Probability of Overload at Signalized Intersections

STAN TEPLY

Quality of service at transportation facilities can be determined from a number of objectively measurable criteria. For signalized intersections, delay has been accepted internationally as the prime measure of performance and, in the United States, as the sole basis for the level-of-service measures. In other countries, several other parameters or their combinations are also used, usually as additional measures related to specific design or evaluation objectives. A pilot project at the University of Alberta, which investigated the potential of a supplementary criterion—the concept of the load factor used in the 1965 Highway Capacity Manual, are used as indications of which probability type would provide the most practical representation of measurable parameters. The results of the pilot project suggest that the probability of an overload in one, two, or both of two consecutive cycles is a strong candidate to approximate the percentage of cycles with discharge overloads. It is mathematically simple and, because it resembles the overload factor, easy to measure.

The quality of service offered by a transportation facility manifests itself to the users and to the designer or analyst in a number of ways. The 1985 Highway Capacity Manual (HCM) (1) employs such criteria as speed, traffic density, percentage time delay, reserve capacity, and so on, depending on the type of facility. The measure used for signalized intersections is the average stopped vehicular delay, which is applied as the sole characteristic of levels of service. In many specific cases, however, the traffic engineer must also analyze other operational features that relate more closely to the problem at hand, such as the capacity, length of the queue in important lanes, probability that an approach or a lane will be oversaturated, delay to passengers, number of stops, emissions, or cost.

The 1965 HCM (2) defined load factor as the measure underlying the level of service. The 1984 Canadian Capacity Guide for Signalized Intersections (CCG) (3) does not specify levels of service but relates the ranges of average overall delay to the quality of operation. It also recommends additional criteria, such as the probability of clearance and queueing. The second edition will include a number of other measures of effectiveness (4). A 1983 Australian Road Research Board report (5) employs overflow queue, delay, number of stops, and queue length.

ENGINEERING EVALUATION AND QUALITY OF SERVICE

An examination of the definitions of various criteria used for signal design and analysis indicates that there are two different, but related, reasons for specific measures describing intersection performance:

- Relating the quality of operation to the general public and politicians,
- Identifying specific features of the operation with the intent of revealing problems, their causes, and contributory factors.

The first reason is clearly related to perceptions. The level-of-service concept represents an attempt to describe the ranges of satisfaction provided by the facility. Traffic engineers apply categories of an objective measure to define different levels of service. The objective measure (e.g., delay in the 1985 HCM) has been selected on the basis of its perceptivity by the user of the facility or a decision maker.

The second reason may employ a whole range of operational characteristics. If, for instance, a frequent lane blockage by a spillover queue has been the problem, the analyst should be in the position to determine both the length of the queue and the frequency of this undesirable event. Therefore, it should be possible to both measure and calculate these operational features. Naturally, the computational process may involve analytical formulas, iterative procedures, or simulation.

The focus of this paper is on cycle overload as a specific engineering operational criterion. It discusses its forms and past formulations and presents potential analytical and measurement methods to make its use practical. The intention is not to substitute delay as a measure underlying levels of service, but merely to provide the analyst with an additional tool.

LOAD FACTOR

The load factor was used as an intersection performance criterion to determine the level of service in the 1965 HCM. It was classified as a part of environmental conditions, that is, conditions that cannot be readily changed by alteration of design or control features of the intersection.
This factor was defined very specifically as follows:

The load factor is a measure of the degree of utilization of an intersection approach roadway during one hour of peak traffic flow. It is the ratio of the number of green phases that are loaded, or fully utilized, by traffic (usually during the peak hour) to the total number of green phases available for that approach during the same period. As such it is also a measure of the level of service on the approach, as discussed later. The load factor for a normal intersection may range from a value of 0.0 to a value of 1.0. (2)

The 1965 HCM defined a loaded cycle as a cycle in which there are vehicles ready to enter the intersection when the signal turns green and in which vehicles continue to be available to enter during the entire phase with no unused time or exceedingly long spacings between vehicles at any time because of lack of traffic. It was pointed out that the ending of a loaded green phase will usually (i.e., not necessarily) force some vehicles to stop, and that any stoppages that occur must be caused by conditions at the intersection under study, not by conditions elsewhere. This description therefore also implies that the just-at-capacity cycles were included (i.e., more than just overloaded cycles were considered).

