SIEMENS

TECHNICAL MEMORANUM #2
MAG ITS/TE On-Call Services
Contract No. 321-I
Glendale Stadium Area Congestion Map
Proof of Concept Project

May 2008

Version 2.0

Submitted To:

Maricopa Association of Governments

Submitted By:

Siemens

INTRODUCTION

Siemens has been contracted by the Maricopa Association of Governments (MAG) to explore the feasibility of generating a congestion map for the arterial street network surrounding the Glendale Stadium Area. The initial concept is of a congestion map to be developed based on occupancy feedback from detectors located at five signalized intersections on the east side of the Glendale Stadium area. Occupancy values would be extracted from the City of Glendale's i2 Traffic Management System. Algorithms would be developed to correlate occupancy values to congestion levels. Varying levels of congestion would be depicted by colored links on a map. The map would be available via the Internet. Congestion information could also be provided to drivers via color LED signs displaying link congestion ahead.

The first task of the project was to perform a field/hardware survey of the five signalized intersections around the Glendale stadium. Results of the survey are documented in Technical Memorandum #1. The second task is to investigate the feasibility of developing the required algorithms and software interface for generating a color-coded congestion map. A proposed algorithmic approach and corresponding modifications to i2 central systems software is the topic of this document, Technical Memorandum #2.

Task 2 was summarized in the Scope of Work as follows:

Task	**Description**
2.	Investigate if the system software can be modified to meet the project needs.
2.1	Define movement/lane group/approach.
2.2	Allow manual threshold inputs to define severe, heavy, moderation, light congestion and no data.
2.3	Allow manually defined smoothing factors.
2.4	Define reference point of the cycle and collect data by cycle.
2.5	Generate color code for each movement/lane-group/approach every 1 minute.

CONCEPT OF OPERATIONS

Background

MAG staff reported that their concept of operations for an arterial congestion map was inspired by a similar effort recently undertaken by the City of Bellevue, Washington.[1] The Bellevue system is briefly summarized as follows:

- The system utilizes existing advance detector loops located between 100 and 140 from the stop line.
- Detector occupancy data is collected and aggregated (per-cycle) values (smoothed over time) are used as the primary input to determine congestion.

[1] Fred Liang, *Development of the Real-Time Arterial Traffic Arterial Traffic Flow Map*. Presented at ITE District 6 Annual Meeting, Honolulu, Hawaii, June 2006.

- Manually configured thresholds are configured for each detector to map the smoothed occupancy values to a corresponding congestion level.
- An online map display is updated once a minute to provide a color-coded indication of the congestion on each monitored link.

Table 1 is an example of typical occupancy thresholds used by the City of Bellevue to select one of four congestion levels, depending on the location of detectors.

Table 1. Typical occupancy-to-congestion mapping (source: Fred Liang [1]).

Congestion Level	Two Lanes Detectors 125' back	One Lane Detector 125' back	Two Lanes Detectors 300' back	Two Lanes Detector 50' back	Two Lane with only one detector
Light	< 50%	< 45%	< 45%	< 47%	< 34%
Moderate	>= 50	>= 45%	>= 45%	>= 47%	>= 34%
Heavy	>= 80%	>= 68%	>= 65%	>= 90%	>= 54%
Severe	>= 90%	>= 78%	>= 75%	>= 95%	>= 64%

Bellevue reports a process of ongoing data calibration, with multiple field visits, user feedback, and continual fine-tuning to improve the accuracy of reported congestion levels.

Workflow Activities

The workflow for setup and operation the Glendale congestion map system is summarized in the following activities:

1. Site detectors
2. Configure system
3. Collect data
4. Process data
5. Update display

The following sections discuss Siemens proposed approach for each of these five activities, which differs slightly from the approach used by the City of Bellevue. Each section discusses the requirements and goals (or preferences), considered options, and proposed approach, making an effort to convey the benefits of the proposed approach.

SITING TRAFFIC DETECTORS

Goals

The goals of positioning vehicle detectors for congestion measurement are as follows:

- Provide an accurate measure of traffic demand/congestion.

- Utilize existing detector locations/technology to minimize deployment costs.
- Be flexible in the location of detection (setback, length, lanes covered).
- Keep detector installation and configuration/calibration costs down.

Discussion

The location of vehicle detection plays an important role in measuring signalized-intersection congestion levels. Figure 1 illustrates vehicle detectors as blue rectangles on all approaches to a signalized intersection. In particular, Figure 1 shows 6-foot advance loops positioned 120 feet from the stopline as used in the City of Bellevue. Figure 1 also includes 20-foot loops positioned at stopline locations, which may be more representative of current detection around the Glendale stadium.

