DECISION SUPPORT FOR INTELLIGENT TRAFFIC AND ENVIRONMENT MANAGEMENT

Margaret Bell[1*], Visalakshmi Suresh[2], Fabio Galatioto[1], P. Watson[2]
[1]**TORG, School of Civil of Engineering and Geosciences, Newcastle University, Cassie Building, Claremont Road, Newcastle upon Tyne, NE1 7RU, UK; +44 (0) 1912226547; margaret.bell@ncl.ac.uk**
[2]**School of Computing Science, Newcastle University, Devonshire Building, Kensington Terrace, Newcastle upon Tyne, NE1 7RU**

ABSTRACT

ITS is fundamental to safe, efficient, control and management of people, vehicles and transport systems and plays an important role in addressing environmental impacts. Recent advances in intelligent pervasive technologies provide a wealth of static and dynamic data which complement legacy ITS systems. There is significant value-added in the integration of existing and emerging technologies especially if achieved across jurisdictional boundaries. However, there are considerable challenges to realise the rewards of integration particularly in real-time. This paper offers an e-science solution providing insight into the delivery of a real-time data management platform for both monitored and modelled sources.

KEYWORDS
Intelligent Transport Systems, Information Management, UTMC, Congestion, Decision Support System.

BACKGROUND

Government policy throughout the world is pressing for improved air quality and reduction in carbon emissions. The responsibility for delivery of these goals rests with local authorities who will need to make tough decisions. The effectiveness of technical and policy decisions depend on the accuracy, timeliness and appropriateness of information being available for the technical manager and operator as well as for the public. Everyone needs to work together to deliver change. Intelligent Transport Systems (ITS) are proving to be valuable tools for managing congestion (UTMC and bus priority) and air quality hotspots (gating and metering of traffic to open space) as well as managing demand (road user charging, ramp metering). However, the value of ITS is two-fold because, as a bye product of its implementation, it provides a rich source of data.

Success in delivering the policy objectives rests with the public changing their travel behavior in Bell (2006) [1]. The research presented here acknowledges not only the value and importance of data available from legacy ITS systems but that the true potential can only be delivered the information is useful and available in real-time. As people and vehicle tracking systems are introduced, essential knowledge of user travel patterns and habits will begin to emerge providing *'marketing'* information to further improve the transport services and potentially achieve seamless integration across different transport modes in the future.

The Project (MESSAGE) Mobile Environmental Sensing Across Grid [2] jointly funded by the EPSRC and Department for Transport, UK, has developed an inexpensive environmental pervasive sensor system for urban road networks. The MESSAGE is a collaborative project between 5 UK Universities (Imperial, Cambridge, Newcastle, Leeds and Southampton) supported by 19 stakeholders representing the ITS profession. Newcastle University has developed wireless sensors which are fixed onto street furniture and carried by vehicles and pedestrians thereby collecting data at a very high spatial resolution. Every minute the captured data is sent, using wireless GPRS links in real-time, to the central server where more detailed analysis, processing and dissemination of information takes place. This forms the basis of a Decision Support System, DSS, and provides information to monitor the current status of the network as hourly, daily, weekly, monthly and annual average concentrations or for longer-term traffic and environmental impact assessments and policy formulation. A user-friendly interface has been designed to meet the needs of decision makers. Capturing, synchronizing, cleaning, manipulating and processing data presents huge challenges especially when information across large spatial networks is needed quickly. For example updating origin-destinations for traffic assignment; to assess congestion incidents; carry out signal optimization to improve network performance; relocate queues from closed to open space to enable the natural ventilation of the built environment to disperse pollution and thus prevent the build-up of pollution hotspots. The data capture has to be supported by quality assurance procedures; validity checks catering for missing values; statistically sound analysis and mathematical modelling.

This paper will describe the components of a DSS. In the section that follows, the design of the MESSAGE systems architecture will be presented. The functionality of the different components will be illustrated by examples. The data currently integrated in real-time includes pervasive sensors, referred to as "*smart motes*", as well as legacy monitoring stations and include, traffic, pollution and meteorological conditions. The data collected is processed and analyzed so that it can be used to drive applications in real-time to monitor network status and to control traffic to reduce congestion, improve air quality and manage noise impacts. The paper will be concluded with an indication of the planned next phase of the development in the DSS.

