Current State and Future Outlook of Traffic Data Fusion in London *

Jun Hu, Ioannis Kaparias, and Michael G. H. Bell

Abstract — **Metropolitan areas today have become more than ever saturated with various types and sources of real-time data. Yet, the unsolved practical challenge how to most effectively combine data sources currently prevents the wide use of this data as a powerful tool to both improve the quality of the transport supply and to influence travel demand. Focusing on London, this paper investigates the current state and attempts to give an outlook into the future of traffic data fusion in dense urban network environments. Successes and gaps in the current state are identified, and extensions are proposed, along with respective deployment scenarios and impacts assessment.**

I. Introduction

The provision of traffic and travel information has long been at the center of development of Intelligent Transport Systems (ITS). Nevertheless, in complete contrast with the prevailing data scarcity that affected the transport sector for many decades, metropolitan areas today have become more than ever saturated with various types and sources of real-time data. Data sources include in the first instance the "traditional" transport-related ones, such as public transport service providers, road sensors and Automatic Number Plate Recognition (ANPR) cameras, but these are increasingly being complemented by a wide range of new information sources, such as mobile devices, web-based platforms and social networking services.

The vast amounts of available data have the potential to improve the quality of the transport supply, through the provision of more efficient and reliable services, but also to influence travel demand, through offering reliable real-time information to travelers, assisting them in their travel choices. A practical challenge that arises, though, is how to make best use of this potential, i.e. how to most efficiently compile and aggregate the various data into a common database on a real-time basis, in such a form that applications and users can access relevant information with appropriate representation and level of detail.

The solution lies in the concept of data fusion, which entails combining data from multiple and diverse sources with the goal of extracting new knowledge and producing better quality of information, estimates and predictions, as illustrated in Figure 1. Data fusion in metropolitan areas faces two main challenges. On one hand, there is the technical challenge, reflecting the issues associated with gathering the data in a timely and consistent manner and computationally manipulating it for different user groups. On the other hand, there is the organizational challenge, portrayed by the difficulties arising from the large number of public and private data providers involved and from issues such as data ownership, financing, privacy, etc [2].

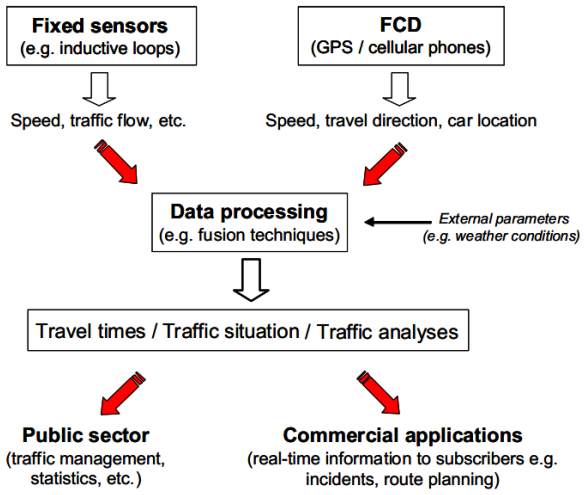

Figure 1. Schematic view of traffic data fusion [1]

The aim of this paper is to identify the role and needs of cooperative ITS in the field of traffic data fusion in dense urban networks. The paper uses London as a case study to analyze specific potential cooperative and non-cooperative ITS technologies that could be used in data fusion, as well as to investigate their feasibility of implementation in the near future. To this end, a primary research approach is adopted, such that existing literature is complemented by interviews with a purpose-assembled panel of experts from industry, academia and local transport authorities, actively working in the fields of urban traffic management and ITS. The views of the experts and the corresponding existing literature are combined with the authors' own views, such that an overall insight into the state-of-the-art and future outlook of traffic data fusion is given.

The paper is structured as follows: Section 2 gives a

* This work has been carried out as part of the NEARCTIS (Network of Excellence for Advanced Road Cooperative traffic management in the Information Society) project, which is funded by the European Commission under the 7th Framework Program for Research and Development.

Jun Hu is a Research Associate at the Centre for Transport Studies, Department of Civil and Environmental Engineering, Imperial College London, London SW7 2BU, United Kingdom. (phone: +44-20-7594-6086; fax: +44-20-7594-6102; e-mail: jh205@imperial.ac.uk).