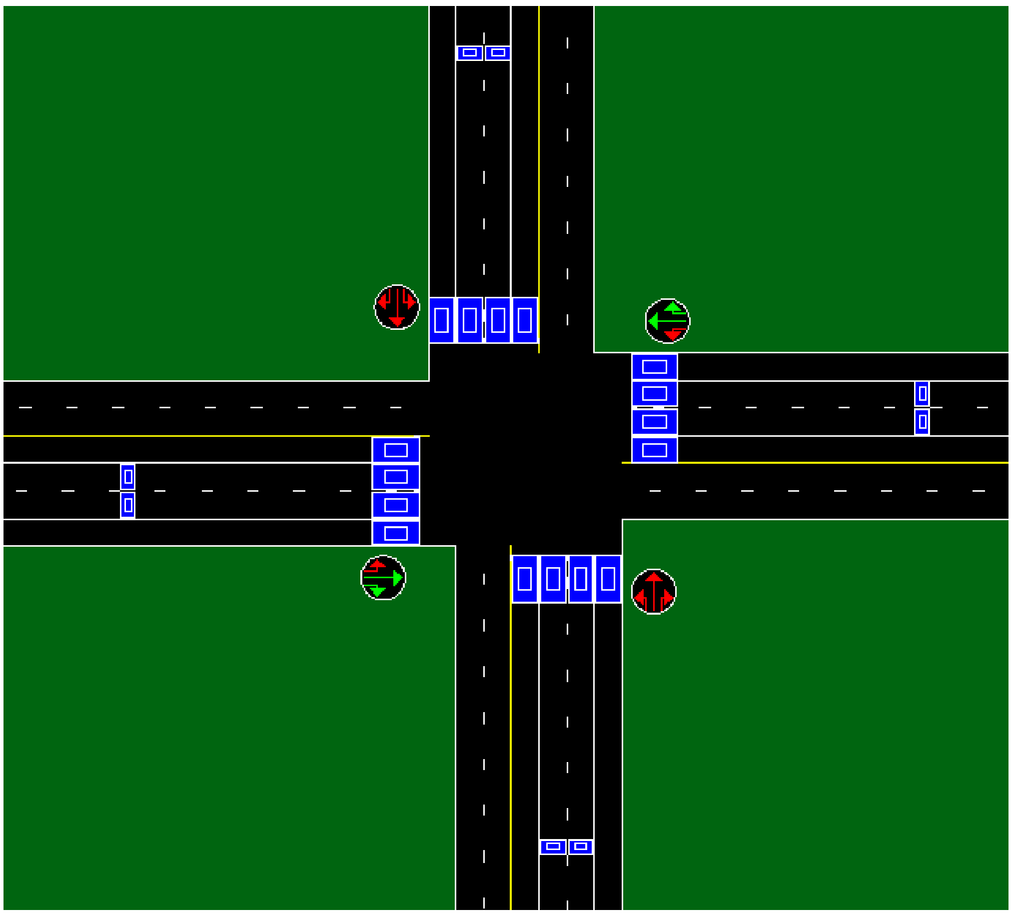

Figure 1. Location of vehicle detectors (blue rectangles) at a signalized intersection.

The City of Bellevue used advanced detectors located between 100 and 140 from the stopline to measure congestion, citing that they preferred to use existing advanced loops instead of following what they termed the "convention wisdom" of using system loops that are typically positioned at least 300 feet from the stopline. By leveraging existing loop location, the city was able to quickly bring 465 roadway segments online and saved an estimated $350,000 by foregoing the installation of system detectors on these roads.

Reviewing the current detection used at the five candidate intersections in Glendale (see Technical Memorandum #1), there are existing inductive loop detectors located at the stopline for two of the intersections, and video detectors on the other three intersections. It is most common (across the United States) for intersections to use stopline detection on left-turn and side-street movements, while main-street through lanes are relatively variable (no detection, advance dilemma zone detector, stopline detection, system detectors only, or a mixture). When using video detection, it is relatively common to define detection zones at the stopline and/or advanced dilemma-zone/extension locations zones for main-street through lanes. However, stopline zones may be preferable source of data when using video detection since cameras are susceptible to a greater frequency of errors when monitoring zones located farther away from the typical farside mast arm location of the camera. The City of Glendale has expressed a willingness to redefine video detection zones as necessary to facilitate congestion measurement.

Stopline detectors, advance detectors (located 50-300 feet from the stopline), and system detectors (located mid-block to just after the upstream intersection) could all be utilized to measure congestion. We propose a general purpose approach that could utilize all three locations, by providing a configurable weighted combination of occupancy and volume data, rather than solely using occupancy as in the City of Bellevue. There are advantages to this approach in terms of cost, ease, accuracy, and flexibility.

Stopline detection is the most prevalent location of detectors across the United States, and thus using these detectors would require the least incremental upgrade cost for most agencies (presumably including Glendale) when instrumenting a network for congestion measurement. We have considerable research experience with the use of stop-line detectors for demand/congestion measurement, and have applied it successfully in our ACS Lite adaptive control system. Our use of stopline data for this project would be much more simplistic than adaptive control applications; however, our experience, prototyping, and experimental results suggest that stopline detection would provide the most accurate and easily configured measure of traffic congestion during undersaturated conditions.

Advance detectors were used by the City of Bellevue to good success. Advance detection can better discern when conditions are *over*-saturated. However, it is more challenging to use in just-saturated or undersaturated conditions as it is much more sensitive to the cycle length in use. Generally, the longer the cycle length, the longer the queues, and thus longer cycle lengths correspond to more queue spillback over advance detectors, and hence higher occupancy than at shorter cycle, despite operating at equivalent volume-to-capacity ratios. The five intersections proposed for congestion monitoring in Glendale currently all run free (see Technical Memorandum #1) and thus there is much greater uncertainty in the actual cycle length and the interpretation of congestion measures for advance detectors at these signals. A combined approach of using stop-line detection and advanced detection would yield the best results.

We do not intend to use or experiment with system detectors for this project; however, it would be possible to use system detectors to estimate congestion levels. This would represent the most challenging option to configure/calibrate, and thus we do not recommend use of detectors at these locations unless there is no alternative.

Regardless of the location of detectors, more accurate information can be obtain from individual lane detectors than from detectors that span multiple lanes. Multi-lane detection could still be used where necessary, though individual lane detection should be used where greater accuracy is desired. In general, multi-lane detectors are likely to suggest higher congestion levels than are actually present. More detail on the use of detection data will be presented in subsequent sections.

CONFIGURING SYSTEM PARAMETERS

Goals

The goals entails in system configuration are as follows:

- Incorporate the information system to form reasonably accurate congestion measures.
- Keep it simple. It would preferable not to have to make several field visits, handle numerous callers saying its wrong, and doing a lot of trial-and-error adjustments … if that's possible to avoid.
- Allow flexibility for user configuration to manipulate/define the measures put out.

Recommendation

For each detector, configure the following parameters:

- Signal (to which it is connected)
- Detector number
- Link
- Movement (left, thru, right)
- Free flow speed (or speed limit)
- Setback distance (feet upstream from the stopline)
- Length (feet from leading to trailing edge)
- Lanes covered (1, 2, 3, or more)
- Sample period (e.g., period = 60 seconds)
- Minimum samples (e.g., at least 8 samples/minutes of data)
- Maximum samples (e.g., consider up to last 15 samples/minutes of data)

- Smoothing factor (suggest tail-weight = 1)
- Volume and Occupancy Weights
  - W_O is a weighting factor for occupancy.
  - W_V is a weighting factor for volume.
- Manual/Automatic Weighting Option
- Thresholds for mapping:
  - L_{max}: The highest measured values for which congestion level Low/Light (*L*) is indicated.
  - M_{max}: The highest measured values for which congestion level Medium/Moderate (*M*) is indicated.
  - H_{max}: The highest measured values for with congestion level High/Heavy (*H*) is indicated.
  - S_{max}: The highest measured values for which Severe congestion level Severe/Extreme (*S*) is indicated. Beyond this value would be considered a fault condition due to impractically high readings.

For each link-movement to be monitored, configure which signal-detectors to use, and whether take the AVERAGE or MAXIMUM of multiple detector indications. Configuration would also include the minimum number of detectors with fault-free data necessary to publish a congestion measure. If not enough detectors have good data, a "no data" indication would be published.

As will be explained in subsequent sections, this configuration should be flexible enough to allow the use of detectors in multiple locations, using multiple techniques. An "automatic" weighting option would allow a default weight to be applied, which would simplify the setup process and potentially alleviate the need for fine-tuning adjustments as explained later.

