cuta actu

Canadian Urban Transit Association

PROCEEDINGS

The International Conference on

AUTOMATIC VEHICLE LOCATION IN URBAN TRANSIT SYSTEMS

September 19 – 21, 1988
Ottawa, Canada

pp. 189 - 200

Transport Canada Transports Canada

Canada

Using Automatic Vehicle Monitoring Systems to Derive Transit Planning Data

Peter G. Furth
Associate Professor of Civil Engineering
Northeastern University
Boston, MA 02115

Abstract

Data valuable to transit planners may be derived from automatic vehicle monitoring (AVM) systems. A theoretical analysis shows estimation errors in schedule adherence at timepoints and in running time to be minor. However, when a simple AVM system transmits passenger activity in real time, the potential for mismatching stops and for wrongly combining the activity of adjacent stops is shown to be high unless a very short polling cycle is used. Measures to improve stop resolution, including odometer stamping, storage buffers, and variable rate polling are analyzed as well. A comparison favors on-vehicle storage of passenger information to real-time transmission. Some negative practical experience in extracting planning data is also reported.

AVM and Transit Planning Data

An often cited potential benefit of automatic vehicle monitoring of bus systems is obtaining planning data. Because they track each vehicle's location in real time, AVM systems should be able to measure schedule adherence (actual vs. scheduled time at specified time points) and segment running time, both of which are valuable to planners. If the AVM system is equipped with passenger counters, it is capable of measuring ons and offs by stop, from which one can derive such measures as maximum load point location and passenger miles. However, each transmission channel between bus and control center is shared by many buses , typically with each bus being polled in turn. Events occurring to a given bus between pollings must be summarized in as few bits of information as possible to minimize the amount of time needed per poll. Depending on the form of summarization and the interval between polls, it may be necessary to estimate what happens to vehicles between pollings in order to derive the desired planning statistics, introducing estimation errors. This paper attempts to quantify these estimation errors, and to evaluate AVM configurations that can reduce estimation error. Other problems with using AVM to derive planning data are described as well.

Common Features of AVM Systems

AVM-equipped buses have on-board microprocessors that receive and store information from various on-vehicle devices. These devices typically include a digital odometer that increments by hundredths of a mile (or kilometer), sensors of mechanical measures such as oil pressure, and a silent alarm. The sensors and alarms control yes/no flags in the microprocessor, and each odometer increment increments the microprocessor's internal odometer. When a signpost is encountered, the microprocessor stores the signpost ID and resets the internal odometer. When requested by radio signal from the control center, the bus transmits back to the control center the stored data, from which the central computer determines the vehicle location and responds to abnormal status flags. The control center sends out one such request to each bus in turn once every polling cycle.

In some AVM systems, buses are also equipped with passenger counters. They may either be infrared detectors or pressure sensitive mats that count boarding and alighting passengers. Each detection of a boarding or an alighting passenger increments its respective counter in the microprocessor. The microprocessor also possesses logic to detect when a stop dwell is completed; typically, it must sense the doors close and the odometer advance by one increment. When polled, the reply will indicate whether a stop has been completed, and values in on and off counters. If the stop has been completed, the stop completion flag is reset, and ons and offs at the stop are calculated by subtracting cumulative ons and offs at the previous stop (or else the on and off counters are reset).

Several problems are endemic to AVM systems as described. Tall buildings and canyons can cause regular transmission problems. Transmission problems can occur intermittently for other reasons as well. And while odometers are generally sufficiently accurate, passenger counter accuracy is far from perfect, although improvements are being made.

Estimating Schedule Deviation and Running Time

Transit planners and schedulers would like to know whether buses are on time at certain specified time points in order to be aware of existing problems and to monitor countermeasures taken to improve on time performance. While an AVM system may be perfectly accurate in locating a bus at the moment of polling, the time at which it passed or departed from a given location must be inferred from the location at the pollings straddling the timepoint. The maximum error depends not only on the length of the polling cycle, but also in the distance covered during the polling cycle. For example, if the cycle is 60 sec and the distance covered is 0.5 mi, and if the maximum cruising speed is 30 mph, then there is only one possible vehicle trajectory during the cycle, a constant speed of 30 mph between the two points of polling. In such a case, the time at which the bus passed any location within that cycle is known without approximation as shown in Figure 1a. However, if the vehicle traveled only 0.2 mi in a 60-sec cycle, the vehicle trajectory is uncertain, as shown in Figure 1b. The vehicle could have rested at the point of the first polling (point 1) until the last possible moment, then accelerated to 30 mph and passed the point of the second polling (point 2) at full speed, as indicated by trajectory (A); or it could have passed point 1 at full speed and then decelerated to a stop at point 2 (trajectory B); or it could have has a trajectory lying in between trajectories (A) and (B), possibly including a stop between points 1 and 2 (trajectory C).

