University of Pennsylvania
ScholarlyCommons

Real-Time and Embedded Systems Lab (mLAB)   School of Engineering and Applied Science

11-2008

Real-Time Traffic Congestion Prediction

Rahul Mangharam
University of Pennsylvania, rahulm@seas.upenn.edu

Insup Lee
University of Pennsylvania, lee@cis.upenn.edu

Oleg Sokolsky
University of Pennsylvania, sokolsky@cis.upenn.edu

Follow this and additional works at: http://repository.upenn.edu/mlab_papers

Real-Time Traffic Congestion Prediction

Real-Time Traffic Congestion Prediction

Rahul Mangharam*†   Insup Lee†   Oleg Sokolsky†

*Dept. of Electrical & Systems Engineering   †Dept. Computer & Information Science
University of Pennsylvania   University of Pennsylvania
{rahulm}@seas.upenn.edu   {lee, sokolsky}@cis.upenn.edu

I. Transportation-CPS Networks

According to the 2007 Urban Mobility Report, delays due to heavy traffic are now costing Americans $78 billion in the form of 4.2 billion lost hours and 2.9 billion gallons of wasted fuel. As seen in Fig. 1, vehicular traffic congestion is a growing problem with more drivers experiencing severe and extreme delay where the travel time is in excess of 1.5 times the free-flow trip time. In Fig. 2, we observe that the dynamic range of regular traffic delays is large and furthermore, the spread of the delay during congested hours is significant. In addition, 2/3 of traffic delays are caused not by recurring congestion but by point-based spontaneous congestion due to traffic incidences. This is a core issue in end-to-end travel time prediction as the network is dynamic and prediction estimates must be made within a short duration or they will be out-of-phase with the current network state.

Vehicle traffic probing and travel-time prediction is fundamentally a Cyber-Physical Systems problem of massive scale as computations must consider real-time trajectory data from ≥ 1 million vehicles in a region along with spatio-temporal environmental variables such as weather, road conditions, sports events, school and factory hours, etc. to accurately estimate a canonical set of fastest paths for all drivers by a deadline. Effective congestion probing and prediction algorithms must consider both recurring traffic patterns and non-recurring spontaneous traffic events. The traffic infrastructure must support computation over a large amount of historic traffic data on one end and acquire timely updates from a large number of vehicles on the other end. Given the dynamical nature and scale of the problem, neither classical centralized nor distributed algorithms will perform well. Furthermore, unlike classical Real-Time Systems, where all tasks are assumed to be compliant, in the case of traffic prediction, the drivers-in-the-loop exhibit a varying degrees of non-compliance.

A new class of data-dependent real-time algorithms and system architectures are necessary to compute across large statistical variations and must continually deliver outcomes with the desired level of QoS within a deadline. If for example, street segments are treated as shared resources in time and space, then scheduling of fastest vehicle paths along all origin-destination pairs requires that drivers conform to a weighted estimate of both centrally and locally computed paths.

To summarize, vehicular traffic congestion is a significant and growing problem that affects the entire nation's population. Traffic congestion is a dynamical problem where both sporadic traffic incidences affect travel time delay in a disproportionately large manner and spatio-temporal environmental factors affect traffic behavior in the region and for the duration of the event. Finally, traffic congestion prediction is a data-dependent CPS-problem of massive scale where traffic assignment decisions have to be made by a deadline based on terabytes of semi-global historic traffic data on one hand and hundreds of thousands of real-time vehicle trajectory updates on the other.

The principal properties required by Transportation-CPS networks are:

(i) Prediction algorithms which execute with real-time guarantees with feedback from large historic traffic data and a large number of trajectory-tracking devices.

(ii) As the problem space spans multiple spatial regions and time scales, there is a need for composable algorithms and system architectures with traffic sensors, actuators and controllers to estimate end-to-end paths and travel times.

(iii) Given the scale, statistical nature and non-compliance of drivers, there is a strong need for on-line
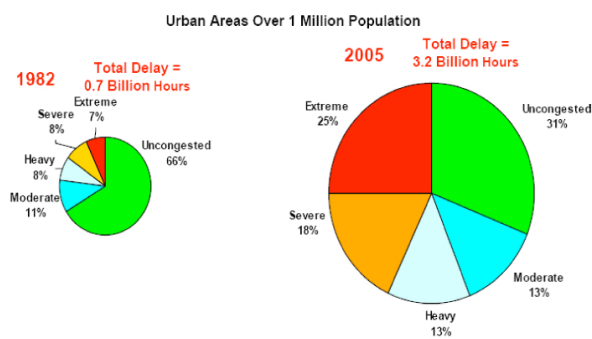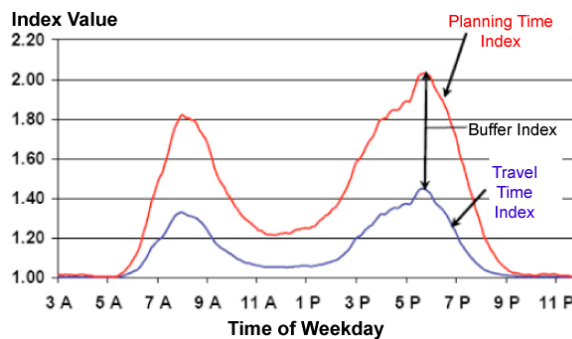

Fig. 1. Growth of Vehicle Traffic Congestion

Fig. 2. Large variation in travel time

validation of any probing and actuation within the road network.