SYSTEM DESIGN

The initial design for a DSS framework involves conceptualization of the system elements at a high level. These include real-time monitoring; statistical analysis to build up a historic picture; modelling impact assessment and the decision support itself. The monitoring applications require high-performance and low latency whereas the historic data warehouse involves the analysis and statistical processing which can be computationally intensive. The design framework starts with a specification for the logical and functional model for individual components and communication between them in order to satisfy the basic requirements for environment monitoring. The interface between the components described in the framework are focused to be compliant with the UTMC [1] specification, described by the Department for Transport. For the architecture to adhere to UTMC standards, all communication across multiple data objects uses the technical standards as specified in the section UTMC TS003 section 4 – 8 [1]. The characteristics of the interfaces will reflect the ITS requirements with respect to volume of data, expected latency, security, resilience and operational management.

The key objective of the logical and functional design is to identify the levels of abstraction in the system design. The users and the database respectively perceive the data as the external and internal level respectively. The data is stored using file based organization and data structures as per the

vendors specifications. Identification of the structural, manipulative and integrity components of the data is described in the logical and the functional model. These are now described in more detail by way of actual applications to real-time simultaneous data capture from pollution monitors, traffic detectors and meteorological conditions. With reference to Figure 1 the logical reference model describes the system architecture as a series of interconnected objects. Individual components can either be legacy systems such as SCOOT, Split Cycle, offset Optimization Technique (TRL, Crowthorne), AURN, Automatic Urban Rural Network (DEFRA) and meteorological conditions and or emerging pervasive sensor technologies [10]. Data communication links are set up by using specific rules that govern the transfer of data from one point to another.
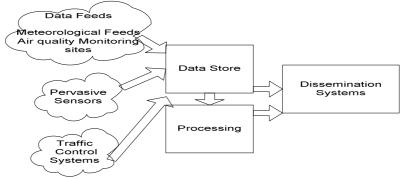
Figure 1: Logical Reference Model

Data Objects
The data objects groups are the elemental units of the logically related attributes of the information. For instance, the data object called pervasive sensing systems as in figure 1 consist of attributes such as carbon monoxide, nitrogen dioxide, noise, traffic flow, temperature and humidity. Temperature and humidity are measured within the pervasive sensor for calibration of the chemical sensors as well as by Met stations which measure actual ambient conditions. Although these measured attributes are the same they measure different environments for different reasons. Whilst they are treated within the logical reference model as identical data observations from multiple data objects (meteorological feeds, pervasive sensing), each observation made by the various devices is distinguished based on the device type and the time of measurement. The system is initially decomposed into several layers namely communications, application and information as illustrated in figure [2]. The functional model presents a narrative overview of the system describing the criteria for real-time processing; historic analysis; system interfaces and future extensions. The data flow is mapped to the specified context as identified in the requirements. The cross references to the input and the output channels from each layer is defined in the functional design.

Communication Layer
Whenever the pervasive sensors start the communication process from the gateway node [10], the data packets are received into the real-time database. The interface design is meant to be generic to accommodate heterogeneous data sources by changing the meta-data of the data packet. An example of the meta-data for the traffic sensors will contain descriptors such as *flow*, *occupancy*, *delay*, *queue length* etc with reference to the time. The communication layer segregates the data packets based on the meta-data description to stream the data into the respective application for further processing.

Application Layer
The application layer is designed to ensure the reliability of the system and more emphasis is given to

the performance requirements of the real-time monitoring systems. In general, the modern system architecture approaches the application design by separation into multiple layers in order to manage the impacts of evolution and changes in management over time. Aspect-oriented application layer design supports the identification, reasoning and description of each component of the user requirement to provide a better interface for the application and encode the knowledge of the individual components into loosely coupled layers within the application. For instance, the traffic sensors use serial communication from legacy systems to pass the data packets into the communication layer. The application layer transforms the data giving emphasis to the real-time delivery of the congestion monitoring requirement by specifying and ordering the data transformations and operands to deliver computational speed. The application layer is designed in a way to enhance the system by wrapping up the application into a web service.
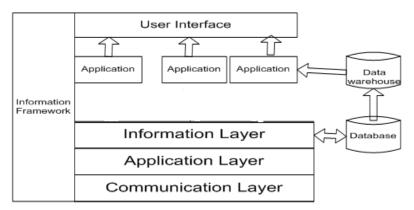