COLLECT DETECTOR DATA

Goals

The goals of detector data collection are as follows:

- Collect detector data (volume and/or occupancy) adequate to measure congestion levels with good accuracy.
- Collect data frequently enough that the resulting congestion measures reflect "current" traffic conditions.
- Use a technique that can collect data with consistently high reliability.
- Use techniques that are flexible/versatile in providing compatibility across a wide-range of controllers (different vendors, different models, different firmware), communications media (serial and IP), and protocols (AB3418, NTCIP, etc.). This facilitates cost saving (ability to use legacy hardware, twisted pair) and allow flexibility to choose from a wide variety of equipment providers.

- Use a technique which keeps the cost reasonably low for central system modifications and new data processing software.

We mentioned previously our success using stopline detection data to measure congestion using ACS Lite. However, our use of stopline detection data for this project would differ from the ACS Lite technique, as would the data collection method. Our ACS Lite work built on prior research from the 1970s for the SCATS adaptive control system from Australia, which also relies on stop-line detection. The SCATS system is primarily oriented towards measurement of occupancy and volume during green intervals (as is ACS Lite). However, such processing presents an additional burden on the signal system, communications systems, and local controllers in order to monitor signal/phase states very closely and match up phase-indications with detector data. Most controllers in the U.S. do not provide adequately reliable second-by-second polling for detector status. Controllers supporting special adaptive traffic control protocols can provide this information; however, this generally requires more bandwidth. With regard to the five candidate signals in Glendale, there are two Econolite ASC2 controllers, which would (with an available firmware upgrade) support the ACS Lite protocol. The ACS Lite protocol is also lower bandwidth than other adaptive protocols, and can be supported over (low-speed) serial communications. However, the ACS Lite protocol is not (currently) supported on the ASC3 controller, used at the other three intersections. The ASC3 has an experimental data logging system (not available on ASC2 controllers) to provide more resolute data, though the data is provided in large hourly files that must be retrieved via FTP, and thus this mechanism is not applicable for serial communications and does not provide reasonably current traffic conditions (which would be less than one hour old). We would propose to use data aggregated at intervals such as 60-seconds (not 1-second or 1-hour intervals) similar to the City of Bellevue (which used per-cycle aggregated occupancy values). We have experimented with a technique and found it to provide very good congestion indications. The data collection is discussed in this section, and processing in the next section.

The proposed data collection technique is to collect detector volume and occupancy data on fixed intervals, such as 30-second or 60-second sample periods. This capability is a standard feature using AB3418 and NTCIP protocols, and is thus supported by the majority of contemporary controllers on the market as an off-the-shelf feature. In the following paragraphs, we briefly review the City of Bellevue's data collection approach, and justify that our proposed approach adequately satisfies project goals.

Bellevue's Data Collection

The City of Bellevue currently collects data once per cycle (collecting data at the yield point). The benefit of collecting data at intervals that coincide with the signal cycle length is that flows over a detector close to the intersection will tend to fluctuate from high to low within the cycle, but are relatively stable from cycle to cycle. For example, queues will build and spillback over the detectors during the red period on each approach (resulting in higher occupancy) and then the queue will dissipate and (hopefully) clear during the green period (resulting in lower occupancy). If a controller had a 60-second cycle length, and data was collected every 30 seconds, then the occupancy could fluctuate substantially from one sample (during the red portion of the cycle for a given approach/detector) to the next sample (during the green portion of the cycle). This is illustrated in Figure 2, where at a signal with a

60-second cycle length, the last 30-seconds of occupancy ranges from 37% to 100% between green and red periods of the same cycle (a range of 63%), whereas samples of the last 60-seconds of occupancy ranges from 61% to 97% (a range of 36%).

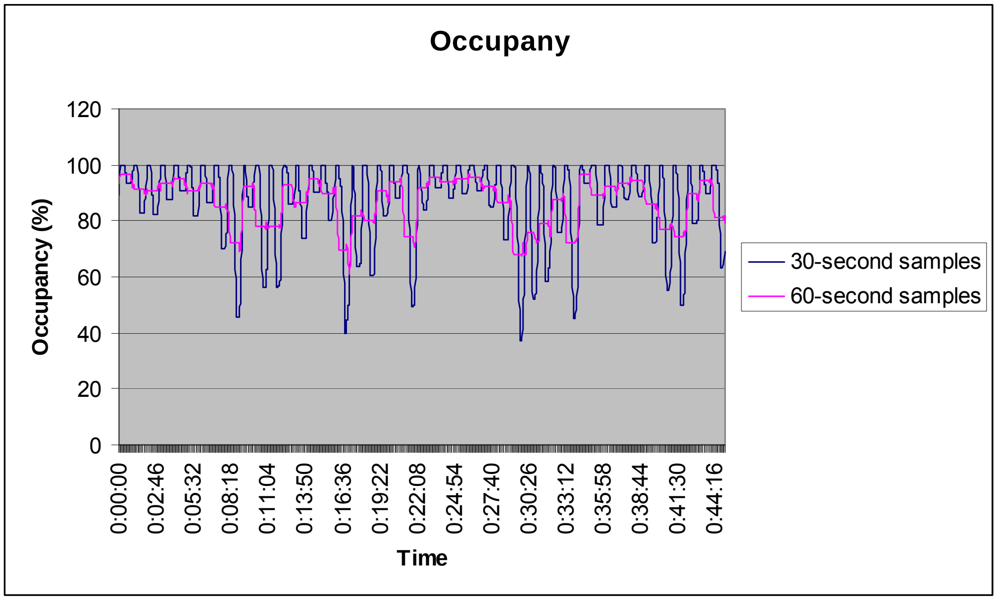

Figure 2. Variability of short occupancy samples.

It is important to note that even when the detector sample period (60 seconds) is equivalent to the cycle length (60 seconds), there is significant variability in the occupancy from cycle to cycle. It would not be appropriate to report congestion values based on a single cycle of data, nor based on smoothing just two cycles of data. In ACS Lite we require a minimum of three cycles or 5 minutes of data, whichever is greater. Table 2 indicates the variability in the occupancy when values are based on 1-, 2-, 3-, 5-, 10-, and 15- minutes (cycles) of data (where the cycle is 60 seconds). This is based on an hour of data simulated at steady/unchanging (mean) traffic volumes. The interpretation of this data is that even averaging over the last 5 minutes (or equivalently 5 cycles in this case), the results can vary significantly from one sample to the next, leading to an erratic indicator that might span three different congestion levels in three consecutive minutes. Thus, we suggest taking a moving average over the last 15 minutes of data to reduce the minute-to-minute variability of the indicator, as sampling at the same period as the signal cycle does not adequately reduce the variance of the congestion indicator. We would also suggest that if fewer than 50% detector poll responses have been successfully received in the last 15 minutes (for all detectors, which is likely since all detectors report in the same poll response) then display a “no data” indication.