(iv) As all traffic data is gathered by anonymous sources on the road, the security and privacy of users' trajectory information must be guarded in a distributed and scalable trust management framework.

To provide such guarantees, there is a need to express timeliness and node-independent coordination across the network. Scalable protocols for vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) need to be designed with timeliness and in-system on-line validation as first-class citizens. To build an end-to-end model-based framework for Transportation-CPS, we require both the abstractions and tools that combine the scale of metro-area simulation with real-world test-beds.

The three primary challenges with Transportation-CPS networks are:

1. How to specify, integrate, validate and verify dynamical and distributed systems.

2. How to identify unintended or unanticipated behavior a priori.

3. How to assure fault tolerance, and provide system reliability despite a less reliable physical infrastructure of vehicles.

II. Position

Our position is that:

A. With a nation-wide infrastructure for real-time V2V and V2I networking, vehicles will be enabled with both Active Networked Safety (on smaller time-scales) and efficient path planning capabilities (on larger time-scales) that make driving more efficient, safer and more enjoyable. Network-enabled vehicles will be able to both probe and predict near-term traffic congestion more accurately and distribute their travel routes more evenly. Current infrastructure-based traffic sensors do not scale well and do not capture vehicle trip-specific speed information. Furthermore, the long round trip duration of infrastructure-based traffic data acquisition, processing and reporting via cellular telephony, radio or satellite results in too little information delivered too late.

B. Vehicle network virtualization, which combines large-scale simulation of virtual vehicles with real vehicles and executes the same algorithms and protocols across both vehicles types, allows for scalable and composable model-based design of Transportation-CPS networks. By modeling the mobility, communications, traffic dynamics, driver behavior and vehicle control of thousands of virtual vehicles we study the scalability and robustness of our system. By designing the same network abstractions and protocols to work in real vehicles, we evaluate the performance of the system with real-world feedback and validate the models themselves. Now by combining the virtual and real vehicles to seamlessly communicate and interact with the same protocols and algorithms on the same street map, we develop richer models and are also able to rapidly prototype Transportation-CPS systems. The validated models may then be used in end-to-end model-based design for real-time traffic prediction algorithms and end-to-end protocols for traffic management.

Such an approach allows us to answer questions such as:

1. At what market penetration does the traffic prediction become effective?

2. What are the metrics to compute the "travel time competitive ratio"? (Street diversity, weather, school locations and hours, road conditions, factory locations and hours, sporting events, etc.)

3. What is the best case of travel time prediction for the top ten cities?

4. How do we identify the hopeless traffic scenarios?

5. How do route closures affect traffic patterns?

6. How can we best assign traffic routes to minimize the induced traffic congestion (and prevent the resultant oscillation in traffic routes)?

7. What is the impact of: **a)** congestion coverage on the accuracy of congestion prediction? **b)** historical traffic information? **c)** mobile gateways and fixed base stations on traffic prediction?

8. What adoption patterns do we expect to see with users?

9. What are the different QoS levels of traffic information?

III. Fundamental Limitations

The complexity of Transportation-CPS is beyond the state of current real-time theory and practice in reasoning about large-scale mobile network protocols with multiple dynamical inputs. The three main factors which contribute to this complexity are: (a) the data-dependent nature of the traffic dynamics and environment variables which require real-time stream processing algorithms to provide traffic updates within narrow time intervals, (b) the inherent non-compliance of system components (i.e. drivers) has previously not be considered in classical real-time systems and finally (c) the scale of the problem with over 100K input streams requiring real-time feedback and distributed trust management.

IV. Research Challenges

Challenge 1- *New Class of Real-Time Network Protocols:* Transportation-CPS networks introduce new constraints on network protocols that distinguish them from conventional point-to-point connection and connectionless protocols. The high rate of change in network topology and the fact that vehicles just entering or exiting a highway drastically alter the routing tables, make path-based and hierarchal address-based route discovery infeasible. If we rely on broadcast-based protocols, there is a need to introduce spatio-temporal structure for per-hop packet scheduling. In cases where traffic incidences result in disrupting traffic patterns for extended periods of time, the protocols must ensure the event alert is persistent in the region of interest. Finally, for bulk data exchange and stream-based communication, the protocols must support 'suspend-resume' operations due to frequent connectivity changes in the network. The major challenge is in developing a suite of high-confidence protocols even though the underlying physical network is unreliable.