Figure 2: Layer Model

Information Layer
The environment and pollution monitoring system is truly data centric. The design of the information layer starts with analysing the format and quality of the data acquisition in real-time. The real-time data acquisition considers the number of individual transactions to be handled and processed over time. Exactly how the data is managed; and the order in which the steps in the data management/computation are carried out; change and expand with the number of on-line transactions and lead to the integrity of the data over time. Irrespective of the acquisition system whether traffic and /or pollution sensor or a model, the data needs to be integrated with respect to both time and space. In real-time, the synchronisation of the data is handled in this layer before depositing the data into the database or the warehouse. As a single entry point to the data, implementation of data services is introduced in this layer to achieve seamless access to the integrated data sources by consumers. The various business rules used to achieve the data transformation are referenced and described in this layer. With a discrete set of rules precisely specified in the information layer, the data is transformed to reduce the complexity and therefore increase the speed of the information processing.

User Interface
The interface design reflects the functionalities to which the information is exposed in a given system to provide a serialized, simple and appealing layout. The design of the user interface starts with the identification of the features to be presented in the given frame. A huge range of choices, in terms of technology selection, appearance and accessibility, are available to deliver an effective presentation of the information. This paper focuses on the design of the user interface for three features namely pollution hotspots, critically congested links and associated statistics.

Database Warehouse

The main aim of the database is to provide the users with an abstract view of the data, hiding the process involved in manipulating and storing the data. The data warehouse handles the linearly increasing data volume and acts as a central repository to archive the data from the real time transactional data store after a specified time interval. The real-world objects are described as entities. The properties of the objects are identified as the attributes of the entities. The relationships between the entities are identified and designed as *foreign* keys. The relational and dimensional design facilitates the transactional and analytical querying capabilities of the system respectively. The historic data warehouse is modelled to achieve an optimised response time for the analytical queries namely *"find a peak hour CO in a particular location for the past 10 years"*. The question posted in the historic warehouse should be balanced across the server load and processers available in the hardware to achieve the optimised result. The design of the system plays a major role in achieving the response time for any analytic queries. In MESSAGE project uses the historic data from Leicester made available from the *i*nstrumented *C*ity facility (ITS, Leeds University) and provides use-case to measure the performance of the e-Science system architecture. The linear growth of the dataset poses another big challenge in the data warehouse.

ARCHITECTURE FOR THE MESSAGE DECISION SUPPORT SYSTEM

The architecture for the Newcastle DSS within the MESSAGE project is illustrated in Figure 3. There are three main components which are integrated by the database and sit at the heart of the DSS. These are real-time data infrastructure, data display marts and modelling. The data capture, based on the pervasive sensors and legacy systems, is requirements-driven. The strategy in MESSAGE was to classify and bifurcate the database into a Real-Time component, which is UTMC compliant, and the Data Warehouse. The raw data from various traffic data sources such as SCOOT, loop detectors, accidents, car parks, traffic signals; pollution data from the AURN, Automatic Urban and Rural Network; meteorological conditions (wind speed and direction, temperature etc) from weather stations and all three types of data from the pervasive sensors, were stored in the UTMC standard way.

The data display marts are driven by the user requirements of the decision maker and make use of the historic and real time data sets as well as the modelling components of the DSS. The warehouse design is meant to address the main problem of Extract, Transform and Load the system from the different data sources. The reporting requirements for the warehouse grow over the years and this issue is addressed by harnessing user needs in the form of data marts. Thus the system evolves and is fine-tuned based on the users' needs. In MESSAGE, the measured data (e.g. Sensors), simulated data (e.g. AIMSUN) and predicted data (e.g. AIRVIRO) were integrated to produce the fine grain report of the measurement system in a very high spatial resolution. Each component of the DSS is elaborated in the remainder of this paper.

