Monitoring a freeway network in real time using single-loop detectors

Xiaoping Zhang*, Yinhai Wang
and Nancy L. Nihan

Department of Civil and Environmental Engineering,
University of Washington, Seattle, WA 98195-2700, USA
Fax: +206-543-5965 E-mail: xpzhang@u.washington.edu
E-mail: yinhai@u.washington.edu E-mail: nihan@u.washington.edu
*Corresponding author

Hu Dong

Department of Civil and Environmental Engineering,
University of California, 1 Shields Avenue,
2001 Engineering III, Davis, CA 95616, USA
E-mail: hdong@ucdavis.edu

Abstract: With the widespread adoption of automated traffic sensors, the increase of computer processing power, and the evolution and growth of the global Internet, our ability to monitor traffic and convey traffic information in real-time has been dramatically advanced. To the authors' knowledge, most of the current traffic information systems mainly provide general traffic flow information. Vehicle classification data, such as truck volumes that should be key inputs for good transportation planning, freight mobility analysis, roadway geometric and structural design and traffic control and operation, are often excluded or only roughly estimated. To address this problem, this paper describes the design and implementation of a traffic monitoring and information system recently developed at the University of Washington that visually conveys speed and vehicle classification information, obtained by processing real-time single-loop measurements, to the general public through the client/server computer architecture. The development of this system makes real-time monitoring of truck volume data on a freeway network possible.

Keywords: freeway; information; internet; traffic; truck data.

Reference to this paper should be made as follows: Zhang, X., Wang, Y., Nihan, N.L. and Dong, H. (2005) 'Monitoring a freeway network in real time using single-loop detectors', *Int. J. Vehicle Information and Communication Systems*, Vol. 1, Nos. 1/2, pp.119–130.

Biographical notes: Xiaoping Zhang is a PhD Candidate in the Department of Civil and Environmental Engineering at the University of Washington. She is also working towards her Master's degree in the Department of Electrical Engineering. She has a Master's degree and a Bachelor degree in transportation engineering. Her research interests include intelligent transportation systems and traffic operations.

Dr. Yinhai Wang is an Assistant Professor in transportation engineering at the University of Washington. He has a PhD in Transportation Engineering from the University of Tokyo, a Master's degree in computer science and Engineering from the University of Washington, and another Master's degree in construction management and a Bachelor degree in civil engineering from Tsinghua University. Dr. Wang serves as the Intelligent Transportation Systems (ITS) Program Coordinator at Transportation Northwest (TransNow), US Department of Transportation University Center for federal region 10. Over the past ten years, Dr. Wang has conducted extensive research in ITS data management and application, video image processing, freeway operations, traffic accident modelling and vehicle tracking.

Dr. Nancy L. Nihan is a Professor in the Department of Civil and Environmental Engineering at the University of Washington and the Director of Transportation Northwest (TransNow), the USDOT University Transportation Center for Federal Region 10. TransNow is a consortium of five universities in the four-state region, the University of Alaska at Fairbanks, Oregon State University, Portland State University, Washington State University, and the University of Washington, which is the lead university for the consortium. Dr. Nihan received her BSIE from the Department of Industrial Engineering and her MSCE and PhD from the Department of Civil Engineering at Northwestern University in Evanston, Illinois.

Hu Dong is a graduate student in the Department of Civil and Environmental Engineering at the University of California, Davis. He has a Bachelor's degree in transportation engineering and a Master's degree in computer science. His research interests include intelligent transportation systems and traffic operations.

1 Introduction

Monitoring traffic in real-time and timely dissemination of traffic information to the general public are critical factors that directly affect the performance of advanced traffic management systems (ATMS) and advanced traveller information systems (ATIS). In the past decade, with the widespread adoption of automated traffic sensors, the increase of computer processing power, and the evolution and growth of the global internet, our ability to monitor traffic and disseminate traffic information in real-time has been dramatically advanced.

Current traffic information systems mainly provide general traffic flow information to the public. For instance, as the most visited traveller information website in the state of Washington, the Puget Sound traffic flow map (http://www.wsdot.wa.gov/PugetSound Traffic/) displays near real-time traffic flow information to the public so that travelers can make better decisions on their mode choices, when to make the trips and trip routes.