Challenge 2- *Distributed and Adaptive Scheduling Theory:* Real-Time models and scheduling theory must be advanced in many ways. Given the spatial scale of the problem and the volume of data that must be churned in real-time, new types of compositional modeling are required. Not only must the problem be partitioned in time and space but consider the statistical nature of the inputs to determine the models and mechanisms to be used for on-line composition.

Challenge 3- *Scalable Dynamic Trust Management* The topological attributes of vehicle-to-vehicle networks permit only unplanned ephemeral relationships between network elements. While reputation-based trust management protocols have been proposed for mobile wireless networks, the scale and potential dangers of misinformation do not permit such 'soft' approaches to vehicular security. The large number of vehicles pose a problem with key assignment and the untethered nature of vehicles does not allow for traditional public key infrastructure.

V. Innovations and Abstractions

A bold new research agenda in real-time software and system design must be set to create innovations and abstractions to meet these challenges.

Data Intensive Network Organization Traffic prediction networks require a combination of *big data* and *large number of devices* to combine semi-global historic traffic information and on-line vehicle updates. This requires the convergence of approaches used by the datacenter community and real-time community. Such dynamical systems with real-time constraints require a new approach to system infrastructure design and machine-to-machine and machine-to-grid network protocols. As the underlying physical network is unreliable, network protocols must not associate routes with nodes but with primitives such time, position, driving direction and average speed. For example, a spatio-temporal grid structure may be overlaid on the street map to determine communication activity such that messages are delay-bounded along highways. By tightly synchronizing vehicles, we can derive reliable logical abstractions of the network through wireless interference control and end-to-end spatio-temporal schedules across a range of vehicle densities and street topologies.

Scalable Real-Time Prediction Algorithms Real-Time prediction algorithms are required in a wide array of CPS applications from financial markets and supply chain management to robot path planning and transportation. As stream processing is applied to larger scale problems, the need for real-time guarantees to data-intensive computation has recently surfaced. New approaches to real-time parallel computation and communication are necessary. For example, we used the Nvidia CUDA computation platform to build the AutoMatrix traffic congestion simulator (see Fig. 3). AutoMatrix is capable of simulating traffic on any street map in the US with first-order mobility, communication and traffic
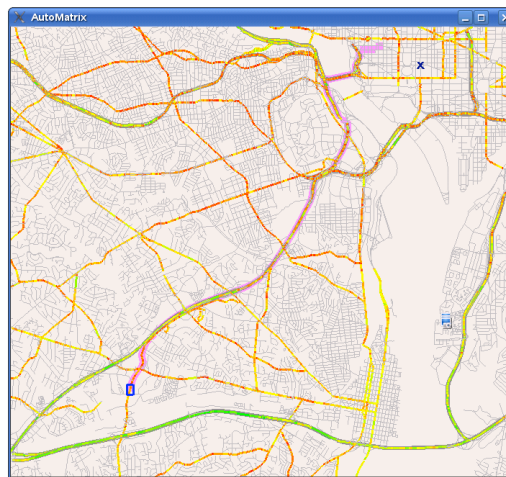
Fig. 3. AutoMatrix traffic simulator with 800K vehicles

queuing models for analysis of recurring and non-recurring congestion. By processing vehicle trajectories in parallel on an Nvidia graphics processing unit we are able to compute the fastest path origin-destinations of over 5 million vehicles. While this operation does not have real-time capabilities, the algorithms required have opened a new class of real-time parallel computation.

Distributed Trust Management Architectures A framework for decentralization of security policies and policy enforcement is required to facilitate distributed authorization management and re-delegation of authorization in vehicular wireless networks. Key establishment mechanisms based on time, space and neighboring vehicles are necessary to ensure the security and privacy of drivers is guaranteed. Finally, proof-based and theorem-proving data validation algorithms must be used to ensure the integrity of traffic updates and to prevent denial of service attacks.

VI. Conclusion

The automobile of the future will be a *Programmable Car* where functionality may be purchased and enabled on-the-fly via software services. The Programmable Car will be networked for network-based active safety and real-time traffic congestion probing and prediction. Given the large scale and statistical nature of vehicular traffic congestion, new research in Transportation-CPS network protocols and real-time stream processing algorithms are required.

VII. Biography

Rahul Mangharam is an Assistant Professor of Electrical & Systems Engineering at the University of Pennsylvania. His interests are in scheduling algorithms and operating systems for real-time embedded and wireless networks. Insup Lee is the Cecilia Fitler Moore Professor in Computer and Information Science at the University of Pennsylvania. His interests include real-time systems and languages, formal methods and model-based design. Oleg Sokolsky is a Research Associate Professor in Computer and Information Science at the University of Pennsylvania. His interests are in the application of formal methods to design and verification of distributed real-time systems.