“MIDTOWN IN MOTION”: A NEW ACTIVE TRAFFIC MANAGEMENT METHODOLOGY AND ITS IMPLEMENTATION IN NEW YORK CITY

Wuping Xin
Chief Scientist
KLD Engineering, P.C.
43 Corporate Drive
Hauppauge, NY
U.S.A.
Telephone: (631) 617-5650
E-Mail: wxin@kldcompanies.com

Jinil Chang, PhD
Senior Traffic Engineer
KLD Engineering, P.C.
43 Corporate Drive
Hauppauge, NY
U.S.A.
Telephone: (631) 617-5650
E-Mail: jchang@kldcompanies.com

Satya Muthuswamy, P.E., PTOE
Senior Traffic Engineer
KLD Engineering, P.C.
43 Corporate Drive
Hauppauge, NY
U.S.A.
Telephone: (631) 617-5650
E-Mail: satya@kldcompanies.com

Mohamad Talas, PhD, P.E., PTOE
Deputy Director of System Engineering
New York City Department of Transportation
Long Island City, NY
U.S.A.
Telephone: (718) 433-3390
E-Mail: mtales@dot.nyc.gov

Submitted for Presentation and Publication
Transportation Research Board 2013 Annual Meeting
Submitted on Nov 15, 2012

5500 words + 8 figures (2000 words) = 7,500 words (including 256 words in abstract)
ABSTRACT

This paper summarizes a new active traffic management methodology developed and deployed in New York City. The design, development and implementation of the methodology are driven by the availability of extensive real-time travel time data, the upgraded ITS infrastructure and the need of a reliable and robust real-time signal control system for congestion management in a complex urban traffic environment. The primary objective is to maintain mobility in the subject control zone by preserving capacity and avoiding breakdown within the zone. To that end, multi-regime adaptive signal control strategies, ranging from strategically regulating traffic demand to balancing queue-storage ratio at critical intersections, work in concert to proactively manage congestion. Large-scale real-time travel time data (nearly 1 million records of individual trips per day) are utilized for effective control. This renders the system perhaps the first large-scale active traffic management framework that uses travel-time based adaptive control for congestion management in an urban grid network. In addition to running autonomously, a new mode of operator-in-the-loop allows the operators to interact with important signal timing decisions when the traffic pattern changes abruptly or switch of control policies is involved. The system has been deployed in a congested urban central business district in Manhattan, New York City. The initial deployment, called “Midtown-in-Motion”, includes a core 110 square block zone of Midtown Manhattan, from 2nd Avenue to 6th Avenue, and from 42nd to 57th Streets. New expansion is currently underway, covering more than 270 square blocks of Midtown Manhattan, from 1st to 9th Avenue, and from 42th to 57th Street.
1. INTRODUCTION
Traffic congestion is a significant problem in urban areas. In 2010, the national congestion cost in the U.S. rose to $101 billion, and is projected to continue to grow to $133 billion in 2015 and $175 billion in 2020 (1). Active traffic management (ATM) is a holistic approach to congestion management that dynamically manages recurrent and non-recurrent congestion based on prevailing traffic conditions (2). In practice, the focus of ATM has been primarily on increasing capacity and smoothing traffic flow on highways, employing techniques such as variable speed limits, hard-shoulder running, managed lanes, or ramp control. In the environment of urban roadways, the application of ATM, especially using diversified real-time traffic signal control technologies for congestion management, has been quite limited.

In New York City, the legacy closed-loop traffic control system known as Vehicle Traffic Control System (VTCS) has recently been upgraded to the newly developed New York City Traffic Control System (NYCTCS) (3). Wired communications are replaced by a dedicated carrier-free mobile broadband called New York City Wireless Network (NYCWiN). Non-intrusive wireless sensors, roadside ETC tag readers, and a new type of advanced solid state controllers (ASTC) are deployed over a traffic network of more than 12000 signalized intersections. With this new ITS environment, an innovative active traffic management methodology has been developed and implemented to manage congestion proactively using multi-regime adaptive signal control. This methodology consists of a combination of real-time control strategies, ranging from strategically regulating traffic demand to balancing queues at critical intersections. Large-scale travel time data are collected in real-time and congestion measures are developed for effective control. The adaptive control strategies, based on the prevailing traffic conditions, are implemented in concert to systematically alleviate congestion.