Ioannis Kaparias is a Lecturer at the Collaborative Transport Hub, School of Engineering and Mathematical Sciences, City University London, London EC1V 0HB, United Kingdom. (phone: +44-20-7040-8746; e-mail: kaparias@city.ac.uk)

Michael G. H. Bell is a Professor at the Centre for Transport Studies, Department of Civil and Environmental Engineering, Imperial College London, London SW7 2BU, United Kingdom. (phone: +44-20-7594-6091; fax: +44-20-7594-6102; e-mail: mghbell@imperial.ac.uk).

physical description of the London site and reports on the various traffic management and information services available. Section 3 then provides an analysis of the current state of traffic data fusion in London, and describes successes, deficiencies and potential improvements. Section 4 goes on to identify the future outlook, including proposed extensions and advances, deployment scenarios and expected impacts. Finally, Section 5 summarizes the key outcomes of the study and concludes the paper.

II. Physical Description of the Site

Greater London is Europe's largest urban area with more than 7.3 million inhabitants and 21 million road journeys taking place every day. Transport for London (TfL) is responsible for managing London's most important roads (580 km of strategic road network, a.k.a. "Red Routes", with the remaining local roads being the jurisdiction of the 33 London boroughs), the public transport system and the congestion charging scheme on a day-to-day basis. It is also responsible for providing information about the administered services and their performance to both the public and the government [3]–[4]. Ensuring the smooth operation of this network and providing all road users with reliable and accurate traffic information is a challenging task.

A range of traffic information services are offered to drivers in London, information for which is mainly derived from the London Traffic Information System (LTIS). LTIS is a real-time database used in the London Traffic Control Centre (LTCC), to operationally manage planned and unplanned road network disruptions. A web-enabled interface offers the media and other stakeholders a live update of events, roadworks, incidents and accidents (Figure 2).

TfL also has a collaboration with TrafficMaster, the Automobile Association (AA) and INRIX to provide real-time traffic information to the public. Various commercial companies also use LTIS to populate their own traffic information services. Radio and in-vehicle systems are the main means of dissemination of on-journey traffic information, whereas the TV, radio and the internet are popular for journey planning. Figure 3 shows the general framework of traffic operations and systems in London.

The scale of London's traffic management system is large. TfL controls all 6,000 sets of traffic signals in London, and 50% of them are dynamically controlled by the SCOOT (Split Cycle Offset Optimisation Technique) system. There are 1,900 Automatic Number Plate Recognition (ANPR) monitoring and enforcement cameras on the London network, as well as 135 Variable Message Signs (VMSs) [5].

In theory, the range of monitoring sensors present in the central London area is sufficient to develop an accurate real-time travel information system for users. However, the underlying challenge remains, how this information can be most effectively captured, processed and delivered in a format that users can easily interpret and make decisions upon.

III. Successes, Gaps and Potential Improvements

The provision of public transport information in London can be identified as the most important success in the field of traffic data fusion, especially through the functionalities offered by London's iBus Real-Time Passenger Information (RTPI) system and by the Barclays Cycle Hire scheme.

With respect to iBus, this is an Automatic Vehicle Locating (AVL) system tracking London's 8,000 buses on 700 routes, and providing real-time passenger audio/visual announcements about bus arrival times and triggering traffic signal priority (TSP) at junctions [4]. iBus data was made publicly available in 2011 to other applications, such as