Table 2. Variability of occupancy averaged over several periods—fixed-time control.

Minutes	1	2	3	5	10	15
Ave	85.1	85.0	84.9	84.7	84.3	84.5
Min	58.4	66.7	71.5	73.5	76.9	79.2
Max	96.4	95.4	94.9	93.9	89.9	87.9
Range	38.0	28.7	23.4	20.4	13.0	8.7
Variance	80.2	49.8	37.3	25.8	9.9	5.6
Std. Dev.	9.0	7.1	6.1	5.1	3.1	2.4
Coef.Var.	10.53%	8.30%	7.19%	5.99%	3.73%	2.81%

The data of Table 2 is perhaps easier to appreciate in visual form. Figure 3 shows how occupancy fluctuated for a detector when averaged over periods from 1 to 15 minutes. This is a detector on an approach served by a fixed-time split, with steady traffic, and a fixed 60-second cycle length. Averaging over the last 5 minutes still results in an erratic measure of congestion. We prefer to consider the last 10 to 15 minutes of data, though we are willing to publish congestion metrics if we only have data for 50% of the last 10 to 15 minutes.

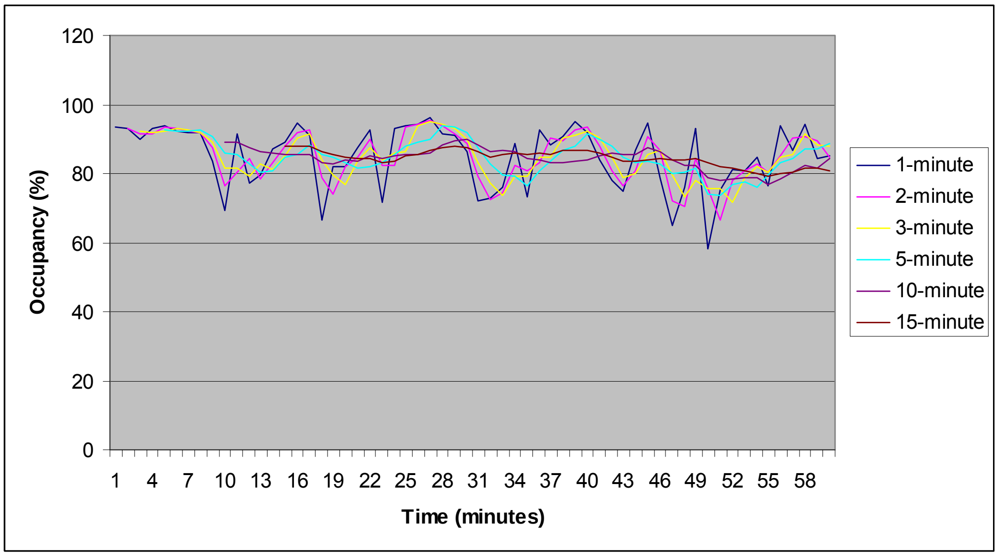

Figure 3. Variability of occupancy for several sample periods—fixed-time control.

Data for the previous tables and figure was derived from a CORSIM simulation, using ACS Lite to collect second-by-second detector data via a special run-time extension (RTE) interface to CORSIM that supports NTCIP messaging with ACS Lite. However, that RTE does not support controllers running free. All five controllers in the Glendale congestion monitoring network are currently running free (see Technical Memorandum #1). To obtain representative data for a controller running free, we used a hardware-in-the-loop simulation technique so that ACS Lite could poll detector data directly from a real Eagle M50 controller running free. However, our controller interface device (CID) was failing after 5-minutes during simulation runs. So, we developed a new, custom run-time extension for CORSIM to collect and output a file with second-by-second volume and occupancy data for detectors in CORSIM; and, it works while the controller is running coordinated *and* free. Figure 4 shows a chart of several occupancy periods (from 1 to 15 minutes) for the same traffic conditions as Figure 3 except that the controller is now running free (i.e., has no fixed cycle length). Under free operation, there is more variability in the actual capacity allocated to a given approach, as that capacity varies from cycle to cycle (unlike fixed-time control).

Comment [LL1]: Could you get some historical data/simulated data as volume and occupancy input for the simulation? It will be helpful to illustrate the performance of the algorithm with the 'real' data.
Yes, please provide your CORSIM FILES.

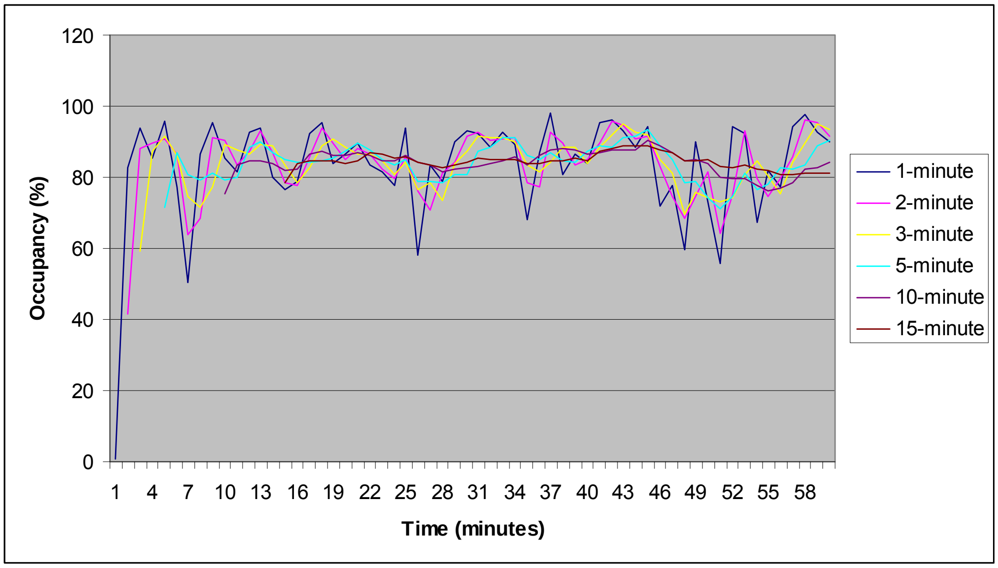

Figure 4. Variability of occupancy for several sample periods—free operation.

Table 3 shows the variance of each of the aggregation periods. Averaging over the last 15 minutes seems to provide measures under free operation that are as "stable" as measures under fixed-time control.

Table 3. Variability of occupancy averaged over several periods—free operation.