Real-Time Data Infrastructure

The real time system is designed based on the UTMC specification, reviewing each step of data extraction, filtering and calibrating to verify the functionality, flow and management of data. The significant real-time data integration issue such as assigning the traffic flow measured on each link during a signal cycle to the correct stage of the signal sequence at the junction into which the traffic is discharged is added as a specialised service in the real-time data capture architecture. The real-time architecture concentrates on functions, business rules and definitions for the incoming streaming data sets which are specific for each ITS. The interface between the pervasive sensors, legacy systems, both

static and dynamic is handled using the UTMC specifications. Enhancements to the UTMC specifications have been needed in the MESSAGE work to capture the dynamic sensor and emerging technologies such as the GPS enabled mobile phone system to capture the vehicle characteristics data to estimate the carbon foot print of the journey.
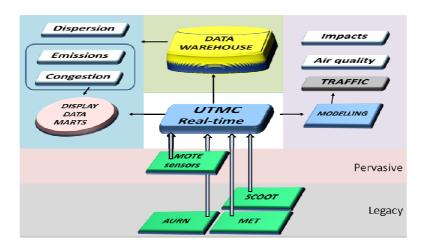

Figure 3 – MESSAGE Systems Architecture

The specification for the real-time system starts with the meta-data definition for the measuring devices such as mote, SCOOT and AURN. The meta-data defines parameters such as the generic features of the measurement device. For example, the meta-data for the static pervasive sensors will include details such as installation location, nature of the measurement device, sampling frequency and height of the measurement from ground level. The thresholds for each measurement system is specified in the measurements configuration section as drafted in the UTMC [1] specification. The legacy systems and the pervasive sensors data need to be synchronised in time and associated with each other spatially by using geographical co-ordinates (northings and eastings) in order to deliver data consistency for the dissemination applications such as a website. The real-time data have a very high flow rate and the volume of data received every second is high. In a few seconds, a raw data will exceed hundreds of thousands of records for a particular region. Therefore the analysis of the data in real-time requires various aggregations, groupings and comparisons. Statistics, such as a rolling, 8 hourly moving average and logarithmic average of noise measurements need to be continuously updated within real-time architecture. This statistical processing in the real-time architecture forms the basis for the local authorities to make queries of the information management system: for example comparison of the AURN with the mote data or the current status with the *'typical'* profile based on the analysis of the historic data within the Data Warehouse. These data operations are delivered using data marts. Some examples will now be given.

Data Display Marts and Applications
The architecture for the data marts is derived from the transport specific requirements. This defines the boundary of the analysis, reports and the questions asked by different users. Data driven subscriptions for the different systems such as pollution and traffic are created as a specialised data mart. Some examples of data display marts for specific applications for the decision support will be described here.
- *Congestion Assessment, CA:* The performance of a network is delivered by algorithms that process SCOOT occupancy and flow data to assess different traffic conditions or flow regimes into states defined as free flow, busy and congested. In real-time these states are displayed onto a Google map.

- *Traffic Tail-pipe Emissions Estimation, TTEE:* Emissions algorithms estimate the pollutant levels for carbon monoxide, nitrogen dioxide and carbon dioxide. These can also be displayed as actual numerical levels. The congestion states and emissions estimates are fed into the data warehouse.
- *Receptor Point Pollutant Levels, RPPL:* This data mart extracts the meteorological conditions data from the real-time database and uses the Canyon dispersion model (OSPM) along with the estimates of pollutant emissions (output from the TTEE data mart) to provide levels of the pollutant concentrations at the location of the mote.
- *Real-Time Event Processor, REP:* The network status data accumulated within the most recent few minutes is compared with the historic profiles to identify, in real-time, incidents, unusual features and events occurring in the network. The process of storing the information every few seconds or minutes and consequently querying the database to retrieve the information will have considerable impact in the performance of the database. Novel methods of processing the real-time events have been designed.
- *Critical Congested Link Assessment, CCLA:* The *CA* measure from various data sources such as SCOOT, pervasive sensors and loop detectors monitored in real-time are compared with the average or *typical* congestion profile obtained from the data warehouse. This is based on statistical analysis of the historic data stored in the Data Warehouse to determine when and where the congestion levels exceed a threshold specified by the traffic signal control operator. To achieve this for a city, millions of records of data need to be scanned in seconds. In MESSAGE, we have implemented a prototype using Oracle Sensor Edge Server which acts as a middleware to handle the real-time event and also integrates the real-time processing of diverse data sets.
- *Visualisation:* The basic data from legacy and pervasive sensors can be visualized according to the user specification in different ways. In MESSAGE, the main focus was on creating spatially rich applications overlaid on the Google Maps. There are plenty of standards such as Sensor Web Enablement Framework, Google focus, to display the information in a web centric manner to explain the functionalities of the system. The system has the capabilities to wrap the data in to any open standards or vendor specific standard.