2. BACKGROUND
Active traffic management (ATM) in this paper refers to the broader concept of dynamically manage recurrent and non-recurrent congestion based on prevailing traffic conditions (2). Though existing ATM applications focus on applying a combination of operational strategies in freeway or corridor congestion management, this paper extends the application domain to urban grid networks using diversified signal control strategies. These control strategies are implemented proactively in real-time to work in concert by time and by places, for the performance goal of alleviating congestion and improving mobility. Specifically, signal control is systematically employed under the active traffic management framework as a mean to regulate demand for better space utilization and to manage queue for improved mobility and travel time reliability. This is different from the conventional objective of adaptive signal control to facilitate progression or minimize intersection delays. The methodology is integrated in a centralized traffic management center environment, either running autonomously or providing the decision support for the control room operators in case of special events, incidents, inclement weathers or switch of tactical control strategies.

From a historical point of view, the concept of adaptive signal control in the U.S. can be traced back to the 1970s when the Urban Traffic Control System (UTCS) had begun to gain wider use for controlling interconnected intersections (4). The first generation UTCS systems are characterized by the open-loop, fixed-time control where timing plans are prepared offline and are selected based on a time-of-day (TOD) schedule (5). The second generation UTCS systems provide the closed-loop capability of online computation of offsets and splits, while keeping cycle length fixed within variable groups of intersections (6). In between, the 1.5 generation UTCS (commonly referred to as traffic responsive plan selection, i.e., TRPS) selects a timing plan based on the traffic pattern matched from a modest number of system sensors (7, 8).

The second generation UTCS systems are the first attempt to optimize offsets and splits in real-time without changing cycle lengths. In recent years, adaptive signal control systems have gain greater attention. These systems are typified by rapid adjustments of signal timings in response to traffic variations (9). Increasing number of adaptive systems has been deployed in the U.S. and other countries. These systems, however, are primarily for the purpose of reducing delays at signalized intersection and improving real-time progression. The application of adaptive signal control technology systematically as a congestion management strategy has been quite limited (10).
Gazis and Potts was probably the first researchers who addressed the signal control problem in congested conditions (11). Using a graphical approach that was later verified by the application of Pontryagin’s maximum principle (12), Gazis identified a two-stage bang-bang control policy, which alternates between maximum and minimum green times till queue dissolves simultaneously on all approaches. Michalopoulos and Stephnopolos placed a queue length constraint to Gazis model and pointed out if queues on all approaches reach the constraint, the problem of minimizing total delay has no feasible solutions (13, 14). Dans and Gazis proposed the so-called “store and forward” model for oversaturated traffic networks (15). It is based on the observation that vehicles are first “stored” in queues then “forwarded” to downstream intersections at saturation flow rates.

In addition to closed-form analytical solutions, signal control under congested conditions is also formulated as parametric optimization problems and solved using simulation tools or numerical approaches. Park et al. proposed a Genetic Algorithm (GA) optimization program using simulation model to generate performance measures to be optimized (16). Lo et al. integrated Cell Transmission Model and developed a Dynamic Intersection Signal Control Optimization (DISCO) method (17, 18). Abu-Lebdeh and Benekohal proposed the formulation and the GA-based solution algorithm of a dynamic signal control for queue management (19-21). Liu and Chang employed detailed queue modeling taking in account shared-lane interactions and the dynamics of queue spillbacks for obtaining optimal signal timing in oversaturated conditions (22). Lieberman et al. presented the problem as a mixed-integer program to yield optimal offsets and the nominal queue length for each approach, and signal timings are adjusted so that the actual queue length closely approximates the nominal one in real-time (23).

SCOOT implements congestion offsets and systematic strategies such as gating (24-28). Congestion Importance Factor (CIF) is specified for each approach representing different level of priorities. The cycle length optimizer in SCOOT is not affected by congestion, as it may be the result of either insufficient or excessive cycle time (29). Notably, the store-and-forward concept has been implemented in the TUC system (Traffic Responsive Urban Control) with the control objective set to balance the ratio of vehicles storage and the link length (30-32).

