of lane boundaries and vehicle tracking using image-processing techniques—Parts I and II, SPIE Mobile Robots IV, 1989.
[58] K.C. Kluge, YARF: an open-ended framework for robot road following, Ph.D. Dissertation, Carnegie Mellon Univ., Pittsburgh, PA, 1993.
[59] K.C. Kluge, Extracting road curvature and orientation from image edge points without perceptual grouping into features, in: Proc. Intelligent Vehicle Symp., 1994, pp. 109–114.
[60] C. Kreucher, S. Lakshmanan, LANA: a lane extraction algorithm that uses frequency domain features, IEEE Transactions on Robotics and Automation 15 (1999) 343–350.
[61] Y. Bar-Shalom, X.R. Li, T. Kirubarajan, Estimation with Applications to Tracking and Navigation: Theory, Algorithms, and Software, Wiley, New York, 2001.
[62] R.R. Murphy, Sensor and information fusion for improved vision-based vehicle guidance, IEEE Intelligent Systems (1998) 49–56.
[63] S.-C. Pei, G. Liou, Vehicle-type motion estimation by the fusion of image point and line features, Pattern Recognition 31 (1998) 333–344.
[64] J. Langheim, A.J. Buchanan, CARSENSE – sensor fusion for DAS, ITS in Europe 2002, Lyon.
[65] C. Stiller, J. Hipp, C. Rossig, A. Ewald, Multisensor obstacle detection and tracking, in: Proc. 1998 IEEE Int. Conf. on Intelligent Vehicles, 1998, pp. 451–456.
[66] J. Mueck, Estimation methods for the state of traffic at traffic signals using detectors near the stop-line, Traffic Engineering and Control, December 2002.
[67] Y. Wang, M. Papageorgiou, Real-time freeway traffic state estimation based on extended Kalman filter: a general approach, Transportation Research B 39 (2005) 141–167.
[68] B. Friedrich, R. Minciardi, Data fusion techniques for adaptive traffic signal control, in: IFAC Symposium on Transportation Systems, Tokyo, Japan, 2003, pp. 53–55.
[69] S.-Y. Sohn, H.W. Shin, Data mining for road traffic accident classification, Ergonomics 44 (2001) 107–117.
[70] S.Y. Sohn, S.H. Lee, Data fusion, ensemble clustering to improve the classification accuracy for the severity of road traffic accident in Korea, Safety Science 41 (2003) 1–14.
[71] L. Breiman, Bagging predictors, Machine Learning 26 (1996) 123–140.
[72] L. Breiman, Arcing classifiers, Annals of Statistics 26 (1998) 801–849.
[73] L. Breiman, Random forests, Machine Learning 45 (2001) 5–32.
[74] D.H. Wolpert, Stacked generalization, Neural Networks 5 (1992) 241–259.
[75] M. Cremer, H. Keller, Dynamic identification of OD flows from counts at complex intersections, in: Proceedings of the 8th International Symposium on Transportation and Traffic theory (ISTTT), Toronto 1981, University of Toronto Press.
[76] M. Cremer, Determining time dependent trip distribution in a complex intersection for traffic responsive control, in: Proceedings on Control in Transportation Systems, IFAC-IFIP-IFORS, Beden-Beden, Germany, 1983.
[77] K. Ashok, M. Ben-Akiva, Dynamic O–D matrix estimation and prediction for real-time management systems, in: Proceedings of the 12th ISTTT, 1993, pp. 465–484.
[78] M. Ben-Akiva, T. Morikawa, Methods to combine different data sources and estimates OD matrices, in: Proceedings of the 10th ISTTT, Cambridge, 1987, MA, pp. 459–481.
[79] J.T. Lundgren, A. Peterson, S. Tengroth, Methods for pre-adjusting time-dependent origin–destination matrices, in: Proc. 10th World Congress and Exhibition on Intelligent Transportation Systems and Services, Madrid, Spain, November, 2003, pp. 16–20.
[80] J.F. Harrison, F. Stevens, A Bayesian approach to short-term forecasting, Operational Research Quarterly 22 (1971) 341–362.
[81] I. Okutani, J. Stephanedes, Dynamic prediction of traffic volume through Kalman filtering theory, Transportation Research B 18 (1984) 1–11.
[82] P.C. Vythoulkas, Alternative approaches to short term traffic forecasting for use in driver information systems, in: Proceedings of the 12th ISTTT, Berkeley, 1993, pp. 485–506.
[83] N.-E. El Faouzi, Nonparametric traffic flow prediction using kernel estimator, in: Lyons, J.-B. Lesort (Eds.), Proceedings of the 16th International Symposium on Transportation and Traffic theory (ISTTT), Pergamon, Elsevier, 1996, pp. 41–54.
[84] W.J. Granger, Combining forecasts: twenty years later, Journal of Forecasting 8 (1989) 167–173.
[85] N.-E. El Faouzi, Combining predictive schemes in short-term traffic forecasting, in: A. Ceder (Ed.), Proceedings of the 17th International Symposium on Transportation and Traffic theory (ISTTT), Jerusalem, 26–28 July 1999, Pergamon, Elsevier, pp. 471–487.
[86] M. Cremer, S. Schrieber, Monitoring traffic load profiles with heterogeneous data sources configurations, in: Lyons, J.-B. Lesort (Ed.), Proceedings of the 16th International Symposium on Transportation and Traffic theory (ISTTT), 1996, Pergamon, Elsevier, pp. 41–54.
[87] J. Sau, N.-E. El Faouzi, A. Ben Aïssa, O. De Mouzon, Particle filter-based real-time estimation and prediction of traffic conditions, in: Proceedings of the ASMDA, 29 May 2007 to 1 June 2007, CHANIA, CRETE, Greece.
[88] D.B. Choi, On multisensor data fusion using attributes association for intelligent traffic congestion information inference, in: Proceedings of the I.V.H.S., 1989.
[89] M.S. Grewal, L.R. Weill, A.P. Andrews, Global Positioning Systems, Inertial Navigation, and Integration, John Wiley, 2001.
[90] M. Wei, K.P. Schwarz, Testing a decentralized filter for GPS/INS integration, in: Proc. IEEE Position Location and Navigation Symposium, Las Vegas, NV, March 1990, pp. 429–435.
[91] W. Li, H. Leung, Constrained unscented Kalman filter based fusion of GPS/INS/digital map for vehicle localization, in: IEEE 2003 International Conference on Intelligent Transportation Systems, Shanghai, China, October 2003, pp. 1362–1367.
[92] N. El-Sheimy, K. Chiang, A. Noureldin, The utilization of artificial neural networks for multi-sensor system integration in navigation and positioning instruments, IEEE Transactions on Instrumentation and Measurement 55 (2006) 1606–1615.
[93] R. Sharaf, A. Noureldin, A. Osman, N. El-Sheimy, Implementation of online INS/GPS integration utilizing radial basis function neural network, IEEE Aerospace and Electronic Systems Magazine 20 (2005) 8–14.
[94] A. Hiliuta, R. Landry, F. Gagnon, Fuzzy corrections in a GPS/INS hybrid navigation system, IEEE Transactions on Aerospace and Electronic Systems 2 (2004) 591–600.