In the past decade, with the increasing emphasis on just-in-time inventories and the growing impact of freight mobility on our regional economy, vehicle-classification data, especially accurate up-to-the-minute data on truck movements have become essential to our regional growth and market competitiveness. Also, because of heavy weights and large turning radii, the characteristics of truck movements are very different from those of passenger cars and must be considered in transportation planning and traffic analysis models. This difference makes continuous collection of truck volume and speed data along our region's roadways and the timely dissemination of these data imperative for a variety of purposes. For example: traffic operators need these data for real-time traffic monitoring, transportation researchers need these data to develop real-time incident detection algorithms, traffic engineers need these data as guidance on where to locate freight-oriented variable message signs, freight modellers need these data to provide calibration information for freight modelling, motor carrier despatchers need these data for routing information (Hallenbeck *et al*., 2003), etc. Therefore, a traffic information system that delivers and displays both real time speed and vehicle classification data, especially truck volume data, is very desirable.

A traditional way of collecting truck data is through traffic surveys conducted at a number of selected locations once every few years. The main shortcoming of this approach is the temporal and spatial discontinuity of the data. Thus, accurate truck-volume variation patterns over time on a freeway network cannot be obtained. The best and cheapest way to collect traffic data continuously is to use existing traffic sensors. So the first objective of this research was to identify a reliable traffic data collection method that has been widely deployed in the nation for continuous speed and vehicle-classification data collection. With such a readily available data collection method, the traffic information system developed in this research could be easily introduced to the majority of the nation's metropolitan areas. The second objective was to design and implement a traffic information system that would process the collected data and display both real-time speed and vehicle classification data, especially truck-volume data over time on a freeway network through the client/server architecture. The availability of this system would meet the need of transportation practitioners and researchers for real-time speed and vehicle-classification data.

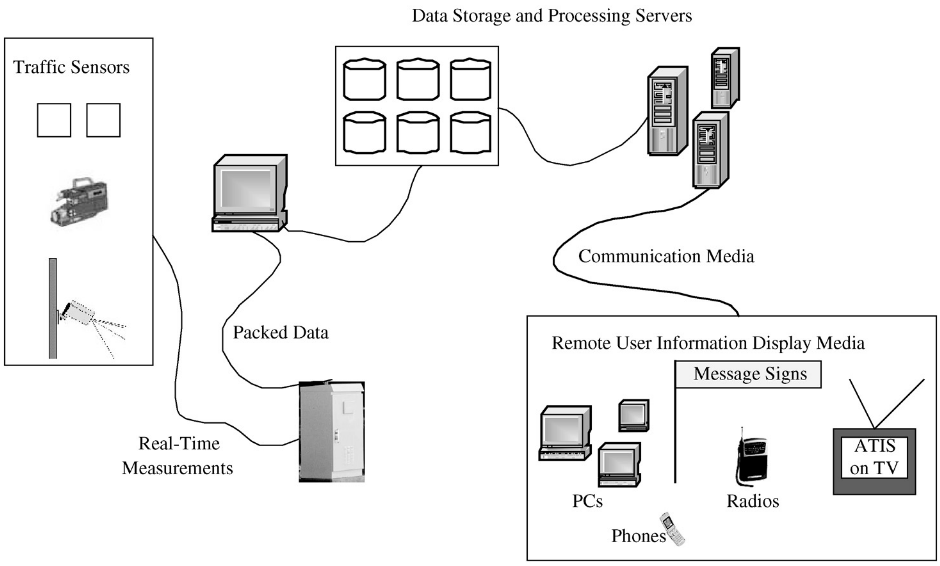
As shown in Figure 1, a typical traffic information dissemination system consists of at least four basic components: traffic sensors that collect traffic data in real-time such as inductive loop detectors, a data storage and processing server, communication media such as the internet and remote client information display media such as personal computers. The terms client and server refer to the two applications involved in a communication. The application that actively initiates contact, such as a personal computer, is called a client, while the application that passively waits for contact, such as servers that process data, respond to clients' requests, and broadcast traffic information to communication media, is called a server (Comer, 1999).

2 Real-time data source

Nowadays, with the widespread adoption of the global internet, agencies and organisations responsible for collecting and maintaining traffic data are able to provide transportation engineers and researchers direct access to file servers, and sometimes even data sensors. In many cases these agencies and organisations are responsible for the accuracy of the data, so users of the data only need to focus on model development rather than data collection and management (Wallace, Zhang, and Wright, 2001).