3. METHODOLOGY OVERVIEW

Midtown Manhattan is among the largest CBDs in the United States with more than 700,000 commuters working in this area. It is characterized by a closely spaced signalized Cartesian grid systems of one-way streets, with short blocks (approximately 200 ft.) along the avenues, and long blocks (approximately 500 ft. or longer) along the cross streets. Substantial traffic demand exists due to the extremely heavy concentration and diversity of activities in this area. Vehicular traffic is often mixed with heavy pedestrian volumes, while intersection spillback arising from overflow queues is frequent, especially in the vicinity of business activity centers. Additionally, Queensboro Bridge (QBB) and Queens Midtown Tunnel (QMT) produce and attract a significant amount of traffic. The problem is further exacerbated by midblock activities such as taxis dropping off passengers, double parking, commercial vehicles loading and unloading. These result in a complex and challenging traffic flow environment for effective signal control and congestion management.

Extensive travel time data, collected by the ETC tag readers deployed along the roadside, are utilized in the development of the methodology in order to derive representative congestion measures for control purpose. In an urban roadway environment, travel times are influenced by traffic flow states, signal controls, and random incidents such as taxis dropping off passengers, double parking, street parking, vehicles entering or exiting garages, crossing pedestrians and cyclists. Travel times vary by time of day and day of week, and the travel time distribution is dependent on the traffic conditions at the time of departure and throughout the trip.

Figure 1 shows a scatter plot of 24-hours travel time observations of an 8-block segment (2103 feet in length) from Lexington Ave/57th St. to Lexington Ave/49th St, New York City. There are clusters representing different congestion levels with some sparse samples of significantly higher travel time (greater than 1800 seconds), indicating outliers possibly caused by the above mentioned random incidents. Apparently these outliers do not reflect prevailing traffic conditions and should be filtered before feeding the data to a control strategy. Figure 2 shows a whole week per-trip travel time
observations in scatter plots. It shows that congestion patterns on weekdays are different from weekends. On weekdays, the traffic is more congested during the day time, while on weekends it is more congested during the evening hours. Among weekdays, congestion patterns also vary by the hour and by the day.

In this work, compiled per-trip travel times are used as a system congestion measure for evaluating overall congestion levels, while queues, estimated from local volume and occupancy data are employed as a local congestion measure. Per-trip travel time data are collected by ETC tag readers. Due to the high usage of ETC tags in the region, approximately 1 million per-trip travel time data are collected daily. These data are used in real-time control, and archived in a central database for offline analysis. Volume and occupancy data are collected by non-intrusive microwave sensors, aggregated in 30 seconds interval. Intrusive detection is not an option due to the high volume of diurnal and nocturnal traffic.

Based on the congestion measures, two levels of control concepts are developed for the purpose of preserving capacity and avoiding breakdown within the zone. The philosophy is to regulate traffic demand to alleviate the concentration of traffic at a network level, while at the same time making signal adjustment at selected critical intersections to prevent propagation of spillovers caused by local queuing. These two levels of control work in concert to accomplish the performance goal of improving mobility and travel time reliability.

Level 1 control, termed “Strategic Adaptive Pattern Reference Control” employs a library of timing plans, each characterized by a different set of offsets and splits for a cluster of intersections peripheral to the target control area. The offsets and splits represent a set of steady-control state measures to exert a tapering and rebalancing effect on the traffic before delivering it to the target control area. Depending on the prevailing congestion level, measured by metrics derived from real-time travel time data, a different level of tapering or rebalancing plans will be applied to alleviate congestions. This is distinct from Traffic Responsive Plan Selection (TRPS) of the 1.5 generation UTCS, in that the plans represent the reference patterns, each with a different degree of tapering and rebalancing effect on the traffic, and the selection of the appropriate pattern is based on systematic congestion measures (i.e., travel time and its derivatives) rather than isolated thresholds of local volume and occupancy.

Level 2 control is tactical critical intersections control inside the target control area. Splits are dynamically adjusted to balance queue-storage ratios at selected intersections. The purpose is to rebalance operations at competing approaches to prevent propagation of spillovers and the developing of “choke” points with stabilized or diminished queues. The control strategy employs a new local congestion metric called Severity Index, which defines a non-linear mapping from occupancy-volume space to discrete control regimes reflecting the relative queuing dynamics.