A number of research papers analyzed load factor applications and their relation to delay using simulation (6), measurements, or analytical methods (7–9).

Experience with the 1965 HCM indicated that the major drawbacks of the load factor application were as follows:

- No computational procedures for design situations were available in the 1965 HCM. Miller (10), in 1968, proposed a formula derived from May and Pratt’s simulation (6) as follows:

\[
LF = e^{-1.36}
\]

where

\[
\phi = (1 - x) \sqrt{sg}/x
\]

and

\[
LF = \text{load factor,}
\]

\[
x = \text{degree of saturation (i.e. } q_c/sg),
\]

\[
g = \text{saturation flow,}
\]

\[
c = \text{cycle time}
\]

- The load factor determination did not include its variability even under identical traffic conditions. For instance, measurement of the load factor during the same periods of similar days with the same volumes and signal timing will bring about a whole range of load factors. A single measurement is only a part of an unknown distribution and, in itself, may not tell much about the average or the deviations.

- No recognition was given of temporary, very short traffic flow peaks or of the fact that the number of overloaded cycles is time-dependent. The chances that a cycle following an overloaded cycle will also be overloaded are higher than for a cycle following a cycle that left no overflow queue. This is recognized in the inclusion of the evaluation time element in some random overflow delay formulas (3, 5).

- The loaded cycles are difficult to identify because no visible demonstration of a zero overflow queue occurs.

**PROBABILITY OF ARRIVAL OVERLOAD**

There were other attempts to overcome the problems of load factor calculations. The most obvious candidate for a similar surrogate measure was the relationship between the arrival distribution and cycle capacity.

This measure is defined as the probability that the number of vehicles arriving during a cycle exceeds the average cycle capacity. It is based on the counting distribution of vehicle arrivals during individual cycles. Under the assumptions of a fixed-time signal operation, a steady mean of the arrival distribution, and the usual conditions at signalized intersections, the arrival fluctuations can be approximated by the Poisson distribution (3, 11–13).

The Poisson distribution also applies to approach lanes at intersections within coordinated signal systems. Although the internal arrival patterns during individual portions of the cycle depend on the quality of the progression (and may follow different counting distributions), the cycle-based arrival distribution can be expressed by the Poisson distribution in most cases. This distribution is well known to practitioners because it has been widely used as a basis for allocating green intervals or evaluation (3, 11, 12).

The probability of an overload in a cycle (i.e., not a loaded cycle in terms of the 1965 HCM), provided that the preceding cycle was not overloaded, equals the probability that the number of vehicles arriving during that cycle exceeds its capacity:

\[
P(x > X) = 1 - \sum e^{-m/i!} \quad \text{for } i = 0 \text{ to } i = X
\]

where

\[
x = \text{number of arrivals in a cycle},
\]

\[
X = \text{cycle capacity},
\]

\[
e = \text{basis of natural logarithms},
\]

\[
m = \text{average number of arrivals per cycle, and}
\]

\[
i = \text{counter}.
\]

The number of arrivals is normally expressed in vehicles (veh) per cycle, although passenger car units per cycle can also usually be used without any loss of representative power.

The formula corresponds to percentage of cycle failures by Drew and Pinnell (13), which was also Poisson-based, relating the average number of arrivals per cycle to the probability of \( X + 1 \) or more arrivals per cycle.