Modelling

There are two types of modelling which support the DSS, *mathematical* and *simulation*. Depending on the application machine learning techniques such as Artificial Neural Networks, ANN, and cluster analysis techniques such as K-means can be used to establish patterns in the traffic and pollution data. The *mathematical* models that result may be used for forecasting the changes to be expected. This is of vital importance to traffic control operators because early warning of potential network problems enables appropriate management strategies to be implemented giving rise to pro-active rather than reactive control. *Simulation* on the other hand provides the framework to allow strategies to be designed to resolve specific network problems. The monitored data initially is used to validate a model of the *current state* of the network commonly referred to as the base-case. The REP and CCLA define a problem which requires a particular solution: a traffic management or control strategy known as a *scenario*. In the proposed architecture, the data described in the scenario is specified in the data marts and fed into the traffic model. This is then used to predict the traffic flow, delays, stops and congestion expected when the scenario is implemented. In this way, a library of intelligent transport system plans to be implemented in response to the specific problem for which they have been designed is created and stored in the data warehouse. The modelling involves mode split, traffic assignment, traffic signal optimisation, air quality, noise and carbon emissions modelling.

Impact Assessment and Evaluation

Other components of the DSS include *impact assessment* and *evaluation*. *Impact assessment* is

delivered by comparing the changes in the measured parameters *before* and *after* the implementation of signal control plans, management strategies or changes in policy interventions. A module of the data warehouse provides a continuous assessment of the performance of the network to changes. *Evaluation* is a component of the scenario modelling. For each strategy, designed to solve a particular network problem, an estimate of the impacts is derived. These data are deposited into the Data Warehouse so that when the strategy is implemented then the predicted changes can be compared with the actual change and thus to provide an evaluation of the effectiveness of the strategy and/or accuracy of the strategy. This feedback enables the credibility of the DSS system to be continually challenged and therefore, with appropriate modifications, to be fine-tuned over time in response to change. As the system knowledge improves so does the quality of the decision making.

CONCLUSION

System architecture is of utmost importance in the design of real time monitoring systems. The system is designed for data intensive applications while features are considered to accommodate the continuous evolution in intelligent transportation systems. Current methodology describes the experiences in the framework design, implementation and results. The business policies are met according to the stakeholders needs. The integration of the multiple systems reveals information about the effective control and added value to the Intelligent Transport System implementation.

ACKNOWLEDGMENTS

REFERENCES
[1] Urban Traffic Management Control – UTMC http://www.dft.gov.uk/pgr/roads/tpm/utmc/
[2] MESSAGE – Mobile Environmental Sensor System Across Grid www.message-project.org
[3] The Data warehouse Toolkit – Ralph Kimball
[4] The identification and control of recurrent urban traffic congestion using SCOOT data – thesis submitted for the degree of Doctor of Philosophy by Raymond G Evans
[5] Estimation and prediction of roadside pollution using real-time data in an air quality model - Thesis submitted for the degree of Doctor of Philosophy by Santosh A Jalihal
[6] Airviro - http://www.westlakes.org/live/html/env/airviro.htm
[7] AIMSUN - http://www.tdg.co.nz/index.php?a=2&b=5&c=3
[8] Oracle - http://www.oracle.com/database/index.html
[9] Bell M.C. Suresh V., Watson P., Blythe P.T. (2008a) UTMC Compliant Database to support technologies of the future. RTIC Conference 20th 22nd May 2008 Manchester
[10] Bell M.C., Neasham, J.A., Suresh, V., Rhys-Tyler, G., Sharif, B., Watson, P., Blythe P.T., (2008b) Inexpensive Pervasive Sensors for Traffic and the Environment Monitoring – Future Network Intelligence. JCT Conference, Nottingham University Jubilee Campus, 17th – 18th September 2008
[11] Bell M.C., Neasham, J.A., Hill G., Tate J.E., (2008c) Validation of carbon monoxide and nitrogen dioxide sensors. Internal Report, Transport Operations Research Group, Newcastle University, November 2008
[12] Bell M.C., Galatioto F., Hill G., Mohammad A., Automatic Evaluation of Environmental impacts of traffic DEMAND management Strategies (TDMS), UTSG, Jan 2009