Traffic data can be collected by a variety of data sensors, such as inductive loop detectors, videos, remote traffic microware sensors (RTMS), etc. RTMS is a relatively new technology for collecting traffic data. But since it is still in the testing stage, only a limited number of RTMS have been installed. Therefore, it cannot be used for wide-area data collection. Video imaging systems have interested transportation researchers in the past decade, but due to their limitations under unfavourable weather conditions, such as bad lighting conditions, they cannot collect good traffic data continuously. Even under favourable lighting conditions, the video imaging technology has not advanced enough to accurately collect vehicle-classification and speed data, especially extracting information from processing images collected by un-calibrated CCTV cameras. In the state of Washington, an inductive loop detection system is still the main traffic data source. As one of the most popular automated traffic data collection methods, inductive loop detector technology was first introduced for detection of vehicles in the early 1960s (Traffic Detector Handbook, 1990), and today, after a 40-year evolution, it has become a ubiquitous means for collecting traffic data in the USA. So, in the state of Washington and a lot of other states, an inductive loop detection system is still the main traffic data source. Inductive loop detectors are frequently deployed as single detectors, i.e., one loop per lane, or as speed traps (also called dual-loop detectors or dual loops) formed by two consecutive single-loop detectors placed several metres apart. Single-loop detectors are used to measure volume and lane occupancy, while dual-loop detectors measure speed and vehicle length. In the state of Washington, field microprocessors aggregate inductive loop actuation signals into 20-second averages of flow, occupancy, velocity and vehicle length data.

Figure 1 General framework of a traffic information system



Real-time speed and vehicle classification data are important inputs for modern freeway traffic control and management systems. However, these data are not directly measurable by single-loop detectors. Although dual-loop detectors can provide speeds and classified vehicle volumes, there are too few of them on our current freeway systems to meet the practical needs of ATMS and ATIS, and the cost of upgrading from a single-loop detector to a dual-loop detector is high. In addition, a study (Zhang, Wang, and

Nihan, 2003) conducted by University of Washington researchers found that the dual-loop detection system in the state of Washington is not a very reliable source for vehicle classification data due to its undercount and misclassification problems. Using video ground truth data for validation of the loop measurements, the study also found that most of the single loops were providing good volume and occupancy measurements. Also because single-loop detectors are far more widely deployed than dual-loop detectors in the nation, they were considered as the best real-time data source for this research. So, the traffic data needed were collected by single-loop detectors on the Seattle Area freeway system and managed by the Washington State Department of Transportation (WSDOT) Traffic Systems Management Center (TSMC). Therefore, an algorithm that could process single-loop detector measurements, such as volume and occupancy, to produce speed and vehicle classification data was also needed.

3 System engine

An algorithm that can process single-loop measurements to produce speed and vehicle classification data has interested transportation researchers since single-loop detectors were first introduced (Athol, 1965; Hall and Persaud, 1989; Pushkar, Hall and Acha-Daza, 1994; Ishimaru and Hallenbeck, 1999; Dailey, 1999; Wang and Nihan, 2002; Wang and Nihan, 2003). One school of study involved applying Athol's speed estimation algorithm, shown in Equation (1), the focus of which was how to choose an appropriate value for the speed estimation parameter, g, for accurate speed estimation:

$$\overline{s}(i) = \frac{N(i)}{T \cdot O(i) \cdot g} \tag{1}$$

where i is time interval index, s is space-mean speed for each interval, N is volume, e.g., number of vehicles per interval, T is time length per interval, O is lane occupancy, e.g. percentage of time a single-loop detector is occupied by vehicles per interval, and g is speed estimation parameter.

In practice, g is assumed to be a constant value determined by the average effective vehicle lengths (EVLs) of the traffic stream. In reality, however, g varies with changes in vehicle composition; therefore, g can only be a constant when the mean vehicle length for each detection interval is approximately equal over the study time period (Wang and Nihan, 2002). So the key to accurate speed estimation while still keeping g constant is to make the input data meet the uniform vehicle-length assumption and the way of meeting this assumption is to use only intervals that contain small vehicles for speed estimation (Wang and Nihan, 2002).

Based on these findings, Wang and Nihan (2002) developed a single-loop algorithm that makes single-loop detectors capable of doing the work of dual-loop detectors. There are three steps in the algorithm: the first is to separate intervals with long vehicles from those without, the second step is to use measurements of intervals without long vehicles for speed estimation, and the third step is to identify long vehicle volumes for the intervals with long vehicles using the estimated speed. Preliminary tests for both spatial transferability and temporal transferability of the algorithm were performed and the results were encouraging. This single-loop algorithm was then chosen as the engine of the system to process single-loop measurements. Therefore, single-loop detector data collected by the WSDOT TSMC were employed as the data source for the proposed system to make the system transferable to other regions or states that only have single-loop detectors installed on their freeways.

