A real-life test bed for multi-agent monitoring of road network performance

Jos Vrancken* and Michel Dos Santos Soares

Sectie ICT, TBM, TU-Delft
Jaffalaan 5
2628 BX Delft
P.O. Box 5015
2600 GA Delft, The Netherlands
E-mail: J.L.M.Vrancken@tudelft.nl
E-mail: M.DosSantosSoares@tudelft.nl
*Corresponding author

Abstract: This paper describes an approach to road network monitoring. The approach has the following main components: a multi-agent hierarchical monitoring system in which the agents are network elements such as crossings, road segments and routes, software engineering techniques to handle legacy sensors and to homogenise their heterogeneous data streams, distributed Kalman filtering to reduce sensor noise, and a traffic simulation model to help in predicting future traffic states and to replace missing physical sensors. The approach has resulted in a real-life test bed for network performance monitoring in which various configurations of agents and various ways of distributed Kalman filtering can be tested. The current, still rather simple monitoring configuration, is already in operational use for traffic management.

Keywords: road traffic monitoring; Kalman filtering; multi-agent systems.


Biographical notes: Jos Vrancken received his BS and MS degrees in Mathematics from the University of Utrecht, The Netherlands in 1979 and 1982, respectively his PhD in Computer Science from the University of Amsterdam, The Netherlands in 1991. Currently, he is an Assistant Professor in the Use of IT in Infrastructures at Delft University of Technology, The Netherlands.

Michel dos Santos Soares received his BS and MS degrees in Computer Science from the Federal University of São Carlos, Brazil and the Federal University of Uberlândia, Brazil in 2000 and 2004, respectively. Currently, he is employed as a PhD student in Software Engineering at the Delft University of Technology, The Netherlands.
1 Introduction

In most countries with congested roads, traffic management is considered an indispensable tool to maximise the capacity of the road infrastructure and to increase traffic safety. Traffic management is best known for a number of essentially local measures such as traffic signals at crossings, ramp metering on motorway entry lanes, and variable message signs with warnings or dynamic speed instructions. However, local measures have obvious limitations. For instance, solving congestion at one place often results in shifting the congestion to some other location. Local measures can only be reactive. Proactive traffic management, i.e., preventing problems rather than solving them, has to look into the future. In a network with moving vehicles, this means that one has to consider a larger part of the network than what local measures currently do. In addition, some societal goals in traffic management, such as improved travel time reliability, are network-related. These considerations have made network-level traffic management into one of the main challenges in the field of traffic management (Taaele et al., 2004).

Figure 1 The belt road around Alkmaar (see online version for colours)

An indispensable component of traffic management at the network level is the monitoring of relevant network quantities, for instance those mentioned in traffic management policies. This, in itself, is already a major challenge, especially in a real-life setting in
which one has to deal with the existing sensors that were not designed to be part of a
sensor network for network performance monitoring. The real-life setting is more or less
indispensable. Simulation does play a role, but the value of laboratory findings can only
be assessed by a comparison with real-life traffic data. This holds for local quantities but
even more so for a network-level quantities as the latter apply to an essentially more
complex object and, therefore, harder to model reliably.

This paper gives an overview of the problems encountered in network monitoring,
making a distinction between traffic engineering, signal processing, institutional and
software engineering aspects, and describing the interdependencies between these
aspects. Then an approach is formulated to obtain an effective network monitoring
system. This problem is far too big to allow definitive solutions, so the essence of the
approach is to create a real-life test bed to support the research in network monitoring.
The approach has been implemented in an operational traffic management system on the
belt road of Alkmaar (see Figure 1), a medium-sized city in the Netherlands (Vrancken
et al., 2007). The implementation is such that various ways of network monitoring can be
tested by rapid reconfiguration of software components.

2 Problems in network monitoring

There are a number of problems making the measurement of network quantities
essentially more difficult than the measurement of local quantities such as traffic speed
or intensity at a single point. These problems can be categorised as either traffic
engineering, institutional, signal processing or software engineering problems.

2.1 Traffic engineering problems

2.1.1 Definition of network performance

A network allows a far greater number of different quantities to be measured than a
single point or even a road segment or crossing. The relevant quantities for network
performance depend on the societal goals that one would like to pursue with traffic
management. For instance, an emphasis on high throughput requires the measurement of
intensities at the boundaries of the network, whereas an emphasis on travel time
reliability would require the measurement of travel times over a number of important
routes in the network. The definition of network performance will have to reflect the
policy makers’ priorities. Practice shows that such priorities tend to have a shorter life
cycle than the systems deployed for monitoring network performance which means that
flexibility is an important requirement for the systems involved.