The range of error in the estimated time at which the vehicle passed (or departed from) a point is given by

$$r = c\,(1 - \bar{u}/u_{max})$$

where

c= polling cycle
$\bar{u}$ = average speed during polling cycle
u_{max} = maximum reasonable cruising speed

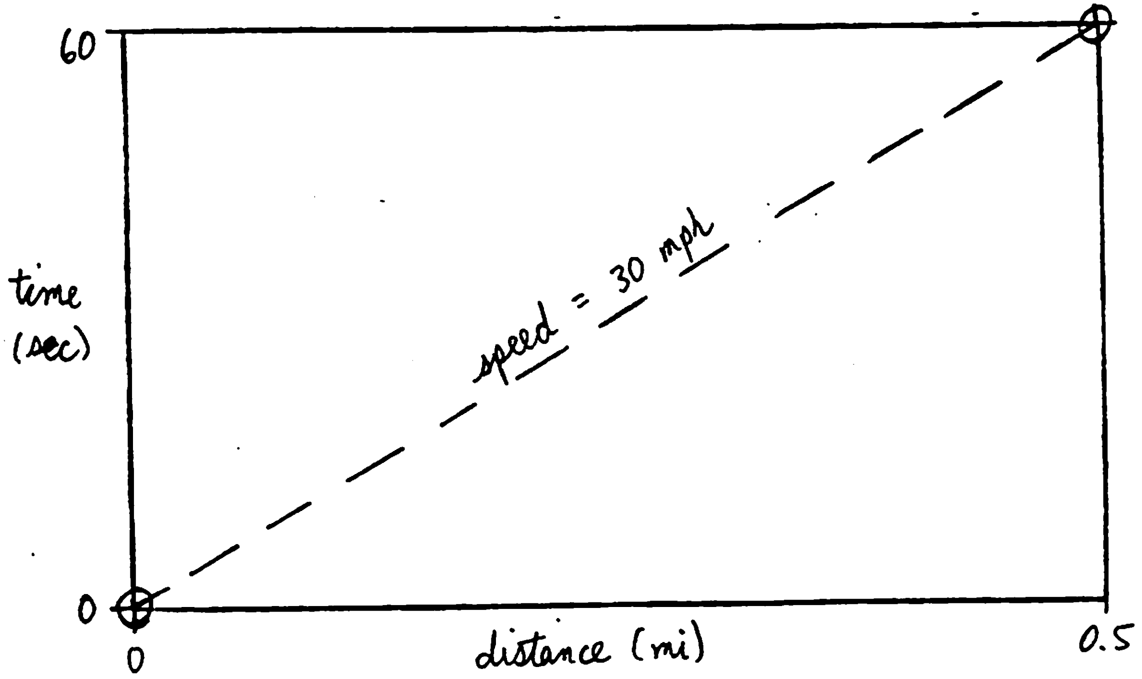
For example, with an average speed of 9 mph (common in city centers) and a maximum cruising speed of 25 mph, the error range is 64% of the polling cycle. If the estimated time

Figure 1
Vehicle Trajectories Between Pollings

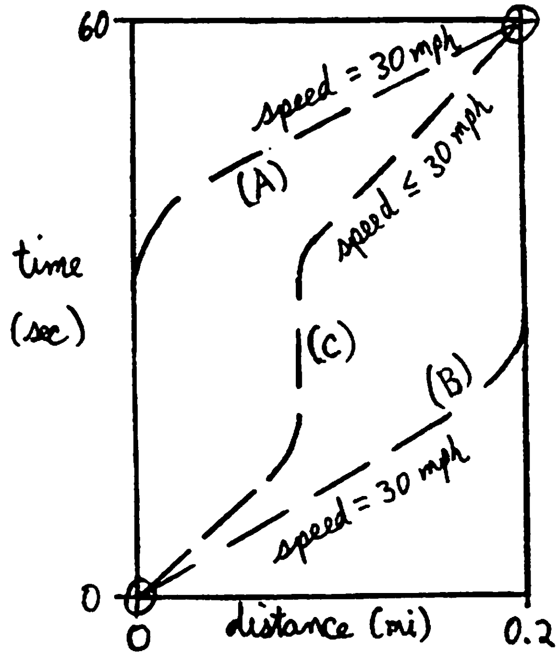
Given: polling cycle = 60 sec, maximum cruising speed = 30 mph