As mentioned previously, the WSDOT single-loop detection system aggregates inductive loop actuation signals into 20-second averages of volume and occupancy. So, we used the smallest aggregation interval possible as our input to the developed single-loop algorithm. The reasoning was as follows:

- When we develop a real-time information system, we want to deliver information in a way as timely as possible. The WSDOT loop detection system broadcasts data every 20 seconds, so a 20-second delay is the best we can accomplish in the delivery of traffic information. In another words, what we are delivering is what has happened on the freeway network during the previous 20 seconds.
- We want to analyse the influence of truck presence on traffic flow as accurately as possible. By using 20-second interval data, we can compare how the presence of a large truck influences the speed during that 20-second period. If we aggregate data into longer periods, the speed variation caused by the presence of a large truck during a 20-second interval will be averaged off by other intervals, so it will be very hard to tell how the large truck influenced the traffic flow.
- We can easily observe truck arrival patterns from data with short-aggregation intervals. Whereas with long-aggregation intervals, it is very hard to analyse the truck arrival patterns. For example, suppose we know that 10 trucks passed by in the previous 15 minutes. With this information there is no way we can tell whether these trucks passed over that freeway segment as a group or each truck arrived randomly.

4 System design

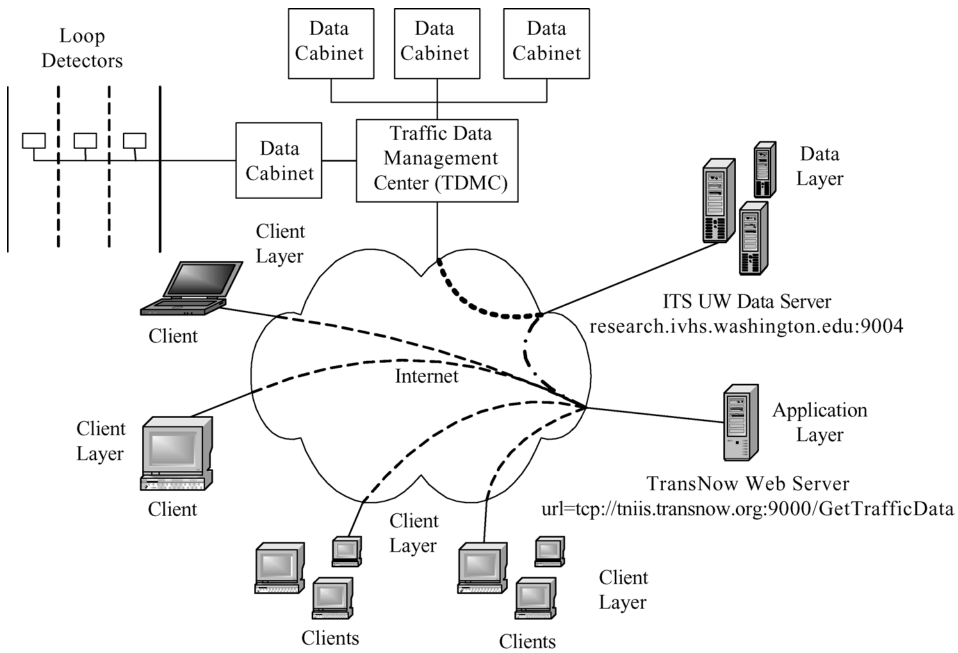
The schematic drawing presented in Figure 2 illustrates the structure of the proposed real-time speed and vehicle classification system and its basic elements. Consistent with the general framework depicted in Figure 1, this application-specific system consists of four basic functional components. Single-loop detectors are real-time traffic sensors, file servers managed by University of Washington (UW) intelligent transportation system (ITS) research group and web servers managed by Transportation Northwest (TransNow) are data storage and processing servers, the global internet is the communication medium, and user computers are the client information display medium.

Traffic data collected by inductive single-loop detectors are sent to the WSDOT TDMC for archiving. The TDMC grants direct access to its file servers to the UW ITS research group, which in turn provides other interested parties access to their data storage servers via a transmission control protocol (TCP) port 9004 (Loop Client, 1997), from where we retrieve the input data for the proposed real-time speed and vehicle classification system. The retrieved data are processed on TransNow web servers to obtain speed and vehicle classification data, which are then broadcast to the global internet. Real-time traffic information will be displayed on clients' computer whenever the clients start the client-side application program.