2.1.2 Types of networks

There are different types of road networks with quite different characteristics. Types most
often distinguished are motorways, rural or provincial roads and urban roads. One of
the consequences of network-level monitoring is that the different types of networks have
to be considered as one network. For instance, door-to-door travel times often involve
more than one type of network. But the differences in characteristics between the
different types of network make this into a major source of complexity.
2.1.3 Basic and derived data

Network-related quantities are usually not directly measurable, but have to be derived from more basic, often point-based, quantities. It is often far from obvious which basic quantities are best to use and how the network-related quantities can be derived from them. It is essential to model the set of quantities, both basic and network-related, to get an overview of possible quantities, to obtain insight into which are dependent on which others, and to avoid dependency cycles.

2.2 Signal processing and data management problems

2.2.1 Sensor noise and failure

As a rule, individual traffic sensors feature a substantial noise level and regularly fail. For instance, in the Netherlands, induction loop data has an error level of 12% (van Lint, 2008), even if the sensor is reported to work properly.

2.2.2 Data heterogeneity

Most often it is too costly and too time consuming to install new sensors, specifically for the purpose of network monitoring, all over a network. One will have to make do mainly with the existing, legacy sensors that are usually part of existing local measures such as traffic signals or ramp metering systems. One of the consequences of this circumstance is that the data streams produced differ from each other in a number of ways, such as syntax, semantics, noise level and update frequency.

2.2.3 Changes in the network

A heavily used road network is constantly being maintained, changed and extended. This means that the network itself regularly changes, which has as a consequence that data sets gathered at different times are not easily comparable.

2.3 Multi-actor problems

When monitoring a network, the number of parties involved is essentially bigger than in the case of local monitoring. Apart from the traffic management authorities involved in the three different types of networks, one also has to deal with the larger number of people living in the area and the larger number of equipment manufacturers involved, to mention but the most important categories. The differences in interests exist even within the traffic management authorities, each of whom would like to get rid of congestion on his/her own network.

2.4 Software engineering problems

2.4.1 Legacy sensors

The legacy sensors mentioned above were developed and installed in various periods in the past. They were certainly not designed to be part of a homogeneous sensor network. As a consequence, their characteristics, in terms of accessibility, reliability and availability vary strongly. The same can be said of sensor density in a network.
2.4.2 Real-time requirements

Traffic management is about controlling a process which implies timeliness requirements or real-time requirements. It means that necessary data must be available in time and that decisions must be made when still relevant. The time available for processing data and for communication is therefore limited. Especially in a distributed system, real-time requirements can be hard to fulfil.

2.4.3 Long life cycle of traffic management and monitoring systems

Due to the large investments in installation of traffic monitoring systems, and the negative impact on traffic from the installation process, systems, once installed, tend to have a long life cycle. With all parties involved, there is a strong preference for changes in the system that only involves software and can be done remotely. It is an important software engineering problem to make systems such that roadside hardware changes are kept minimal.

2.5 Relations between the various problem aspects

The problems mentioned above are not independent. The most important dependency is as follows:

2.5.1 Software engineering implications of the traffic engineering problems

Regular traffic policy changes imply an uncertainty in which definition of network performance is most appropriate. This, together with the expected long life cycle, strengthens the flexibility requirement for monitoring systems. The derived data, the necessary noise reduction and the homogenisation of the many heterogeneous data sources imply a high computation and communication load on monitoring systems which, on the one hand, implies the necessity of a distributed architecture. But these two factors, on the other hand, make it harder to fulfil the real-time requirements. The higher number of parties implies that the design process, in which one would like to involve all of these parties, is essentially more complex than in the case of local monitoring.

3 Research objective and problem statement

The main research objective is to find out to what extent it is possible to design a monitoring system that can support network-level traffic management and can address the problems mentioned in the previous section. We want to find out what elements from the key disciplines involved, traffic engineering, signal processing and software engineering, are useful in such a design. A problem at this level of complexity cannot be solved at once. It is rather a matter of supporting the future research process in network-level traffic monitoring and control. The flexibility requirement in essence expresses that it should be able to support a research process, while also being used in real-life operation. This objective can be expressed in more detail as follows:
• How does one model the set of relevant traffic quantities such that interdependencies become visible, that the model can be adapted and extended in the future and that the model allows straightforward implementation in software?
• Which known filtering techniques should be applied to the modelled quantities and how can this be done without causing too much overhead in computation and in communication?
• How can data heterogeneity be tackled?
• How can historic data remain useful, even if the road network has undergone substantive changes?
• Which software architecture can serve the purpose of connecting a wide variety of legacy equipment, allowing the system to become strongly distributed, while meeting the real-time requirements, providing the system with high flexibility and making it likely that most future changes can be done in software and remotely, without disturbing traffic.