The following sections elaborate on the structural components of the actual system implemented in New York City.
Figure 1. Per-trip travel time observations (blue) with 15-minute Box and Whisker chart (black).
Locations: Lexington Ave/57th St. To Lexington Ave/49th St (2103 feet in length), New York City.
Figure 2. Whole-week per-trip travel time observations.
Locations: 3rd Ave/49th St. to 3rd Ave/57th St (2013 feet in length), Manhattan

Date: the week of Jun 18 - Jun 24 2012. Total 36260 trips; trip length: 2013 ft, 3rd Ave/49 St. to 3rd Ave/57 St., Manhattan

Figure 2. Whole-week per-trip travel time observations.
Locations: 3rd Ave/49th St. to 3rd Ave/57th St (2013 feet in length), New York City.
4. SYSTEM ARCHITECTURE

The active traffic management methodology is implemented as a software system called ACDSS (Adaptive Control with Integrated Decision Support System), which works as a general plugin for a traffic management centre (TMC) environment, through a novel web-service based application programming interface (API). Figure 1 illustrates the system architecture of ACDSS, integrated in a TMC. As shown in this figure, field controllers, sensors, and cameras communicates with an ATMS deployed at the TMC. ACDSS interfaces with the ATMS through a web-service based API, which is specifically designed in this work to meet the needs of reliable and robust control. Real-time data are processed and feed to the control strategies. Signal control parameters are generated and sent back to the ATMS through the web-service APIs. The latter translates the information in NTCIP protocol to be implemented by the field controllers. Figure 4 shows this architecture with more details on the active traffic management component.

4.1 Communication by Web-Service based APIs

Integral to the system architecture is a web service based Application Programming Interface (API). It is depicted as Component G in Figure 4. This API allows the active traffic management logic to be transparent to hardware, and enables future upgrades of the software and hardware independent of each other. ACDSS connects to the web service server using the XML-based SOAP standard. The web service server, on the other hand, connects to one or more NYCTCS servers that communicate with the controllers and sensors in the field use NTCIP, one of the standard protocols of the ITS industry. From system engineering point view, this “loose coupling” of different components greatly improves the reliability and reduces maintenance and upgrade costs. No proprietary standards are employed, nor does any customized MIB (Management Information Base) need to be installed in controllers.

In terms of functionalities, the interface has two goals. The first is to provide real-time intersection and detector data from the NYCTCS systems to ACDSS, and to enable latter to calculate optimized signal timings. The second goal is to allow ACDSS to send the optimized timing data to NYCTCS systems for implementation in the field. For Level 1 control this means the ID of the relevant plans and for Level 2 control, optimized splits. New signal timing parameters are implemented by the controllers at the earliest next local zeros.

4.2 Data and Information

*Detector Data and Phase Status Data*

The web service interface provides sensor data (volume and occupancy) and controller status data to ACDSS when requested by the latter in real time. By design the interface does not limit the resolution of the data, and it is possible that the sensor and controller status data are provided on a second-by-second basis. In the NYC implementation, the active traffic management logic requires 30 seconds aggregated occupancy and volume data to be transmitted in real time. The aggregation interval is set as 30 seconds because higher resolution second-by-second data are not required by the underlying control algorithms. Controller status data are reported as soon as a cycle finishes, including the start and end time of each cycle, observed splits and offsets. The controller status data also include flags of whether the controller has just recovered from a communication failure, in transition, offline, running TOD or adaptive, and if running TOD, the current plan ID. These flags are reported immediately when the relevant event occurs. Volume and occupancy data are aggregated in the field in 30 seconds interval so that communication channels are not overburdened with details. Key information is available and can be aggregated at the receiving end as needed.
Travel Time Data

Segment travel times are collected using ETC tag readers deployed along the roadside. Individual per-trip travel time is identified by matching the tag ID at the two ends of the trip, and published to an internal secured telnet server as soon as the trip finishes (See Figure 3). Each travel time record includes a scrambled tag ID, starting and ending points of the trip, and the duration of the trip. Due to the high usage of ETC tags, the system receives approximately 1 million records of per-trip travel times from 70 travel time segments, ranging from 200 feet to 5424 feet in lengths. High volumes of per-trip travel time data provide the statistical basis for identifying prevailing congestion levels required by the strategic Level 1 Control.