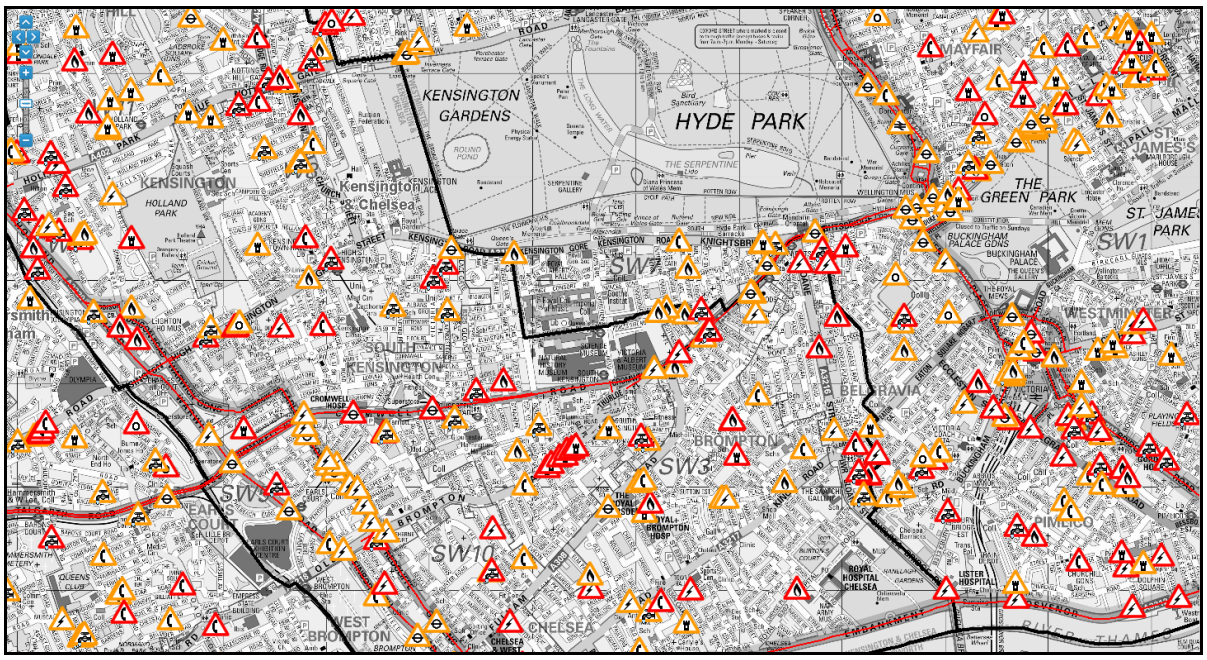

Figure 2. TfL public web interface [5]

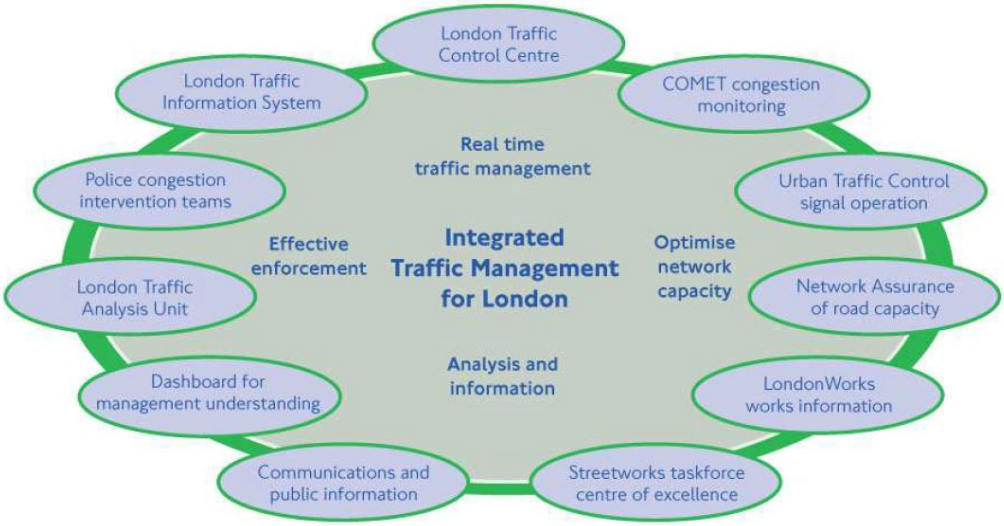

Figure 3. An overview of the traffic operations and systems in London [6]

mobile phones. Namely, it became possible for users to access arrival time information at any bus stop on their mobile phone, in such a way that they could plan their journey even before arriving at the stop. The service also became available through text messaging, such that sending a stop code by SMS enabled receiving real-time bus arrival information at that stop. Furthermore, iBus also provided an additional source of real-time traffic information for estimating road network conditions through more accurate bus location and speed information.