⊕ = known point in time and space (time, location when polled)

a) Distance covered = 0.5 mi: only one possible trajectory



b) Distance covered = 0.2 mi: many possible trajectories



is placed in the middle of the error range, the maximum absolute error will be 32% of the polling cycle. For a 120-sec cycle, this represents a maximum error of only 38 sec. Since the timepoint is likely to be a stop with passenger activity, the extreme trajectories (A and B in figure 2(b)) are unlikely, reducing the expected error. Because schedule deviation is typically measured and reported by whole minutes (or, at the finest, by half minutes) this level of error should be insignificant, especially when considering the vastly greater number of observations an AVM system will yield for calculating summary statistics compared to manual data collection. With polling cycles of 4 min or greater, the estimation errors can be significant; however, cycles this long are rare because they are inconsistent with other purposes of the AVM system.

To estimate segment running time, the difference is taken between estimated time at two timepoints. Because errors can compound, the maximum error range for running time is twice the range for time at a timepoint, or about 76 sec for a center city segment using the prior assumptions. However, the chances of this worst case error occurring are small. Assuming a uniform error distribution, the 90 percentile absolute error is only 68% of this maximum error (or 52 sec for the example), and the mean absolute error is only third of the maximum error (25 sec). Again, considering the number of observations an AVM system will yield, the estimation error in any summary statistics (such as mean running time in a given period) will be minor.

One device that can increase accuracy in measuring time at a timepoint is for the on-board microprocessor to record its arrival or departure time (whichever is desired) at each stop and to include this item in the transmission to the central computer. With this device, time at which a bus reaches or leaves a timepoint can be known without inference if the bus stops there and the stop location is correctly determined. Transmitting stop time will require more time for each poll, however, decreasing slightly the number buses that can be polled at a given rate with one radio channel.

Stop Matching Without Odometer Stamp

In the simplest AVM configuration that includes passenger counting, is is assumed that stop location will be derived from the regular transmission of location information. However, regular location information only pinpoints (with imperfect accuracy) the location at the beginning and ending of each polling cycle. When a bus reports at a polling that it has completed a stop, the location is known only to lie before the current location and at or following the previous location. If therefore more than one stop has been passed during the polling cycle, there can be a matching error. For example, suppose there is passenger activity at stop 1. The bus is polled after departing stop 1, just before the first odometer increment, and so stop 1 activity is not reported at this poll. The next poll occurs after the bus has passed stop 2, at which no passenger activity took place. The passenger activity may be incorrectly attributed to stop 2. To prevent this type of error, it is necessary that the cycle length be short enough that the bus cannot reach stop 2 before the next poll. Assuming that the odometer increment is small enough that the bus will still be accelerating when the odometer first "clicks", and assuming constant acceleration to a constant cruising speed, the maximum polling cycle can be derived thus:

minimum polling cycle = time needed to go between stops at cruise speed
+ delay due to acceleration
- time spent during first odometer increment

$$= d/u + u/2a - \sqrt{2i/a} \qquad (1)$$

where

i = length of odometer increment
a = acceleration rate
u = cruising speed
d = interstop distance

With typical center city values of $d = 0.125$ mi, $u = 25$ mph, $i = 0.01$mi, and $a = 3$ mph/sec, the minimum polling cycle length is found to be 17 sec. Since AVM systems typically aim for a polling cycle of around 120 sec, adjusting the system to get a 17 sec cycle will require seven times as many radio channels, seven times as much data processing,and so forth, increasing the system cost and threatening its feasibility.

Stop Matching and Stop Splitting With Odometer Stamp

If the on-board microprocessor stores the odometer reading when the stop is completed and transmits this reading with the poll, then stop matching when only one real stop occurs in a cycle is no problem. There is a cost to this enhancement, however; a few more bits must be transmitted with each poll, decreasing the system polling capacity, thereby decreasing the number of buses per channel that can be served at a given polling cycle.

However, it two stops are completed during a polling cycle, their ons and offs will be combined. The first problem is how to know that two stops have occurred. Adding a stop counter (again, costing another bit or two of transmission time) can solve this problem; otherwise a guess must be made, perhaps based on average speed during the cycle. Even if it is known that multiple stops occurred, splitting the ons and offs cannot be done without estimation error. To avoid estimation error, then, the polling should be no longer than:

maximum polling cycle = minimum polling cycle from equation (1)
+ deceleration delay at second stop
+ dwell time at second stop (2)

Using values from the previous example, and with a deceleration rate of 3 mph/sec and a minimum dwell time of 9 sec, the maximum polling cycle is only 30 sec.