Travel time data are processed on a rolling window of 15 minutes for each road way segment where data are available. Median of the 15 minutes travel time samples are taken as the prevailing travel time measure because it is the natural center of the clusters of data points, it is also not biased by outliers (i.e., unusually longer travel time caused by double parking, vehicle entering and exiting garage etc.) without the need of prescribing an artificial travel time distribution beforehand.

Signal Commands

ACDSS uses a set of signal control strategies for the active traffic management. For Level 1 Strategic Pattern Reference Control, the signal plans are selected from a pre-defined library, hence only the relevant plan IDs and their associated schedule (i.e., the duration of time that the plan should be active) are send in real time to the field controllers. Level 2 Control requires more communication overhead as it needs to send back the detailed splits for each controller. Regardless, these data are first converted to SOAP objects then translated into NTCIP objects and sent to the controllers to be implemented in the immediate next local zero.

Inventory Data

Inventory data such as the list of sensors deployed, and the static information of each sensor are available via the web service interface. For each controller, the TOD plans and schedules are also accessible in real time. This static information is used for offline configuration and useful for online verification.

Real-time Data Warehousing

Millions of records of real time data are collected daily. To manage these data, a master database server (See Figure 3) is deployed that acts as the central data warehouse archiving all detection data, travel time data, controller status data and inventory data. Operation logs, system errors, diagnostics, and summary statistics are also handled by this database server. The master database server also provides data mining functionality for the ACDSS system to explore patterns and for offline analysis. The master server is replicated on two other servers for data redundancy. In order to boost run time computational performance, a high-speed in-memory database is also implemented (See Figure 3). This in-memory database only stores short-term data up to a few hours and acts as a high-speed cache for the data needs of the active traffic management.

4.2 Measures

Even though travel times are a reliable indicator of system congestion levels, they may not be promptly evident for operators in the control room that may have to make on-line decisions, given only a limited window of time for decision. Number of stops per trip is a more practical yet intuitive metric. It reflects both average queuing (which correlates to mean travel time) and the dynamics of queues (which correlates to travel time variability). By examining the clusters in the travel time data, empirical procedures are developed to convert sample groups of travel time data to number of stops. This provides the basis of different demand regulating strategies in the next section.
A new metric called Severity Index is developed in this work to reflect local congestion level. It is based on estimated queue storage ratio (QSR), computed as queue length divided by relevant link length. Severity Index is thus defined as:

\[
SI = \begin{cases} 
1, & 0.0 \leq QSR < 0.3 \\
2, & 0.3 \leq QSR < 0.6 \\
3, & 0.6 \leq QSR < 1.0 
\end{cases}
\]
Figure 3. ACDSS System Architecture.
Figure 4. Active Traffic Management using Adaptive Signal Control.
4.3 Control Strategies

Two levels of control are implemented under the proposed active traffic management methodology. Level 1 Control uses signal offsets and splits to exert a tapering or balancing effect on the traffic. This way, traffic is regulated so that the storage space is better utilized before delivering the traffic to the subject control area. It is also anticipated that by virtue of regulating traffic using Level 1 Control queues inside the target area can be better managed at the local level, with reduced travel time variability and improved throughput. Specifically, the control strategies under Level 1 control include:

- NBP (Network Balancing Plan) that simultaneous offsets are implemented on approaches to the target control area with minimal green window tapering;
- AC1 (Level-1 Access Control) that simultaneous offsets on approaches to the target control area with increased green window tapering;
- AC2 (Level-2 Access Control) that reverse offsets on approaches to the target control area with increased green window tapering.

These strategies have an ascending order in terms of the severity of congestion levels. Each strategy has a corresponding triggering condition based on the system congestion measure introduced above. The system assesses the prevailing traffic conditions every 3 minutes. If a higher level of strategy is justified, new strategy will be implemented, or presented to the control room operator for his or her decision. Once a new strategy is implemented, a 30-minute “pattern hold” period becomes active during which the system dwells on the new strategy for at least 30 minutes till the prevailing traffic conditions dictate a new pattern that calls for a different strategy.