As shown in Figure 1, this probability is not symmetrical (because of the probability of 0 arrivals). If the average number of arrivals equals cycle capacity, the probability of more vehicles arriving than capacity is still less than 0.5. For example, for \( X = 6.0 \) veh/cycle, and \( m = 6.0 \) veh/cycle,

\[
P_{\text{arrival overload}} = P(x > X) = P(x > 6) = 0.394
\]

In instances for which capacity well exceeds the average number of arrivals, this simple distribution can be used as an approximation of the number of cycles that will be overloaded. Nevertheless, as the average number of arrivals increases (and capacity remains constant) the number of over-
Probability of Overload

![Probability of Overload diagram](image)

FIGURE 1 Poisson distribution for probability of number of arrivals exceeding capacity of a cycle (probability of arrival overload) for cycle time as time basis for mean of arrival distribution \((m)\).

Loaded cycles may exceed this probability measure by a wide margin. For example, in the preceding case of an average number of arrivals equal to six and capacity of six, the queueing system is unstable. Depending on the starting conditions (the presence of an initial queue), the probability that a cycle will be overloaded may reach 100 percent despite the fact that the arrivals at the end of the queue still follow the Poisson distribution. Naturally, if the average number of arrivals is greater than the cycle capacity, the queue will be consistently growing.

The computation becomes meaningless because all cycles will be overloaded.

For this reason, it is advisable not to confuse the arrival distribution with the probability that a cycle will be overloaded. It is, however, still a useful simple measure. It has been suggested during the review process of the CCG to call it probability of arrival overload, representing the frequency with which more vehicles than the cycle capacity may arrive. Keeping in mind that since in a random process the number of vehicles arriving during one cycle does not depend on the number of arrivals in the previous cycle, the probability of arrival overload in any cycle is independent of the probability of arrival overload in previous cycles. As a result, probability of arrival overload does not indicate the likelihood that the sum of vehicles waiting and arriving exceeds cycle capacity.

PROBABILITIES OF DISCHARGE OVERLOAD

Miller also suggested a formula for the proportion of phases in which the queue is cleared \((10)\). For consistency, a complementary probability term will be used in this paper as follows:

\[
P_{\text{cycle overload}} = (1 - P_{o}) = e^{-1.56\phi} \tag{4}
\]

where

\[
\phi = (1 - x) \sqrt{sg}/x
\]

and

\[
x = \text{degree of saturation (i.e., } qc/sg),
q = \text{arrival flow},
s = \text{saturation flow},
g = \text{effective green interval}, \text{ and}
c = \text{cycle time}.
\]

Messer and Fambro \((14)\) have shown a comparison of Poisson probability of arrival overload, Miller’s load factor, and Miller’s probability of cycle overload.

Obviously, because of the number of cycles considered in the numerator

- Load factor: \((x_{\text{arriving}} + x_{\text{waiting}}) \geq X\) (where \(X\) denotes cycle capacity),
- Probability of cycle overload: \((x_{\text{arriving}} + x_{\text{waiting}}) > X\), and
- Probability of arrival overload: \((x_{\text{arriving}}) > X\), the following relationship must exist:

\[
\text{LF} > P_{\text{cycle overload}} > P_{\text{arrival overload}}
\]

Equations proposed originally for the second edition of the CCG also featured an empirical negative exponential form \((4)\) but included a calibration parameter for the series of the means of arrival distributions as follows:

\[
P_{\text{cycle overload}} = 100e^{-\{X-(m-1)\}/a} \% \tag{5}
\]

where

\[
X = \text{cycle capacity},
m = \text{means of arrival distribution}, \text{ and}
a = \text{calibration parameter}.
\]

The problem of sequential overloads merits closer inspection. Figure 2 illustrates possible cycle overload situations in a sequence of probabilities for an average of two arrivals per cycle and a cycle capacity also equal to two. The fast growing extent of this somewhat extreme case of a probability tree is the reason that such small numbers were chosen.
Given that there was no leftover queue in the previous cycle, the probability that the first cycle will be overloaded is equal to the probability of arrival overload, that is, to the probability that more vehicles arrived than the cycle can discharge.