Minutes	1	2	3	5	10	15
Ave	83.2	83.9	84.0	84.0	84.1	84.4
Min	0.8	41.7	59.1	71.1	75.4	78.6
Max	98.2	96.0	94.9	93.5	90.3	89.0
Range	97.4	54.3	35.8	22.3	15.0	10.4
Variance	232.6	94.5	52.0	26.8	11.2	5.1
Std. Dev.	15.3	9.7	7.2	5.2	3.3	2.3
Coef.Var.	18.32%	11.59%	8.58%	6.16%	3.98%	2.68%

To summarize some of the discussion so far, our testing suggests that to obtain accurate and "stable" measures of occupancy, it appears necessary to average over the last 5 to 15 minutes just to smooth out the randomness of traffic flow from cycle to cycle. (This is randomness about the steady mean rate, not randomness due to truly changing conditions in the average rate of traffic arrivals.) When running under free operation (no fixed cycle), the data tends to be more variable from cycle to cycle, but smoothing over 10 to 15 minutes provide measures of occupancy that appear to be as stable as the same measures under fixed-time control.

These experiments suggest that it is not necessarily critical to collect occupancy data over periods that match the cycle length (when there is a known, fixed cycle length), because data must be averaged over longer periods (several cycles) to contend with normal cycle-to-cycle variations in traffic. It is also evident that in collecting over longer periods (10 to 15 minutes) that good measures can be obtained for uncoordinated signals. This suggests that it would be adequate to poll controllers for detector data at fixed intervals, such as every 30 or 60 seconds. That would eliminate the complexity of changing the polling period from one pattern next to match the cycle length of the active pattern, which does not seem to drastically improve performance. Furthermore, protocols such as AB3418 allow fixed-period detector polling, but do not provide a dynamic means to change the polling period through a simple message. A larger block of data would have to be downloaded at every pattern change to change the polling period. Polling at fixed intervals would significant reduce the costs to modify the central system, and to configure the system and controllers correctly.

PROCESS DETECTOR DATA

Goals

The goals in processing detector data are as follows:

- Obtain reasonably accurate and stable/reliable measures of congestion.
- Develop easy to configure, calibrate, and interpret congestion measures to keep costs of field visits and user feedback (i.e., complaints) to a minimum.
- Keep down costs of software development for modifications/enhancements.
- Retain user configuration flexibility.
- Use techniques that are applicable to a broad range of controller hardware/firmware.

Discussion

Bellevue applied user-defined thresholds as shown previously in Table 1 to map occupancy levels to congestion levels. There appear to be pros and cons to this approach. We propose extending this to be more flexible by calculating a congestion measure based on a weighted combination of detector occupancy and volume, as illustrated in the following equation.

$M_i = W_O \times O_i + W_V \times V_i$ **[1]**

Where:

M_i = A congestion measure for detector *i*.

O_i = The aggregated (fault-free) occupancy of detector *i*.

V_i = The aggregated (fault-free) volume of detector *i*.

W_O = A weighting factor for the occupancy.

W_V = A weighting factor for the volume.

The occupancy and volume values would be based on the last several samples of data, assuming that at least the minimum number of samples was obtained within the maximum number of sample periods considered. For example, the minimum could be set to 8 and the maximum could be set to 15 samples, where each sample is 60 seconds in duration. In addition to having enough data, it is assumed that the number of volume samples and occupancy samples are equivalent (it would be unusual to have one and not the other, but negligent to use occupancy from one sample and volumes from another). Occupancy values in most protocols are coded to indicate faults identified on the controller side. Both volume and occupancy values collected during a fault sample would not be utilized. If a fault occurred, there may have been erroneous data leading up to that fault identification. For example, the detector is typically stuck on, stuck off, or counting excessively for sometime prior to the controller reporting a diagnostic fault. Thus, we would use only samples occurring after the fault towards the requirement to meet the minimum number of samples occurred after a fault condition cleared. The system would not be purely "event" driven by the receipt of new data. It would also track time, such that as time moves forward and no samples data is obtained for one or more sample periods (due to communications failures), a communication fault would be identified. We would not continue reporting the last known data once we have reached a point where there are not enough current samples (say 8 of the last 15 minutes) to substantiate a congestion indication, and thus a "not available" indication would be posted until there is adequate data again.

Comment [LJN2]: You're only concerned about skipping data before detector faults, correct? I don't understand the significance of mentioning communication faults. This simply limits the number of samples that you have within a period.

Assuming there was adequate fault-free data, calculate measure M_i and then use user configured thresholds to calculate the congestion index, generally as follows.

Congestion	Condition				
No data	Not enough good data (but no faults).				
Fault	Detector faults detected.				
Low	0	≤	M_i	<	L_{max}
Medium	L_{max}	≤	M_i	<	M_{max}
High	M_{max}	≤	M_i	<	H_{max}
Severe	H_{max}	≤	M_i	≤	S_{max}
Fault	S_{max}	<	M_i	Out of realistic range values.	

Table 1 gave an example of setting thresholds for occupancy data only. Users could set the occupancy weight to 1.0 and the volume weight to 0 to utilize that scheme. In situations where users want to apply existing upstream system detectors, they might use volume only as a measure of congestion, and set the volume weight to 1.0 and the occupancy weight to 0; or, the traffic engineer might prefer to use a tradition V+kO measure, though these measures are more difficult to accurately calibrate [2] In using stopline detectors, we suggest considering a weighted combination of occupancy and volume, inspired by ACS Lite and SCATS. The basic concept in ACS Lite is similar to a gap/extension timer, whereby if a detector becomes unoccupied, the signal holds the phase green for a short-time gap/extension time until the next car arrives. During saturation flow, headways are typically about 1.8 to 2.2 seconds, with some variability, so a gap/extension timer might be set conservatively for about 3.6 to 4.4 seconds. This accounts for the expected “unused” space between vehicles in saturation flow, traveling over short detectors. Figure 5 illustrates raw occupancy (1 = occupied and 0 = unoccupied) of a detector placed 125 feet from the stopline. As the queue clears, there are gaps between the vehicles which correspond to gaps in the purple presence timeline below. The maroon-colored areas illustrate the effect of a gap/extension timer, filling in the time of a typical headway between vehicles in saturation flow (such as a dispersing queue). Where the maroon extension does not completely fill in the gap in (purple) detector occupancy, there are exposed gray background areas in the chart, and this represents larger gaps where “stragglers” are arriving after the queue has already dispersed, sometimes sparsely and sometimes in multi-car clusters/platoons. ACS Lite operates by filling in gaps in the occupancy timeline to compensate for expected gaps in saturated flow, which might be larger or smaller depending on how long the detectors are. This allows a more precise read on traffic demand during green. We refer to this gap-filled occupancy as “utilized occupancy”. To collect this sort of data would require much more sophisticated polling and processing to match up effective green time with detector presence and fill in the gaps.