5 System implementation

Since the system retrieves traffic data from file servers instead of accessing data sensors directly, the system structure illustrated in Figure 2 can be simplified into three layers based on their functionalities. As shown in Figure 3, the three layers include the data layer, the application layer and the presentation layer. These three layers are connected by TCP channels through the internet. The data layer is the layer where the data are stored and accessed, the application layer is the layer where the data are processed and client requests are handled, and the presentation layer is the layer where traffic information is displayed. These three layers are described in detail in the following subsections.

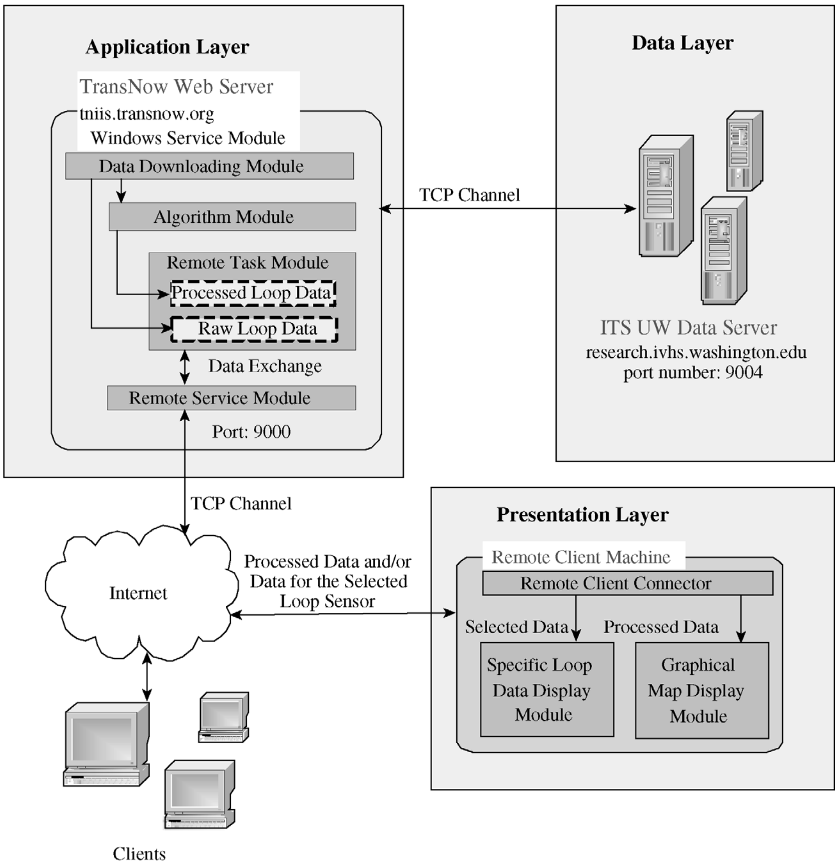
Figure 2 Structure of the real-time speed and vehicle classification system



5.1 Data layer

The data layer, maintained by the ITS Research Group in the Department of Electrical Engineering at the University of Washington, is running on the servers located at research.its.washington.edu. The UW ITS Research Group provides a software package called 'Loop Client' for requesting real time loop data from these servers on a UNIX platform. In our current study, we modified this software so that we could access the data on a Windows platform. The WSDOT loop detection system aggregates loop actuation signals into 20-second averages of volume and occupancy and sends the data back to TSMC for archiving and distribution. The TSMC data servers broadcasts loop data every 20 seconds, so the most recent data we can retrieve from the data layer are the measurements collected during the previous 20 seconds.

Figure 3 Illustration of functional layers



5.2 Application layer

The application layer, also called server-side software, is a special-purpose program designed to provide a single service, i.e., process downloaded data and handle clients' requests. This process is initiated when the computer that hosts the servers boots, and then runs continuously (Comer, 1999). The application layer can be divided into five major functional components as follows: the Windows service module, the data downloading module, the algorithm module, the remote task module, and the remote service module.

The Windows service module runs in its own Windows session as a Windows service. Whenever it starts, it registers a TCP channel to the data layer to retrieve real-time 20-second single-loop data. Meanwhile, it initiates the remote service module and registers a TCP channel so that the remote service module is able to listen for the client requests. This module is also responsible for passing downloaded data to both the algorithm and remote task modules.

The data downloading module connects to the UW ITS group's data servers to acquire real-time single-loop data over the internet. It was developed by modifying a software package called 'Loop Client'. As mentioned previously, 'Loop Client' was developed by the UW ITS Research Group to access the group's data servers on a UNIX platform. The data downloading module for the current study is a modified version of 'Loop Client' which accesses the data servers on a Windows platform.