Because the number of arrivals in individual cycles can be treated as independent events, the probability that the capacity of the second cycle will be exceeded is the sum of two probabilities as follows:

1. If fewer or equal numbers of vehicles than capacity arrived during the first cycle, the probability that more vehicles than capacity arrive in the second cycle,
2. If more vehicles than capacity arrived during the first cycle, the sum of the probabilities that the number of leftover vehicles together with the number of vehicles arriving exceeds capacity.

Because the overload in the second cycle depends on the outcome of the first cycle, events must be treated as dependent. The dependence is reflected in the change of the probability assigned to the dependent event.

Figure 2 shows that the resulting probability can be expressed as

\[
P_{\text{discharge overload in 2nd cycle}} = P(x \leq X)_{1st} \cdot P(x > X)_{2nd} + [P(x = X + 1)_{1st} \cdot P(x = X - 0)_{2nd} + P(x = X + 2)_{1st} \cdot P(x = X - 1)_{2nd} + P(x = X + 3)_{1st} \cdot P(x = X - 2)_{2nd} + P(x = X + 4)_{1st} \cdot P(x = X - 3)_{2nd} + \text{and so on}]
\]

where \(x\) is the number of arrivals in a cycle and \(X\) is the cycle capacity.

The probability of a discharge overload in the first, second, or both cycles (i.e., at least one overload in these two cycles) can be determined as follows:

\[
P_{\text{discharge overload in at least 1 of 2 cycles}} = [P(x \leq X)_{1st} \cdot P(x > X)_{2nd}] + P(x > X) \cdot 1.0
\]
Or, in an easier way, examining the probability tree, as one minus the probability of no overload in either of the two cycles, that is

$$P_{\text{discharge overload in at least 1 of 2 cycles}} = 1 - [P(x \leq X)]^2 \quad (8)$$

Similarly, the probability of at least one overload in three consecutive cycles can be expressed as

$$P_{\text{discharge overload in at least 1 of 3 cycles}} = 1 - [P(x \leq X)]^3 \quad (9)$$

In an analogous way, one can also use the probability tree to determine the probability that two of two, two of three, or one of two cycles, and so on, will feature a discharge overload.

Figure 3 shows examples of some of such Poisson-based calculations as functions of capacity as an independent variable. Figure 3 shows (a) a simple Poisson distribution for the arrival of overload (equal to the probability of discharge overload in the first cycle); (b) distribution of discharge overloads in the second of two consecutive cycles; (c) distribution of discharge overloads in the first, second, or both (i.e., in at least one of two) consecutive cycles; (d) distribution of discharge overloads in both of two consecutive cycles; and (e) distribution of discharge overloads in at least one of three consecutive cycles. It can be seen that for the greater degrees of saturation reflected here in the greater differences between the number of arrivals per cycle and cycle capacity, the curves are close enough. Any of them could be used for practical considerations (perhaps with the exception of the overload in both cycles). This includes the curves representing arrival overload, shown as probabilities of discharge overload in the first cycle, that is, simple Poisson distributions. As capacity approaches the average number of arrivals (or vice versa in the real world), significant dissimilarities emerge.

A natural question is, What function would be a realistic representation of a true set of events during a longer period of time? Theoretically, many more than two or three consecutive cycles may be overloaded, or many combinations of overloads and number of cycles are possible. The probability of at least one overload in \( n \) cycles has been used as an example. Figure 4 shows that this probability converges when many cycles are considered. Note that with the increasing degree of saturation (i.e., \( m/X \)) the distributions become flatter and the mean simulated overload factor (OF) increases.

However, there appear to be two practical limitations to the number of cycles that should be used: (a) within a real set of events, there may be many cycles with no overload (as a result, the probability tree restarts itself); and (b) longer periods of consecutive overloads (say, 10 min or more, i.e., 5 to 10 cycles) usually result from a sudden surge in traffic demand (from a change of the arrival mean). As a consequence of consecutive overload, events are not random any more, and this condition may be hidden from the analyst, even when a peak hour factor (I) is considered.