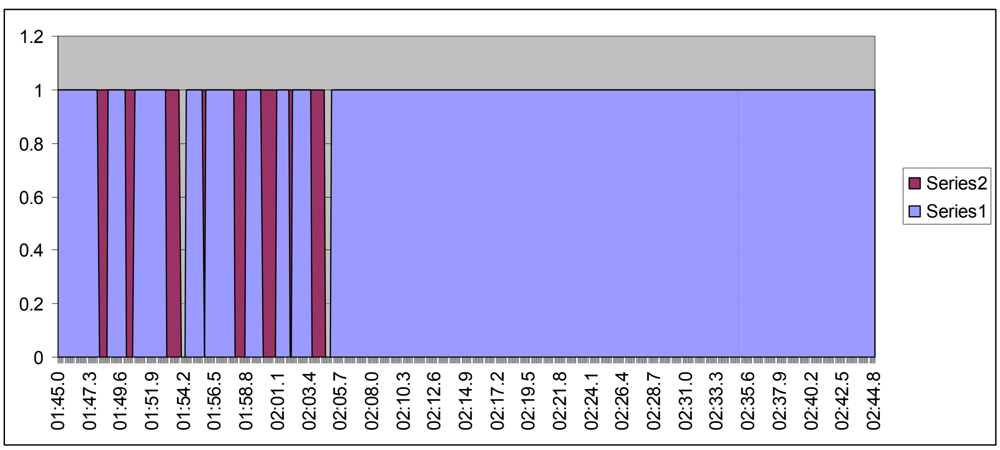

Figure 5. Illustration of raw presence (purple) and normal trailing gap-space (maroon).

---

[2] The *FCTrip* map in Fort Collins, Colorado uses this approach. The local traffic engineer reports areas where the V+kO technique seems to be fairly inaccurate. (http://fcgov.com/fctrip/)

SCATS is slightly simpler than ACS Lite by simply considering only a single aggregate value of volume and occupancy over the whole green period. In the chart above, the detector was occupied 72.9% of the green time, and there were 6 gaps during green (a volume of 6). It can be calculated that for a 6-foot loop, the average occupancy time of a passing vehicle at free flow speed is 0.4 seconds, so if the average headway between vehicles is assumed to be 1.8 seconds, then the average expected gap could be assumed to be 1.4 seconds (1.8 seconds between vehicles minus 0.4 seconds while crossing the detector). SCATS would add 6 average gap times to the total occupied green time (6 gaps X 1.4 seconds/gap = 8.4 seconds) to yield an adjusted occupancy during green of 122% (yes, the values can be over 100%, which SCATS sales literature states is because they have an ability to measure oversaturation).

We do not propose that it would be a low-cost modification to the central system to collected high resolution volume and occupancy and phase status and do the processing to discern green occupancy and green volume as SCATS does. However, it is possible and effective (for the less-precise-than-adaptive-control purposes of discerning low, medium, and high congestion) to use full-cycle (green, yellow, or red) occupancy and volume in this manner … inflating the recorded occupancy over the last 15 minutes by an "average headway gap" for each gap (the volume) counted during that same period. To test this hypothesis, we simulated a fixed-time intersection with a 60-second cycle length, with four approaches operating at 100%, 90%, 80%, and 70% saturation levels (or volume-to-capacity levels). Some findings (for a single simulation run) are shown in Table 4.

Table 4. Comparison of congestion measures for fixed-time control.

V/C Ratio	Phase Failures	Subjective Congestion	Stopline Detector		Stopline Adjustments		O+g(V) Indicator	Advance Detector		Advance Adjustments		Bellevue
			Occ.	Vol.	Sec/Gap	Adj. Occ.		Occ.	Vol.	Sec/Gap	Adj. Occ.	Indicator
1.00	68	High	90.31	330	1.17	101.03	High	60.11	488	1.41	79.20	Moderate
0.90	28	Medium	85.62	298	1.17	95.30	Medium	25.76	477	1.41	44.41	Light
0.80	18	Low	81.66	259	1.17	90.08	Low	14.39	434	1.41	31.36	Light
0.70	8	Low	79.88	229	1.17	87.31	Low	10.59	396	1.41	26.07	Light

The data in Table 4 is based on a 1-hour simulation, with 60-second cycle lengths, and thus 60 cycles of the signal; however, each approach has both a through phase and a left-turn phase, which both contribute to the phase failure count. The congestion level was determined subjectively (in the 3rd column) by considering the phase failure count for each approach and by observing the simulation animation. The first approach (V/C = 1.0) had a either a left-turn or through-phase failure during most cycles and was certainly in high congestion; however, the queues were not growing (it was not over-saturated or in severe congestion). With 20-foot detectors at the stopline, the adjusted occupancy (occupancy + gaps X average secs/gap) indicator tends to top out around 100 (percent occupancy) suggesting high congestion. However, a stop-line detector is not in our opinion well-suited to detector over-saturation (despite SCATS marketing to the contrary). Table 5 provides thresholds that might be used to delineate congestion levels based on different performance measures.

Comment [LL3]: The formula is a little over simplified. (3600*90.31%+330*1.17)/3600=1.0103

Table 5. Congestion thresholds for different measures.

Congestion Level	V/C Ratio	Saturation Level	Stopline Adj.Occ.	Bellevue Adv.Occ.
Severe	> 1.0	> 100%	-	> 78%
High	> 0.9	> 90%	> 100%	> 68%
Medium	> 0.75	> 75%	> 95%	> 45%
Low	< 0.75	< 75%	< 95%	< 45%

Returning to Table 4, note that in using a 6-foot advance loop located 125 feet from the stopline, the average occupancy over the hour indicated only moderate congestion. We believe the benefit of using the stopline detectors and the adjusted occupancy technique (O+g(V)) provide a much easier to understand metric that probably (we speculate) will not require nearly the calibration/tuning effort of using occupancy only at an advance loop. The occupancy experienced at an advance loop would seem subject to the queue lengths, which in turn are a function of the cycle length. Thus, the stopline detectors seem to give a more consistent read that is insensitive to cycle length changes throughout the day and might be less prone to users calling in to suggest changes. We believe that the stopline detectors with the adjusted occupancy were more accurate at all measured congestion levels in this experiment; however, we note that advance detectors are capable of revealing severe congestion (i.e., oversaturation) which is not possible to distinguish (from high congestion) with stopline detectors.

Figure 1 charts the adjusted occupancy from two stopline detectors for an approach operating just at 100% saturation, as well as occupancy-only measures from two advance detectors. Using the congestion level threshold values for these measures suggested in Table 5, it is evident that the stopline detectors measure high congestion until about 45 minutes into the simulation, when values drop into the medium congestion level. However, the advance detectors vary quite widely in their assessment, bouncing from high to severe to high to medium to low. Contrast these assessments to the chart of phase failures for this approach shown in Figure 7, which shows a relatively steady rate of phase failures throughout the simulation.