On the other hand, the Mayor's 'Cycling Revolution' for London has brought the successful Barclays Cycle Hire scheme to Central London. Since the launch of the scheme in 2010, a sharp increase in the numbers of people cycling in London has been recorded, expressed through an overall rise by 15% of the number of cycles counted on TfL's Road Network (TLRN) in 2011 [7]. By the end of 2010, more than 130,000 people had become members of Barclays Cycle Hire and around 25,000 journeys were made by hired bicycle users every weekday, the vast majority of whom would have previously not cycled. Real-time information on bike availability at docking stations was made available to users through mobile phones and internet. Currently, users have instant access to this information and are able to plan their journey accordingly, as illustrated in Figure 4 [7].

In contrast, the most important deficiency in the current situation of traffic data fusion in London is the fact that while the accuracy and quality of the data is good on the major roads, there are significant data inconsistency problems on the minor roads, especially with respect to roadworks and disruption events. Most of the minor roads are managed by the local boroughs and are not covered by TfL's monitoring

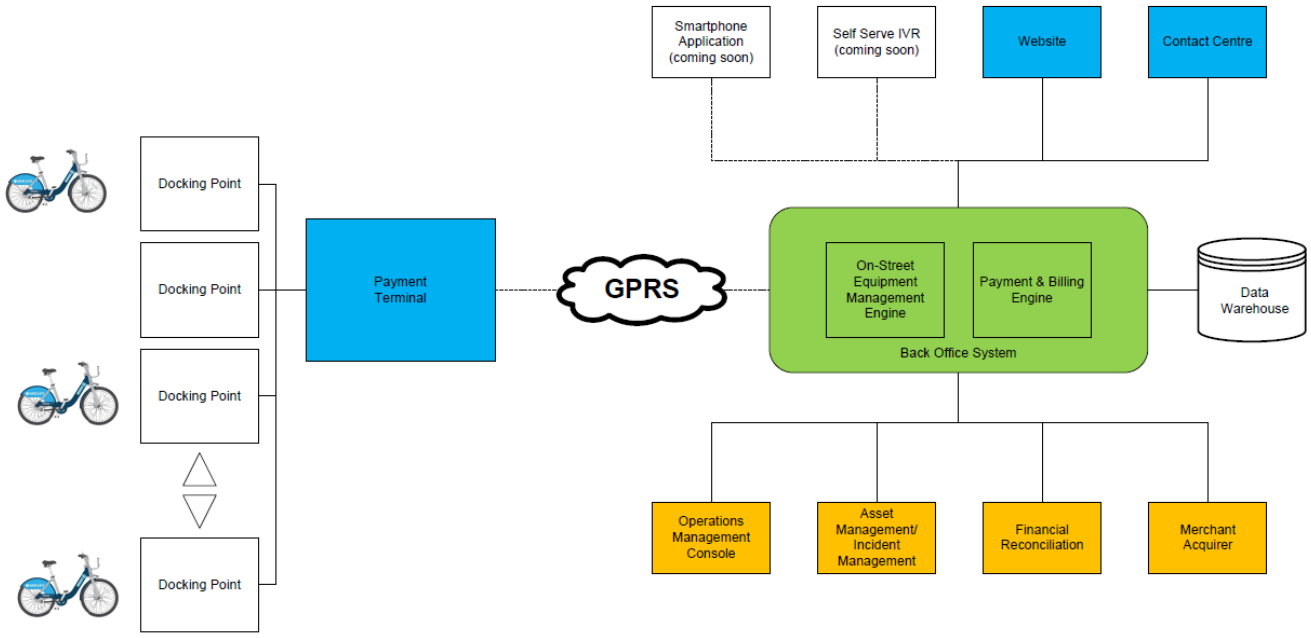

Figure 4. London cycle hire scheme [7]

system. As such, traffic information on these roads is unreliable and inaccurate, and better coordination between the boroughs and TfL is required in order to improve data quality, aside from carrying out regular consistency checks when drawing data from different sources.

Another deficiency that can be identified is the lack of a good parking guidance system (PGS) in London. While a number of ad-hoc electronic signboards are present, informing drivers about availability in some car parks, these are not positioned at strategic locations and are currently not linked to London's central traffic management system. It would be extremely useful if this information could be integrated into existing systems.