The algorithm module processes the downloaded data to calculate average speed, and counts of short vehicles and long vehicles. These processed data are then sent to the remote task module for temporary storage. The algorithm employed in this module is the single-loop algorithm described previously in the system engine section. In this research 15 consecutive 20-second single-loop data intervals, including the most recent 20-second interval, are used to estimate vehicle speed and to classify vehicles. In the first step, these 15 consecutive 20-second intervals are sorted by the average occupancy-per-vehicle value in an ascending order and then classified into two groups based on an empirical occupancy-per-vehicle threshold value, with Group A containing those 20-second intervals that only include short vehicles and Group B containing those 20-second intervals that have at least one large vehicle traversing the detector during the interval. In the second step, intervals contained in Group A are used to calculate speed by applying Athol's speed estimation algorithm. In the third step, the calculated speed is used to estimate how many small vehicles and large vehicles are contained in each of the 20-second intervals in Group B by applying a nearest neighbour method (Wang and Nihan, 2003). The calculated speed and numbers of small and large vehicles will then be sent to remote task module for temporary storage. The remote task module temporarily stores the processed data and raw loop data and responds to clients' requests. The remote task module receives client's requests forwarded by the remote service module, interprets the details of the clients' requests, and returns the resulting information to the client program through the remote service module.

The remote service module is responsible for registering a TCP channel (Microsoft Net Remoting, 2003) so that the client-side machines can connect to the TransNow web server via this TCP protocol. the remote service module is also connected to the remote task module, to where it forwards all the requests of the clients'. This connection provides an approach for client machines to access the remote task module for data exchange.

5.3 Presentation layer

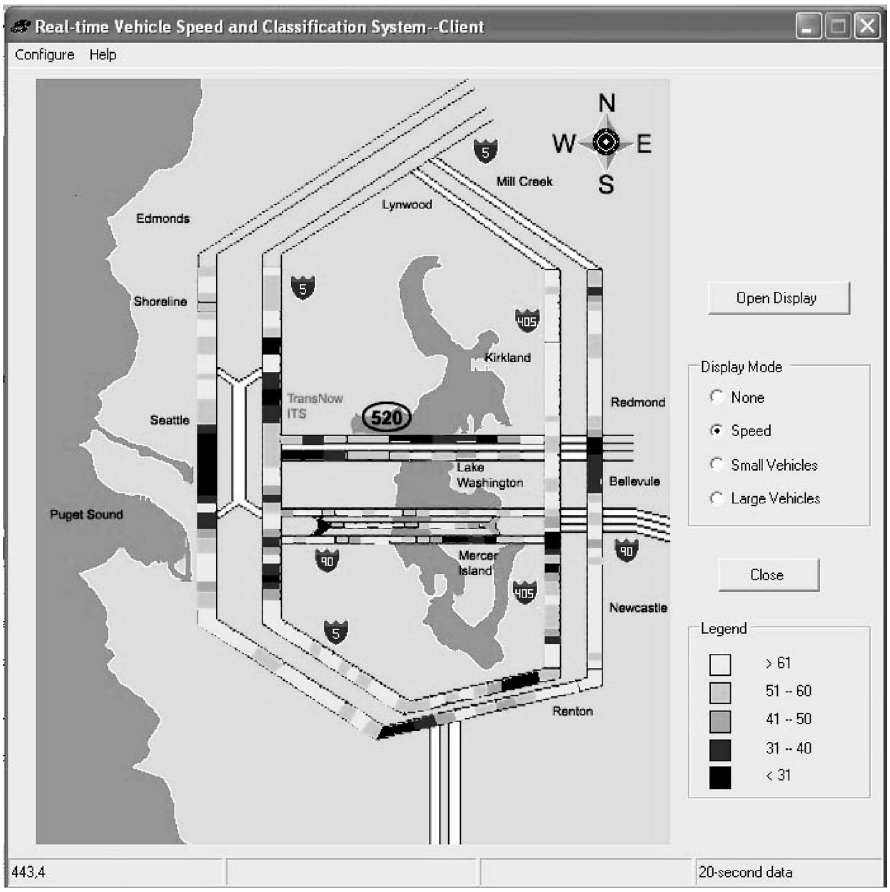
The presentation layer is the client program. A client program is usually a general-purpose application program that becomes a client temporarily when acquisition of remote data is necessary. It also performs local operations (Comer, 1999). This program runs on any client's Windows computer with internet access and is active for a single user session. In this system, during the execution of the client application program on the user's local computer, the program launches onto the internet a request for real-time vehicle speed and classification data using a TCP protocol. The program then retrieves the data and visually presents it on a digital map. Based on their functionalities, the presentation layer has three functional components including the remote client connector, the map display module, and the specific loop display module.