Comment [LL4]: What is the scale and X,Y?

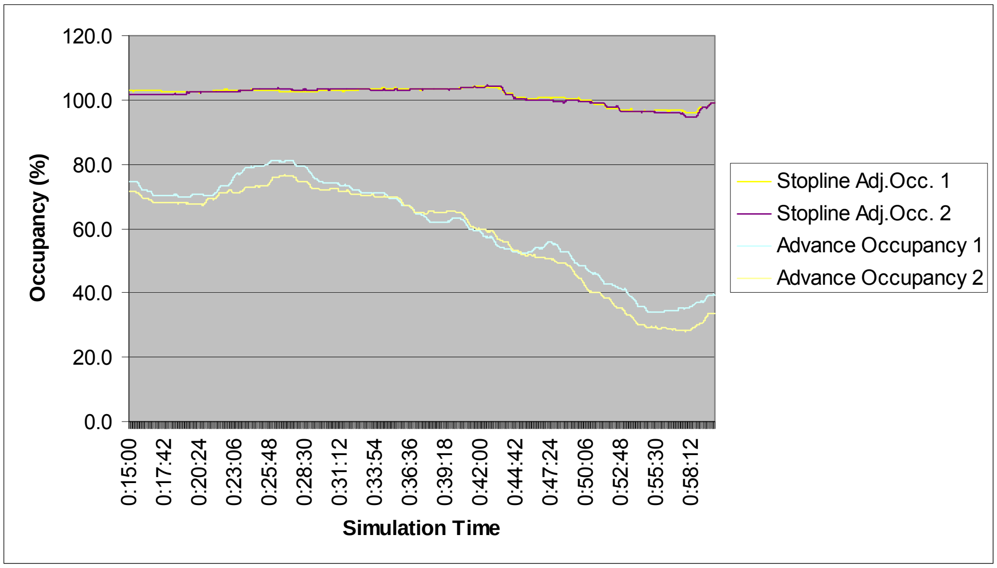

Figure 6. Comparison of detectors measures.

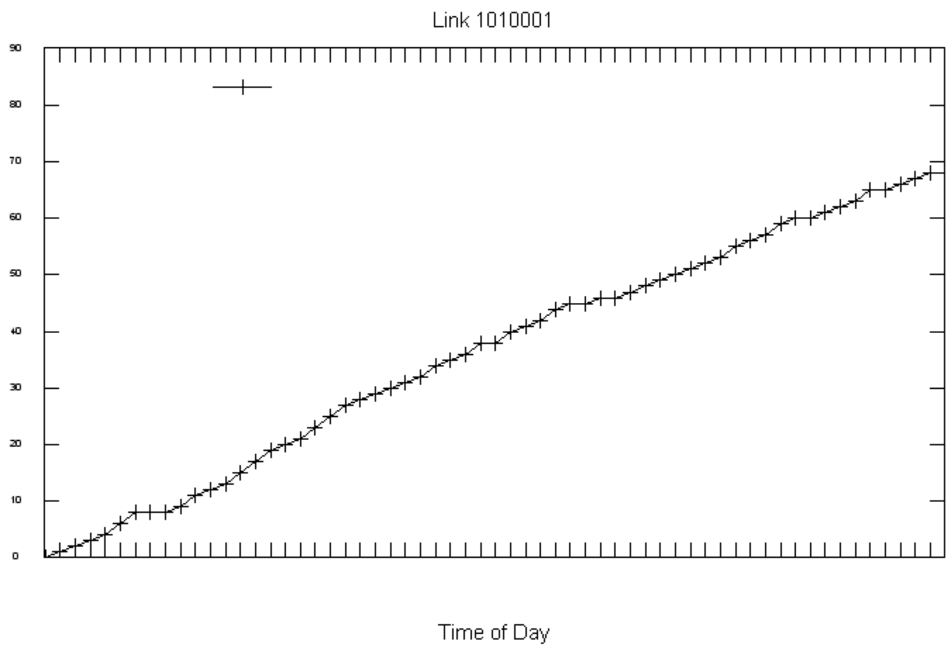

Figure 7. Phase failures on 100% saturated approach over 1-hour simulation.

In summary, we feel that adjusted occupancy values from stopline detectors could give a better read on data, and appear to be less prone to any need for calibration/tweaking of thresholds than possible using occupancy-only at advance detector locations. In particular, it would seem that occupancy thresholds for advance detectors are sensitive to the prevailing cycle length, and the 60-second cycle studied here is shorter than the typical cycle lengths used to establish default thresholds in Bellevue. We have also experimented with 90-second cycles and free operation. Both approaches are certainly workable, and both approaches have their pros and cons. For example, the stopline detectors are not able to pickup severe saturation, whereas advance detectors can. Using both detectors on the same approach would be preferable, where stopline detectors would pickup a more accurate saturated or undersaturated congestion levels, and advance detectors would indicate when oversaturation appears. This would seem particularly relevant to the Glendale stadium, which is likely subject to severe saturation surrounding major sporting events.

Comment [LL5]: The method and measurement by combining stopline and advance detectors for oversaturated condition was not spell out. Please add.

To calculate the weighting or "correction" factor used for our adjusted occupancy values above, we calculate the time an average length vehicle (say 17 feet) would take to travel over the configured detector length (20 feet in this case) while traveling at the free flow speed (or speed limit). In this case, the free flow speed was 40 miles/hour, or 58 feet/sec, and thus a vehicle traveling at that speed would actuate the detector for 0.6 seconds while traversing it (where the front bumper travels 20+17 = 37 feet at 58 feet/sec). Assuming an average headway during saturation flow of 1.8 seconds from front bumper of a leading vehicle to front bumper of the following vehicle, we subtract the detector traversal time and yield an average expected gap in occupancy of 1.2 seconds. Thus, the correction factor is a straightforward calculation from the known detector information specified in the configuration, and this weighting factor does not require excessive tweaking or tuning to "find". This correction factor is what we would propose to calculate for an "automatic occupancy adjustment" feature, which would keep configuration simplistic—a benefit for large-scale deployments. This technique scales easily to different detector loop lengths (which are not so easy to modify after-installation) and may eliminate the need for field visits and trial and error tuning. Multi-lane detectors can be handled as well, though we would expect degraded performance and some need for trial and error adjustment of weights and thresholds. The "automatic adjustment" feature (checkbox) could be disabled and the user would be free to manually configure weights as desired. We do not recommend multi-lane detection if it can be avoided.