Finally, a further deficiency that can be pointed out is the limited understanding of end users' needs. The important questions needing to be answered are what types of information do the users want, how will users actually respond to the information generated and made available through data fusion systems, and whether they will utilize the information to make better travel decisions. There is clearly a gap under the current traffic information provision process to understand the users' requirements.

IV. Future Outlook

A. Proposed Extensions and Advances

There are many possible extensions and advances that can be proposed through existing cooperative ITS, five broad areas of which are highlighted.

1) Better integrated traffic management system

Traffic control and management in real-time is a complex task, especially in dense urban networks like London. The operation of an integrated traffic management system is the key to the smooth running of traffic on the network and the provision of accurate and reliable information to users. Traffic management involves many functions, including network monitoring, traffic signal control, incident management and public transport operation; however, each one usually involves a different agency, which often results in a lack of coherence between the various functions. For example, the London Traffic Control Centre (LTCC) is the main control centre for monitoring and managing TfL's road network in real time. Within TfL, though, there are many agencies providing information and working closely with the LTCC, such as London Buses, London Traffic Analysis Unit etc. It is crucial to ensure that these agencies share data and/or functionalities across the departments, as a first step towards better data fusion.

2) Better use of mobile sensing data

The majority roads in London managed by TfL are well-covered by sensors and ANPR cameras. Data collected through these monitoring devices can be used to derive journey time information on the network, as well as for analyzing network disruptions. However, the current monitoring network does not offer enough "granularities" in the data (i.e. spatial and temporal resolution of the data), and there is no monitoring of flow and speed data from many roads. ANPR data in London is temporally good, but spatially poor. It would be extremely valuable to gather some traffic data from roads that are not well-covered by the ANPR and SCOOT systems, and since nowadays most of the vehicles in London are equipped with some kind of mobile device (e.g. driver's mobile phone), they have the ability to act as anonymous traffic probes. The positions of mobile phones are regularly transmitted to the network by means of triangulation or by other techniques, such as handover, and then travel times and further data can be estimated over a series of road segments. This data source provides a potential cost-effective alternative to the expensive conventional monitoring systems. Some research has already taken place by TfL and by some mobile operators, as well as by other relevant stakeholders (e.g. Google) in order to explore methods to extract traffic information from that data source. Nevertheless, a number of technical difficulties have been reported so far, demonstrating a strong research potential in the field. Examples include the issue of positioning accuracy, the problem of the position of the user being lost for a period of time due to unavailability of transmission towers in certain areas, and the difficulty of distinguishing the mode of the road user and of filtering vehicle drivers from other transport users.

3) Better integration between users and transport operators

One of the biggest challenges for transport operators is to understand the needs of end users and to integrate them into traffic management systems. This means that they have to find the best way to not only present meaningful information to the users, but also to anticipate how they will respond to this information. The existing traffic information provision model is based on data providers collecting, processing and publishing data, which then end users receive and react accordingly. Recent advances, however, mean that pervasive computing environments have the ability to change the role of end users from traditional information receivers to information providers, both in terms of supplying actual data and of participating in generating different ways of using the data, such as developing smartphone apps. This creates a platform to engage the interaction between the transport users and the traffic management systems, and has the potential to transform the transport decision making process from "top-down" to "bottom-up".

4) Better integrated journey planning tool

Nowadays there are many different types of journey planners available to travelers in London through the internet and mobile phone applications. The capability of these journey planners is very limited, however, as it is only restricted to users planning their journey at one point of time, usually at the start of their journey, and it is almost impossible to use them when an unanticipated incident occurs along the trip and users decide to change their original travel plan. Therefore, it would be extremely useful to create an integrated trip planner, which could run on a hand-held device and provide real-time guidance to the users en route. In case of an incident, the journey planner would be able to suggest alternative routes or different modes of transport, and also provide the timetables for the relevant public transport information. The information required for such a task is gradually becoming available, but it needs an intelligent algorithm to make use of it, as well as a simple

design so that everyone can use it. It is envisioned that such an application will appear in the near future, though there is still a long way until travelers become able to just "speak" their destination to their mobile device and let it do the rest and guide them with all the real-time travel information in hand.