It should be recognized that the consequences of incorrectly splitting a stop or matching a stop incorrectly to a neighboring stop are minor. If reasonable guesses are made, errors in statistics such as passenger-miles or mean load at a branching point will be small. Nevertheless, planners and schedulers are known known to be very suspicious of data that include estimates, and may choose not to use it, making it useless (except for some statistics like total boardings which are not subject to these types of errors). Furthermore, the magnitude of these errors may be in fact significant if the polling cycle is four times or more longer than the value that prevents matching errors, perhaps allowing three or more stops to be mixed. On the other hand, during congested periods, when the data are perhaps most crucial, cruise speeds decline and dwell times expand, reducing the potential for mismatching stops.

Variable Rate Polling

Recognizing that planners do not need stop-level data on every trip every day, at least one manufacturer offers a system that will poll a subset of buses at a faster rate. By rotating the fast poll group around the system, a sample of days with detailed data can be obtained for every route. Even sampling a route once a week or every other week is still far superior to most manual sampling programs, which rarely provide more than one ride check and/or eight point checks per year. In days when a route is in the background poll set, it is still possible to measure total boardings per trip and to get a good estimate of peak load. The sampled days with a fast poll would be used to generate average ride check profiles, passenger-mile statistics, and other ride-check based data.

The polling burden of a variable rate system is given by

$$B = 60N\,[(1-p)/c_b + p/c_f]$$

where B = polling burden (total polls per minute)

N = number of vehicles
p = fraction of fleet on fast poll
c_b = background polling cycle (sec)
c_f = fast poll cycle (sec)

Table 1 shows the polling burden per 100 buses for different parameter values. The number of radio channels needed and many other system size measures are proportional to the polling burden.

Applying variable rate polling imposes an extra planning and administration burden. Because it is desirable to fast poll entire routes on the same day, bus ID's cannot be selected for the fast poll either at random or according to any preset scheme. The ID's of the buses operating on the routes to be sampled on the next day must be reported every night, and entered into the system. Last minute substitutions should be accounted for as well. Fast poll data should be kept separate from background poll data, with the capability of combining them if desired.

Multiple Stop Transmission

Another way to overcome the need for a short polling cycle is for the system to store and transmit two stops' worth of data, including an odometer stamp at each stop. The maximum polling cycle will be more than twice the maximum for single stop transmission with odometer stamp (calculated as 30 sec in a previous example), since the chances of two worst-case stops following one after another is remote. The maximum polling cycle for two stop transmission, using the previous example's data, should be around 70 sec, allowing for at least 10 sec of traffic delay. The length of each poll will increase, however, to convey the needed information; the number of bits transmitted from the bus to the polling center will increase by 30 to 50 percent, increasing total transmission time per bus by around 20 to 35 percent (since time needed for signalling each bus and for signal spacing is not affected). Overall, then the practical capacity increase of two-stop transmission will be around 75%, since the longer transmission time is offset by the more than doubled allowable polling cycle; that is, the same amount of radio transmission resources could cover about 75% more buses without danger of erroneously mixing or mismatching stops. Using three-stop transmission will allow still more economical use of transmission resources. With a 110 sec polling cycle and an estimated increase (relative to one-stop transmission) in transmission time per bus of 50 to 70 percent, the number of buses that could be covered per radio channel more than doubles.

TABLE 1

POLLING BURDEN USING VARIABLE POLLING RATE

Background polling cycle (sec)	Fast poll cycle (sec)	Fraction with fast poll	Polling burden per 100 buses (polls/min/100 buses)
NA	15	1	400
NA	30	1	200
60	NA	0	100
120	NA	0	50
180	NA	0	33
60	15	0.1	130
60	15	0.2	160
60	30	0.1	110
60	30	0.2	120
120	15	0.1	85
120	15	0.2	120
120	30	0.1	65
120	30	0.2	80
180	15	0.1	70
180	15	0.2	107
180	30	0.1	50
180	30	0.2	67