The charts and tables used in this section were intended to convey the technique, though a greater appreciation comes from viewing the simulation animation and comparing charts of the recorded measures and corresponding congestion indicators with the animation. These simulations are available for review if desired.

UPDATE CONGESTION DISPLAY

Congestion measures for each link or link-movement will be updated periodically as new sample data becomes available (or it becomes evident that communications have failed). The congestion map can be updated periodically as well, such as once-per minute. It does not seem overly crucial for the update period of the online display mechanism to match the update period of the detector data collection.

Figure 8 shows a screenshot of Bellevue's online congestion map. Bellevue has also experimented with utilizing Google maps for the underlying map display. A proposed congestion display will be the subject of the next task and Technical Memorandum #3.

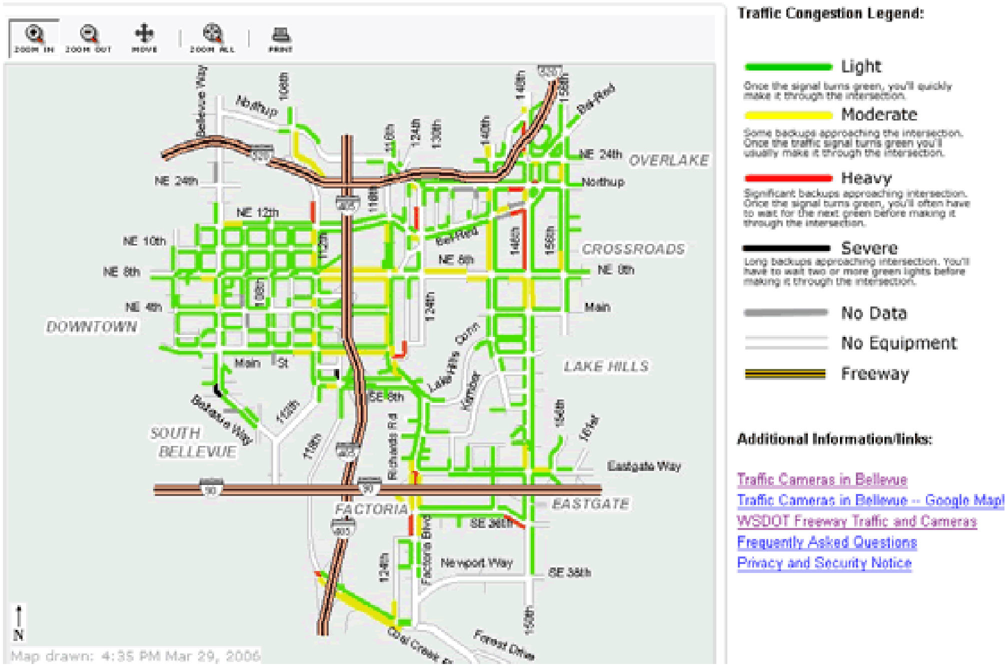

Figure 8. Bellevue's congestion map (source: Fred Liang [1]).

Congestion Level Definition

The City of Bellevue did mention receiving user feedback to suggest fine-tuning of their thresholds. It would seem that clear definition of what is intended by different color indications might alleviate a number of questions/suggestions about the system, as each user may have their own subjective notion of low-, medium-, and high- congestion levels.

Here are the definitions used by Bellevue:

1. **Light** Traffic: Once the traffic signal turns green, you'll quickly make it through the intersection.

2. **Moderate** Traffic: Expect some backups approaching the intersection. Once the traffic signal turns green, you'll usually make it through the intersection.

3. **Heavy** Traffic: Expect significant backups approaching the intersection. Once the traffic signal turns green, you'll often have to wait for the next green light before making it through the intersection.

4. **Severe** Traffic: Expect long backups approaching the intersection. You'll have to wait for two or more green lights before making it through the intersection.

When using advanced detection and occupancy-only measures, the different congestion levels become a little more blurred from signal to signal or from hour-to-hour as advance measures are more sensitive to the current cycle and split lengths. Using stopline detectors instead of advance detectors, we expect to be able to distinguish the first three states slightly more accurately and consistently, and offer the following definitions. These congestion states are defined in terms of the expected number of cycles or green displays that may be required for a vehicle to pass through the traffic signal at the downstream end of the link, as follows:

- **GREEN:** No congestion (or low congestion) will be indicated by the color green when it is expected that all vehicles will get through the traffic signal at the downstream end of the link on the first display of green.
- **YELLOW:** Borderline congestion (or medium congestion) will be indicated by the color yellow (or perhaps "amber", which is slightly darker and tends to provide more contrast and a more easily discernable display on a computer screen). This condition is indicated when it is expected that vehicles arriving to an already green signal will sometimes have to wait for a second display of green before proceeding through the intersection at the downstream end of the link. This is perhaps more easily thought of as in-between "always getting through on the first green" and "the last vehicle(s) never get through on the first green".
- **RED:** High congestion will be indicated by the color red. Under this condition, it is expected that the last vehicle(s) arriving during green every cycle of the signal will encounter queued traffic preventing passage through that first display of green and can expect to wait for a second display of green before proceeding through the intersection.
- **MAGENTA:** Severe congestion will be indicated by the color magenta (which stands out on a link-display beyond the color red, and is less subject to being disregarded than the color black). Under this condition, it is expected that at least one or more of the vehicles already in the queue at the onset of green will fail to clear the intersection each cycle and can expect to wait for two or more displays of green before proceeding through the intersection. Thus, every vehicle can expect to be stopped by the signal at least once.
- It is also expected that for a state of failed communications, inadequate data, or failed detectors, the links might be displayed in another color, such as **GRAY**.

Comment [LJN6]: How does this relate to the 1,2,3 levels indicated earlier?

A useful term to distinguish a basic level of service is a *phase failure*. A *phase failure* is defined as a scenario where vehicles that are in the queue when the signal turns green do not make it through the intersection during that first display of green. If all vehicles queued at the onset of green are able to make it through the light, then it is not a phase failure, even if one or more vehicles joining the queue just after the onset of green do not make it through. Lacking any existing terminology for this secondary condition, where it is only the arrivals after the beginning of green that fail to clear the signal, we might refer to this as a *flush failure* (indicating that the full queue did not completely flush/clear, even if cars in the *green* *onset queue* did clear the stop line). The red/high-congestion state indicates consistent *flush failures*, but not necessarily consistent *phase failures*, which are more problematic. Where flush failures indicate just being at saturation, consistent phase failures would indicate oversaturation/severe saturation.

Only with advance detection can “severe” saturation be reliably detected. For example, if the adjusted occupancy of an advance detector was 100%, it means that the queue is never shorter than the length between the stopline and the detector, and thus the approach is oversaturated.