An additional argument lies in the basic notion of the probability of independent events—probability defined as the number of cases favorable divided by the total number of cases. Whereas this principle applies within the different levels of the probability tree, it is not valid for the whole chain of events because the individual probabilities change with the levels (i.e., are dependent). Figure 4 illustrates this point. For the first few cycles the probability that there will be an overloaded cycle is represented by a definite number for any capacity. If, however, capacity equals or is less than the number of arrivals (i.e., \( X \leq 6 \)), the chances that a cycle within the series will be overloaded quickly approach certainty with increasing time (i.e., with the number of cycles considered).

The last deliberation also illuminates the difference between the 1965 HCM load factor and the probability of discharge, defined as the probabilities of overloads in the second or the third cycle, one overload in two cycles, two overloads in three cycles, or in any other way. The load factor was considered more as a sequence of independent events than a chain of dependent events. It was also frequently interpreted as the chance of an individual cycle being overloaded. This way of looking at the problem led quite naturally to its representation by the probability of arrival overload, that is, by a simple Poisson distribution.

**FIGURE 3** Probability distributions of overloads for mean number of arrivals in cycle \((m)\).
SURVEYS AND SIMULATION OF OVERLOAD FACTOR

The importance of knowing how many cycles are or may be overloaded justifies a search for an appropriate measure. To seek guidance as to which of the probabilities of discharge would represent a practical calculable measure, overload factors were studied by both simulation and surveys in a pilot project.

The overload factor was defined as the 1965 HCM load factor with one important exception: only those fully loaded cycles in which at least one vehicle was unable to discharge at the end were included [i.e., similar to Miller’s percentage of cycles in which the queue is not cleared (10)].

SURVEYS

In summer 1992 two intersection lanes in Edmonton were surveyed during three types of traffic conditions. The arrival rates for individual surveys were tested on the stability of the mean and were found satisfactory. The cycle capacity was determined directly as the discharge during fully loaded cycles. A total of 10 surveys included 292 signal cycles. Table 1 gives the results.

Simulation

Additional overload factors were obtained from a spreadsheet-based simulation program that used a Poisson distribu-

<table>
<thead>
<tr>
<th>Survey Number</th>
<th>Intersection</th>
<th>Period / No. of Cycles</th>
<th>Direction / Lane</th>
<th>Arrival Mean (per cycle)</th>
<th>Measured Capacity (per cycle)</th>
<th>Measured Overload Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>87 Avenue / 109 Street</td>
<td>PM / 21</td>
<td>EB / 2</td>
<td>10.8</td>
<td>12.9</td>
<td>0.286</td>
</tr>
<tr>
<td>2</td>
<td>87 Avenue / 109 Street</td>
<td>PM / 22</td>
<td>EB / 2</td>
<td>9.3</td>
<td>12.2</td>
<td>0.227</td>
</tr>
<tr>
<td>3</td>
<td>87 Avenue / 109 Street</td>
<td>PM / 26</td>
<td>EB / 2</td>
<td>9.2</td>
<td>12.6</td>
<td>0.154</td>
</tr>
<tr>
<td>4</td>
<td>72 Avenue / 114 Street</td>
<td>AM / 22</td>
<td>WB / 1</td>
<td>5.9</td>
<td>7.6</td>
<td>0.455</td>
</tr>
<tr>
<td>5</td>
<td>72 Avenue / 114 Street</td>
<td>PM / 27</td>
<td>WB / 1</td>
<td>5.1</td>
<td>7.7</td>
<td>0.111</td>
</tr>
<tr>
<td>6</td>
<td>72 Avenue / 114 Street</td>
<td>PM / 26</td>
<td>WB / 1</td>
<td>5.3</td>
<td>7.1</td>
<td>0.192</td>
</tr>
<tr>
<td>7</td>
<td>72 Avenue / 114 Street</td>
<td>PM / 36</td>
<td>WB / 1</td>
<td>5.8</td>
<td>6.8</td>
<td>0.222</td>
</tr>
<tr>
<td>8</td>
<td>72 Avenue / 114 Street</td>
<td>PM / 59</td>
<td>WB / 1</td>
<td>5.2</td>
<td>7.0</td>
<td>0.169</td>
</tr>
<tr>
<td>9</td>
<td>87 Avenue / 109 Street</td>
<td>PM / 27</td>
<td>EB / 2</td>
<td>8.7</td>
<td>12.2</td>
<td>0.074</td>
</tr>
<tr>
<td>10</td>
<td>72 Avenue / 114 Street</td>
<td>AM / 26</td>
<td>WB / 1</td>
<td>6.6</td>
<td>7.0</td>
<td>0.500</td>
</tr>
</tbody>
</table>