Level 2 Control is tactic critical intersection control inside the subject control area where splits are dynamically adjusted to prevent spillovers caused by local queues.

4.4 Operator-in-the-Loop

In addition to running autonomously, the ACDSS system implements a mode of operator-in-the-loop, where the operators can review, override and interact with important signal timing decisions in response to traffic pattern changes or switch of control policies. This mode helps achieve robust and reliable signal operations for eventful situations, confidence building, and making tactical control decisions such as determining the point of shifting strategies. For example, for Level 1 control, the operator with the aid of real-time GIS map which has all travel time segments and sensor data color-coded and live video feeds, can decide whether to continue with a NPB (Network Balance Plan) or turn to a more restrictive AC1 (Level 1 Access Control). Figure 5 and Figure 6 illustrate the decision dialog and the control intervals with operator-in-the-loop.

Figure 5 shows that at the decision point, the operator is presented with the proposed control strategy, the existing control, and the current traffic conditions manifested by the color coded travel time map and converted number of stops. The operator can verify the situation by examining live camera video. If the operator rejects the proposed strategy, the system continues with the existing one, while keeps assessing the traffic conditions every 3 minutes, till the prevailing conditions request a new strategy again (See Figure 6). If a new strategy is implemented, the system holds the pattern for 30 minutes before restarting the assessing process.
Figure 5. Decision Support System – Level 1 Control Display.

Figure 6. Decision Support Process.
5. DEPLOYMENT

The new active traffic management methodology has been implemented in Midtown Manhattan, with project code named “Midtown in Motion” (MIM). The target control area of the current phase is a core 110 square block zone in the central business district, from 2nd to 6th Avenues, and from 42nd to 57th Streets, inclusive. For the study area, a total of 110 intersections are involved, with travel time data supplied by 32 ETC tag readers, and volume/occupancy data by 100 microwave sensors. The study area highlighted by the block is shown in Figure 7.

The first phase of Midtown in Motion has resulted in an overall 10% speeds improvement as suggested by the E-ZPass travel time data and verified by independent taxi GPS data. Currently it is being expanded to more areas, with a total of 210 microwave sensors, 56 traffic cameras and ETC tag readers at 59 intersections, covering the Midtown Manhattan area from 1st to 9th avenue and from 42nd to 57th street, more than 270 square blocks (more than 300 controlled intersections). Figure 8 shows the expanded control area.

6. CONCLUDING REMARKS

This paper summarizes the new active traffic management methodology developed and implemented in New York City. Beyond the implementation scale, unique ITS environment and technical challenges it has resolved, the work is perhaps the first active traffic management framework that uses travel-time based multi-regime adaptive control for congestion management. The design, development and implementation of the methodology are driven by the availability of extensive real-time travel time data, the upgraded ITS infrastructure and the need of a reliable and robust real-time adaptive signal system for congestion management in a complex urban traffic environment. This paper focuses on the system architecture. While the system is currently being expanded to cover more areas in Midtown Manhattan, comprehensive evaluations and further analysis are beyond the scope of this paper and are explored in other and future publications.

ACKNOWLEDGEMENTS

The work was sponsored by Federal Highway Administration (FHWA), New York City Department of Transportation (NYCDOT) and KLD’s Research and Development program. The authors wish to acknowledge Steve Galgano, John Tipaldo, and other engineers from NYCDOT including Rami Khashashina, Boris Skulskiy, Rachid Roumila, and Jimmy Haro. John Tipaldo’s insights on the ETC travel time data are especially appreciated. Special thanks to Robert Rausch, Jeff Mayo, and James Martin of TransCore for the ACDSS system integration and to Raymundo Martinez and others at JHK Engineering, P.C. for their support in the ACDSS system development and testing. The KLD authors would also like to acknowledge New York State Energy Research and Development Authority (NYSERDA) for the funding (Joseph Tario, project manager) of the original IMPOST signal optimization algorithm and of the on-going ACDSS commercialization. The KLD authors acknowledge their senior colleague W.R. McShane for his valuable advice and significant guidance throughout the work.
Figure 7. – Initial Midtown in Motion Study Area (Phase A), Microwave Sensors, Tag Readers, and Cameras.
Figure 8. Midtown in Motion Expansions
REFERENCES