4 Approach

The approach to network monitoring that we applied in the Alkmaar belt control system has the following components:

4.1 Hierarchy of measuring agents

Each physical sensor and each virtual sensor (i.e., a missing physical sensor simulated by a traffic simulation model) (see below) is represented in software by an agent. Moreover, most derived quantities are related to composite features of the network such as routes (composed of a number of connected, nonintersecting road segments), origin-destination pairs or subnetworks. These network features can also be represented by agents which then serve as the sources of the related quantities. For instance, a (rough indication of) travel time over a given route can be calculated by the route agent by summing the travel times of the individual road segments that are part of the route. Likewise, the minimal travel time for an origin-destination pair can be determined by taking the minimum of the travel times over all routes between that origin and destination. This hierarchy offers an efficient and highly versatile way of measuring and calculating both basic and derived quantities of the state of a network.

4.2 Data pool, based on asynchronous middleware

Efficient, real-time data communication can be achieved by a so-called data pool (Doest et al., 2003), in which all data sources and data consumers are being represented and can be connected to each other most easily by dynamic, run-time configuration. The data pool is based on asynchronous publish/subscribe middleware (Eugster et al., 2003; Fiege et al., 2006; Lea et al., 2006) that offers real-time data exchange, both push and pull communication (i.e., the initiative may be with the producer as well as with the consumer), efficient network use and run-time updates of connections. In addition, this
kind of middleware has excellent properties for connecting to legacy systems, such as
legacy sensors, and in converting heterogeneous data streams towards streams with
uniform syntax and semantics.

4.3 Uniform data format and semantics
The required uniform syntax and semantics for basic and derived data is expressed in
a single data model, covering all the data items involved in network performance
measurement, together with rules to further extend the model with new data items.

4.4 Kalman filtering
Kalman filtering is a well-known mathematical technique to reduce noise in data streams
(Haykin, 2001). A Kalman filter combines a prediction based on previous measurements
with a new measurement. In general, both pieces of data are not certain but have
probability density distributions, usually assumed normal (Gaussian). The Kalman filter
can derive the most likely state estimate of the system involved from the latest
measurement and a prediction based on previous measurements. Moreover, it offers an
efficient, recursive scheme to do the necessary calculations, necessitating the storage of
only a minimal amount of data about the past. In the case of a distributed system,
monitored by a network of sensors, there is a multitude of schemes in which Kalman
filtering can be applied. It can be applied to single sensors, to groups of communicating
sensors, to the system as a whole and even to hierarchical compositions of Kalman filters
at different levels (see below). Factors to take into account in such configurations are,
among other elements, the amount of calculations, the predictability of the subsystems
considered, the communication overhead, and the complexity of the filtering
configuration. Schemes may involve cascading (where a subsystem’s state can serve as
input to other subsystems) or consensus strategies, by which subsystems decide together
about each others’ most likely state.

4.5 Traffic simulation
In network monitoring, traffic simulation can serve various purposes. As mentioned
before, it can replace missing physical sensors. A second purpose is that it can supply the
predictions needed for Kalman filtering at the various levels when the subnetwork
considered is too complex to allow straightforward mathematical modelling. A third
purpose applies to the frequent changes in the network. In order to be able to compare
current performance with data in the past, when the network was more or less different
from its current configuration, traffic simulation can be applied to generate the missing
performance data for the bygone network. In addition, the simulation model offers a
homogeneous data model that can serve the purpose of homogenising the heterogeneous
data streams from the various legacy sensors.
5 Related work

There is ample research on different configurations of Kalman filtering in distributed systems. Rao and Durrant-Whyte (1991) described an approach to fully distributed Kalman filtering in a sensor network. The authors compared this with an overall system estimate by centralised computation. They managed to avoid the centralised computation which is not attractive for its communication overhead, it being a single point of failure and a computational bottleneck. The authors still managed to obtain useful results.

Lendek et al. (2007) offered a comparison of centralised Kalman filtering with filtering in decentralised, cascaded systems. A cascaded distributed system is such that the behaviour of one subsystem is considered as input for other subsystems. This is obviously highly relevant for traffic monitoring applications, where the output of one crossing directly influences the behaviour of adjacent crossings. The decentralised approach is theoretically suboptimal but practically comparable, while having advantages like those mentioned in Rao and Durrant-Whyte (1991). Carli et al. (2007) combined local Kalman filtering with a so-called consensus strategy where individual agents join their estimates via a consensus process. In Wang and Papageorgiou (2005), the application of extended Kalman filtering (a generalisation from linear to nonlinear systems) is applied to traffic monitoring on a motorway segment. The paper demonstrates how the prediction part of Kalman filtering can be done by a macroscopic traffic simulation model and how a legacy monitoring system can be used to estimate essential parameters of the system.