The remote client connector retrieves the traffic data, such as speed and vehicle counts for each loop station and for each selected individual loop. The map display module presents the preprocessed data on a colour-coded digital map. The Specific Loop Data Display Module displays traffic information for any selected loop.

5.3.1 User interface design

The user interface, shown in Figure 4, was designed to visually present traffic information to clients and to facilitate clients' queries. The traffic information shown on this map gets updated every 20 seconds. The freeway network is colour-coded so that different ranges of speed or vehicle counts are displayed in different colours. There are four display modes – none, speed, small vehicles, and large vehicles – for clients to select. The background map was digitised so that users can click on any point on the freeway network to query traffic information for any individual loop at any loop station.

Figure 4 Real-time vehicle speed and classification system client interface (speed)

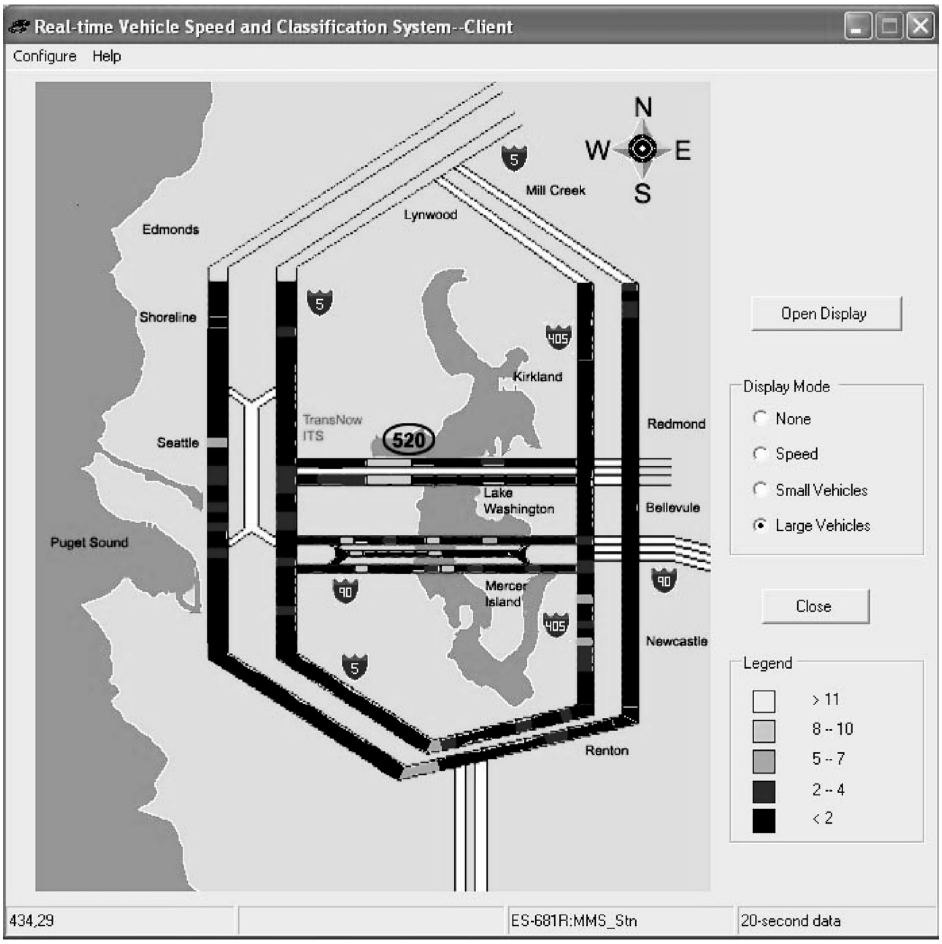


6 System application

This system can be applied to monitor traffic and disseminate traffic information in real time. As mentioned previously, there are four display modes-none, speed, small vehicles, and large vehicles – for clients to select. If the mode 'Speed' is chosen, the average speed calculated based on 20-second single-loop measurements for each loop station is displayed on the freeway network. For example, Figure 4 shows a snapshot of speed information on the freeway network during an afternoon peak-hour period on a typical weekday. As can be seen in this figure, some segments of the freeway network are experiencing slow-moving traffic, especially the westbound traffic on the SR 520 floating bridge (also called evergreen point floating bridge), where the average speed is less than 30 miles per hour. This, to some extent, reflects the fact that the SR 520 floating bridge is now carrying twice as many vehicles as its original design capacity. If the mode 'Small Vehicles' is chosen, the number of short vehicles (length < 12.2 m, 40 ft) (Wang and Nihan, 2002) detected by each loop station is shown on the freeway network.