Using a Stop Buffer

Another device for improving transmission efficiency that has been used by at least one manufacturer is adding a storage buffer to the on-board microprocessor that allows several stops worth of data to accumulate. Only one stop worth (the oldest one in the buffer) is transmitted at each poll. There is therefore no additional polling burden, except of course the need for an odometer stamp. The presence of a buffer allows the polling cycle length to more closely approach the average interstop time rather than the minimum interstop time, giving occasional short segments and random fluctuations less influence over the polling rate. In a typical center city operation, average speed, including dwell time, rarely exceeds 8 mph, while stop density rarely exceeds 8 stops per mile, yielding a minimum average interstop time of 56 sec. The larger the buffer, the closer the polling cycle can approach this limit without there being a significant probability of buffer overflow (in which case the ons and offs of two stops would be combined, which is still not a major error if it happens infrequently). Using the example data presented earlier, satisfactory system performance with a 4-stop buffer could probably be achieved with a polling cycle whose length is 80 to 85 percent of the average interstop time, or around 45 sec. This increase from the previous value of 30 sec represents an increase in practical transmission capacity of 50% over single stop transmission. Using a buffer is particularly effective if bus routes go through relatively small regions where stops are close together (in time). In such cases, the allowable polling cycle can perhaps double, doubling the practical transmission capacity.

Using a buffer puts additional requirements on the on-board microprocessor, increasing system cost. Another impact is that the stop information can be lagged with respect to real-time location information, making it harder to know passenger loads in real time. For example, when a 4-stop buffer is full, if the polling cycle is 45 sec, stop data can be lagged by as much as 3 min. This should be of little consequence to real-time control, however. Furthermore, it is unlikely that the buffer will be full if the system is configured correctly; most of the time the buffer should be near empty.

On-Line Versus Off-Line Stop Data Transmission

An alternative to on-line transmission of stop data is to store the stop data on board the vehicle, and transmit it in batch mode either daily or every few days to a central computer for processing. This is the way automatic passenger counter (APC) systems that do not have real-time vehicle monitoring operate. If a transit system has, or plans to acquire, an automatic vehicle location system and desires automatic passenger counting capability as well, it is certainly economical for the passenger counting system to share some components with the AVL system. Wayside signposts should certainly be shared, and it is probably economical to share the signpost receiver and a microprocess or on each vehicle as well. But real-time transmission of stop-level data puts an additional burden on the AVM system, particularly on the radio transmission component, if it is to be done to the same level of accuracy that can be attained with batch transmission.

There are three benefits of real time transmission. One is that the data is available in real time for control. However, some experienced dispatchers report that once they know (from location observations) how buses are behind schedule and/or bunched, they are fairly confident of which ones are overloaded without using direct load observations, and can issue instructions for appropriate remedial action. Direct load information would probably improve the quality of such decisions, but the level of the associated benefit, while hard to quantify, is probably small. The second benefit is that a separate transmission system is not needed. If the passenger counting data is stored on vehicle, it must be transmitted somehow to a central computer, either by batch radio communication, or batch telephone communication, or by dumping the data onto tapes or some other portable device and transporting them to the computer. Batch radio is probably the most difficult system. APC vendors can supply telephone or portable device transmissions systems that are compatible with almost any central computer. The third benefit is that the turnaround delay required by batch processing is eliminated, speeding up reports by one to several days.

Off-line storage and transmission of stop-level data has several advantages. The first is that the transmission component of the AVM system is not interfered with. Radio channels are rationed by government agencies, and some are usually desired for voice communication, and so even if the hardware cost of doubling or quadrupling the polling burden can be absorbed, doubling or quadrupling the number of radio channels devoted to data transmission may be infeasible. Second, errors of mixing or mismatching stops due to too long a polling cycle disappear. Third, transmission data losses due to tall buildings, canyons or hills, and equipment malfunction are eliminated. Fourth, customizing an automatic vehicle location system to handle passenger data, unless it has been originally built to do so, can be very expensive and time consuming and involve unforeseen complications.

While a cost analysis has not been done, based on a technical analysis it appears that, on balance, off-line transmission of stop-level data is superior to on-line transmission. Accuracy and quality of stop-level data will be better, cost is likely to be less, and there will be fewer difficulties in trying to integrate the systems. The only significant advantage of real-time transmission is that it offers real time load information, but the value of this capability is, as explained, debatable.

Extracting Planning Data Files and Report Generation

It would seem to be a straightforward task for a system that monitors vehicle location in real time, and perhaps passenger information as well, to write records to a file from which schedule deviation, running time, and passenger use statistics can be derived. To the disappointment of several U.S. transit systems, developing this reporting capability has proven either extremely difficult or impossible with AVM systems that did not originally have this capability. This author's experience is limited, and it may be that there is no such problem with other AVM systems. Those interested in extracting planning data from an AVM system should be warned, however, that this ostensibly simple task may not be simple after all, and should seek evidence of a vendor's capability to deliver this capability.