The articles illustrated that the application of Kalman filtering and its combination with traffic simulation models can vary in the following aspects:

- The level at which Kalman filtering is applied varies from application at the smallest units in the system to application only at the top, for the whole system, or application at several levels of subsystems simultaneously.
- The way in which subsystems influence each other can be used in order to improve state estimates for individual subsystems. For instance, by considering subsystems as inputs to other subsystems, or by making use of subsystem state estimates in the prediction of the state of other subsystems. The order in which this can be done can also vary strongly and is probably quite important for the result.
- Apart from the quality of the state estimates, the optimal Kalman configuration in a distributed system is also determined by the resulting degree of system complexity, by system modularity, by the amount of communication needed, by the amount of computation and by the presence of computation bottlenecks.
- A number of parameters needed in a monitoring system have to be estimated after the system is installed.

The Alkmaar system has been developed such that all of these different configurations can be implemented with minimal changes in the software. This makes the system into a real-life test bed for advanced network monitoring. The flexibility stems for the largest part from the use of the agent hierarchy, implemented in a data pool, based on publish/subscribe middleware.
6 Current monitoring system in Alkmaar

The monitoring system currently installed in Alkmaar is part of a traffic control system for the belt road called Het Alkmaar Regelsysteem (HARS, the Alkmaar control system in Dutch (Vrancken et al., 2007)). Traffic management policy, defined by policy makers, is expressed in a so-called reference framework which gives minimal speeds for each road segment in the network and maximal travel times for each shortest route between boundary points of the network. The reference framework prescribes the boundaries of traffic states which are considered unproblematic. It also assigns a priority order to roads. Deviations from the reference framework indicate local (imminent) overloading of the network. If deviations cannot be avoided, then they should occur first on lowest priority segments. The reference framework determines the way the network is monitored and which quantities are being measured. Speed can be measured directly, travel times are a derived quantity. The responsibility for calculating travel times and comparing them with reference values has been assigned to so-called route-agents.

Figure 2 The HARS system (see online version for colours)

The control system has no sensors (nor actuators) of its own, but applies the roadside equipment of the existing traffic signals and Variable Message Signs (VMSs). The communication with the traffic signalling systems was among the most laborious tasks in the development of HARS because several manufacturers were involved. It has been
solved by the application of the IVERA protocol (the name is derived from the two
organizations, IVER and ASTRIN, involved in its definition),\(^1\) which offers a
standardised interface to all traffic signals. With the IVERA interface, all of the traffic
signals involved in HARS could then be connected to the HARS data pool for both
monitoring and control. Kalman filtering is applied only at the lowest level, the level
where sensors are connected to the HARS system. HARS also applies the traffic
simulation model MaDAM, part of OmniTRANS (Goudappel Coffeng, 2006). This
model is mainly used for control purposes and for replacing missing physical sensors. It
is not yet used as the predicting part of any Kalman filtering in the system.

The system is already in operation (see Figure 2), but frequent changes in the road
infrastructure made it difficult to obtain reliable performance data. Interruptions of the
system for maintenance purposes, however, have demonstrated that it already has a
noticeably positive effect on traffic.

7 Future research

Future research with the HARS test bed will concentrate on testing various monitoring
configurations, improving the traffic simulation model and applying it in more advanced
applications of Kalman filtering. HARS is also a test bed for traffic control. Experiments
with monitoring will be combined with experiments in traffic control, as traffic control is
the main purpose of the monitoring function of HARS.

8 Conclusions

Traffic control at the network level requires network monitoring. Effective ways of
network monitoring, which means estimating traffic states in a network, are still largely
unknown. Moreover, there is a vast number of possible ways to explore. Therefore,
flexibility is a highly important requirement in developing and installing network
monitoring systems. The development of the HARS traffic control system in Alkmaar,
which included a system for network monitoring, showed that the multi-agent approach
to monitoring, implemented by means of publish/subscribe middleware, is an effective
way to obtain the necessary flexibility.

Acknowledgements

The research leading to these results has received funding from the European
Community’s Seventh Framework Programme (FP7/2007-2013) under grant agreement
no. INFSO-ICT-223844, and from the Next Generation Infrastructures Foundation in
the Netherlands.
References

Note