If the mode 'Large Vehicles' is chosen, the number of large vehicles (length > 12.2 m, 40 ft) (Wang and Nihan, 2002) detected by each loop station in the previous 20 seconds is shown on the freeway network. For example, as can be seen in Figure 5, six segments of the freeway network had over five large-size vehicles pass by in the previous 20 seconds. By using this system, traffic operators and researchers can visually identify the locations where traffic is moving slowly and the locations where truck volumes are high.

Figure 5 Real-time vehicle speed and classification system client interface (number of large vehicles)



7 Summary and future directions

This system visually conveys speed and vehicle-classification information, obtained by processing real time single-loop measurements, to the general public through the client/server computer architecture. The development of this system makes real-time monitoring of truck volume data on a freeway network possible. It also provides an example of the integration of advanced transportation technologies to meet the growing needs for monitoring traffic in real time and the timely dissemination of traffic information.

The author is currently upgrading the system so that more traffic information such as historic speed and truck volume variations can be displayed in response to clients' queries based on the data collected at each loop station or even at each single-loop detector on the Seattle freeway network. The graphical interface of the system will also be refined to make it more user friendly. Once completed, the system will be released to interested parties and clients to meet their needs for real time and historic speed, and vehicle classification information.

References

Athol, P. (1965) 'Interdependence of certain operational characteristics within a moving traffic stream', *Highway Research Record*, Vol. 72, pp.58–87.

Comer, D.E. (1999) *Computer Networks and Internets*, ISBN 0130836176, Prentice-Hall, Upper Saddle River, NJ.

Dailey, D.J. (1999) 'A statistical algorithm for estimating speed from single loop volume and occupancy measurements', *Transportation Research, Part B*, Vol. 33, No. 5, pp.313–322.

Hall, F.L. and Persaud, B.N. (1989) 'Evaluation of speed estimates made with single-detector data from freeway traffic management systems', *Transportation Research Record 1232*, TRB, National Research Council, Washington, DC, pp.9–16.

Hallenbeck, M.E., McCormack, E., Nee, J. and Wright, D. (2003) 'Freight data from intelligent transportation system devices', *Research Report WA-RD 566.1*, Washington State Department of Transportation.

Ishimaru, J.M. and Hallenbeck, M.E. (1999) 'Flow evaluation design technical report', *Technical Report WA-RD 466.2*, Washington State Department of Transportation.

Loop Client (1997) http://www.its.washington.edu/software/loop_cli.html, accessed May 26, 2003.

Microsoft Net Remoting (2003) 'A technical overview', http://msdn. Microsoft .com /library/en-us/ dndotnet/html/hawkremoting.asp?frame=true accessed May 26, pp.313–322.

Pushkar, A., Hall, F.L. and Acha-Daza, J.A. (1994) 'Estimation of speeds from single-loop freeway flow and occupancy data using cusp catastrophe theory model', *Transportation Research Record 1457*, TRB, National Research Council, Washington, DC, pp.149–157.

Traffic Detector Handbook (1990) *Institute of Transportation Engineers*, Second Edition, http://www.ite.org ISBN 0-935403-14-0.

Wallace, R.M., Zhang, Y. and Wright, J.R. (2001) 'Distributed system for coastal infrastructure modeling and assessment', *Journal of Computing in Civil Engineering*, Vol. 15, Issue 1, pp.67–73.

Wang, Y. and Nihan, N.L. (2002) 'A robust method of filtering single-loop data for improved speed estimation', *CD-Rom for the 81st Annual Meeting of TRB*, paper 02-3843, TRB, National Research Council, Washington, DC.

Wang, Y. and Nihan, N.L. (2003) 'Can single-loop detectors do the work of dual-loop detectors?', *ASCE Journal of Transportation Engineering*, Vol. 129, No. 2, pp.169–176.

Zhang, X., Wang, Y. and Nihan, N.L. (2003) 'Investigating dual-loop errors using video ground-truth data', *ITS America's 13rd Annual Conference (CD-ROM)*, May, Minneapolis, Minnesota.