5) Better dissemination of traffic and travel information

Traffic information is currently broadcast via the Traffic Message Channel (TMC), which, however, is due to be replaced soon by a recently developed new standard format for delivering real-time traffic information, called TPEG (Transport Protocol Expert Group. In contrast with TMC, TPEG takes advantage of high bandwidth in digital radio broadcasting and can provide richer content of information and a wider range of services. For example, TPEG messages can provide much more detail and accurate description of incidents, so that users and other ITS applications can respond more efficiently, as well as information on weather and congestion (Figure 5). The better exploitation of TPEG through its integration in available dissemination platforms (e.g. navigation systems, smartphones, websites, etc) can be an important building block for the successful deployment of cooperative ITS systems.

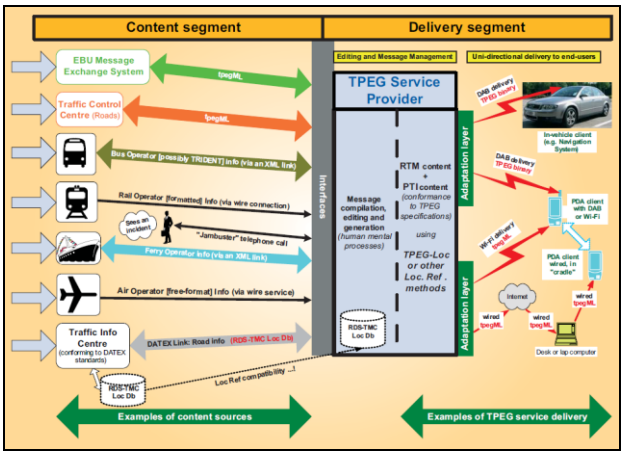

Figure 5. TPEG service diagram [8]

B. Deployment Scenarios

Considering the proposed extensions and advances described deployment scenarios with respect to cooperative systems are conceived in this section. Cooperative systems provide the capability to extend internet-based information sharing into the mobile environment, which can open a large market for third party/end users to make use of publicly available traffic information and create different user-oriented APIs (Application Programming Interfaces). In fact, as TfL has a policy of releasing raw data to the public, end users are likely to be encouraged to participate more actively in the traffic data provision process. On the other hand, traffic operators can more effectively focus and utilize their resources on providing better quality real-time traffic data, while letting the users themselves or private third-party developers to decide how to use it. This is a possible deployment scenario that can bring transport operators, transport users and private third party developers together to provide better traffic management solutions.

A second deployment scenario could be the use of mobile phone data or wireless ad-hoc network data as a potential source of information on road network conditions. This is particularly applicable for dense urban networks, such as London, due to the wider coverage of communication networks and the shorter distances between antennas. Through the development of intelligent algorithms to overcome the problems of distinguishing between different transport users and of computational complexity, transport operators will be provided with more spatial and temporal coverage of the network. However, a number of potential problems associated with using mobile phone data will still exist. These include data ownership issues (i.e. who will own the resulting large databases and whether should these be shared between the different stakeholders, such as mobile phone companies, local traffic authorities, etc), business viability issues (i.e. who will invest the resources to carry out the necessary data mining work, given that most data sources are a by-product of the mobile communications industry and raw data will have to be cleaned, and given that traffic authorities are reluctant to do it themselves), and privacy issues and user acceptance (i.e. the management of personal data related to traffic probes has to be addressed through clear policy messages to gain the people's trust).

A third deployment scenario could be the full integration of the journey planning tool. This would naturally require the integration of the various traffic management systems, such that, for example, the system can be used to warn about current or short-term predicted congestion and advise an alternative route. This system can be linked with other systems monitoring the availability capacity in car parks to advise drivers where best to park and hence minimize the mileage associated with looking for an available space. All this information can be integrated into a single journey planning tool and users can run this tool on their mobile devices on-the-go.

C. Expected Impacts

Traffic data is at the basis of any road traffic management application, so the proposed advances and deployment scenarios could have a significant impact on the transport system in London. Namely, the availability of larger quantities and higher quality traffic data will first and foremost improve the provision of transport services. It will be possible to obtain Origin-Destination data, which will make it possible to run real-time traffic assignment procedures and make a better use of the network capacity in the short term, and will offer valuable input to the planning of infrastructure and services in the long term. More accurate traffic usage data will also enable the application of advanced signal control strategies in all parts of the network (complementary to the SCOOT UTC system), ensuring smoother traffic flow and better incident management. More accurate travel information will further enable travelers to make more informed travel choices, either themselves or through their advanced journey planning applications, thus offering a better customer experience and increasing the reliability and satisfaction of the public transport system.

This could induce a modal shift away from private transport, with associated environmental, energy-efficiency and quality of life gains.

D. Socioeconomic aspects

Most social concerns relating to traffic data fusion are about loss of privacy and user acceptability. For personal data protection, common operational rules must be created and respected by service providers handling personal data. International standards should be developed to establish the basic principles for personal data protection in these services, as the lack of common standard procedures to all service providers might generate public distrust in this technology [9].

In terms of economic aspects, the development and maintenance of the systems for traffic data fusion could prove to be expensive [10], as the underlying technology evolved rapidly and there is a need to put in place organizational and financial structures that can work at the same pace. Otherwise there is a risk that maintaining obsolete technology will become a drain on resources [11]-[12]. On the other hand, data fusion technology also creates many new opportunities for business, as many companies may become involved in providing data processing services and application development.

V. Conclusions

Focusing on the development of cooperative ITS, computational models, and user applications that allow access to real-time information about the state of transport-related resources, this section investigated the future of traffic data fusion in dense urban networks, and in particular London. As a broad conclusion, it can be said that although much of the necessary technology exists, the use of data fusion in transport has yet to fulfill its potential, as both technical and organizational challenges remain.

A number of successes and deficiencies in data fusion in London were identified. Successes were recognized in the implementation of the iBus project and in the operation of the Barclay's Cycle Hire Scheme, primarily due to their ability to collect and make use of real-time information. Gaps, on the other hand, were mainly found in the lack of consistency in the data on minor roads, the lack of parking guidance information and the limited understanding of the requirements of end users.

Based on these findings, future avenues for extensions and advances were investigated, which were grouped in five main areas: 1) the operation of an integrated traffic management system ensuring that the different functionalities and applications share data between each other; 2) the resolution of the open questions relating to the use of mobile phone data as a potential source for data fusion, particularly as regards privacy and computational complexity; 3) the better uptake of the recent advances in computing technology, enabling the increased participation of the end users to the planning and operation of the transport system; 4) the creation of an integrated journey planning tool to improve the efficiency of the information dissemination from data fusion systems to end users; and 5) the full exploration of the potential offered by the TPEG technology to improve the dissemination of traffic and travel information.

Finally, the study identified three deployment scenarios for data fusion, relating primarily to data ownership, user acceptability, privacy concerns, and business viability. These issues require adequate attention before data fusion applications can reach their full potential.

Acknowledgment

The authors would like to thank the four expert interviewees, namely A. Emmonds, N. Hoose, S. Robinson and A. Stevens for kindly participating in this study and providing their valuable insight and suggestions.

References

[1] G. Leduc, "Road traffic data: collection methods and applications," European Commission paper on Energy, Transport and Climate Change, 2008.
[2] A. Amey, L. Liu, F. Pereira, C. Zegras, M. Veloso, C. Bento, A. Biderman, "State of the practice overview of transportation data fusion: technical and institutional considerations," MIT & CISUC, 2009.
[3] Greater London Authority, "Mayor's transport strategy," Government report, 2009
[4] TfL, " Travel in London, report 4," Transport for London, 2011
[5] TfL, "London RoadWorks," accessed on March, 2012 from: http://public.londonworks.gov.uk/roadworks/home.
[6] N. Barton, "Real-time traffic management systems and operations in Transport for London," Transport for London report, 2004.
[7] A. Stannard, G. Wolfenden, "Putting in place a new public transport system for London," Serco Company report, 2011.
[8] M. Bev, "TPEG – what is it all about," EBU technical review, 2005
[9] TfL, "Traffic Operations in London," Transport for London report, 2007.
[10] K. Cowan, B. Sultan, J. Burrow, "Real-time Fusion of UK Traffic Data," Road Traffic Information and Control Conference 2008, Manchester, UK.
[11] N. Hoose, "Future of Road Network Management and the role of Intelligent Transport Systems," 2009.
[12] TfL, "London Traffic System Vision," Transport for London, 2011