Numbers in () represent rounded values.
Percentage in Class

Survey 5
\[ m = 5.1 \quad X = 8 \]
\[ m/X = 0.66 \]
\[ OF = 0.12 \]

Survey 1
\[ m = 10.8 \quad X = 13 \]
\[ m/X = 0.84 \]
\[ OF = 0.32 \]

Survey 10
\[ m = 6.6 \quad X = 7 \]
\[ m/X = 0.95 \]
\[ OF = 0.61 \]

FIGURE 5  Probability distributions of overload factors simulated for arrival and capacity conditions from surveys \((m = \text{average number of arrivals per cycle and } X = \text{cycle capacity})\).
tion for vehicle arrivals. All surveyed sets of mean arrivals and capacities were included in the simulation process of this pilot study. Each set was simulated in 10 series of 20 computer runs representing 20 cycles each, thus incorporating $10 \times 10 \times 20 \times 20 = 40,000$ cycles and resulting in $10 \times 10 \times 20 \times 20 = 2,000$ overload factors.

Because of the differences in the surveyed conditions (mean arrivals and capacities) as well as the small number of surveys, it was not considered prudent to investigate the distributions of measured overload factors. On the other hand, simulation provided sufficient data to that end. Probability distributions of the simulated overload factors for three conditions are shown in Figure 5. They exhibit remarkable consistency for each of the three simulated conditions and document the suspected wide range of possible values for higher degrees of saturation. This observation can be interpreted as a greater risk of surprising system failures for closer to capacity conditions. The practical meaning is that on some days, under identical arrival and signal timing conditions as before, many more overloaded cycles may happen.

A related interesting observation was made for the three surveys for which the measured or simulated overload factor exceeded 0.3. 1965 HCM identified load factor over that value as approaching unstable flow. Surveys and many simulation runs showed a number of consecutive overloads, which, indeed, are indicative of unstable conditions.

COMPARISON OF OVERLOAD FACTORS WITH PROBABILITIES OF OVERLOAD

Figure 6 shows the comparison of simulated overload factors with the various forms of probability of discharge overload. The probability of arrival overload is included as the probability of discharge overload in the first cycle. It is encouraging to see the consistency of the patterns. Generally, the simulated overload factors fall between the probabilities of a discharge overload in the second cycle and a discharge overload in at least one of two cycles. Probability of at least one overload in two cycles would be a good approximation for all overload factors and offers a somewhat conservative estimate.

Figure 7 depicts the overload factor values generated by individual simulation series and identifies the ranges of ±1 standard deviation. Because the measured overload factors are included, comparisons are possible, keeping in mind the previously discussed problems associated with a one-point measurement in an unknown distribution and the wide range of values for identical input values. Under the assumption of a normal distribution of overload factors, about two-thirds of all values should fall within the indicated range. The measured overload factor falls well within that range for most of the surveys. An examination of the distributions in Figure 5 indicates that even for Surveys 7 and 4, the measured values
Probability of Overload

FIGURE 8 Overload factors shown in a tentative system of probability of discharge overload represented by probabilities of at least one overload in two consecutive cycles.

ACKNOWLEDGMENTS

The pilot study reported in this paper was funded under an operation grant by the Natural Sciences and Engineering Research Council of Canada. The author would also like to express his appreciation and thanks to two University of Alberta students: Sylvia Dzido assisted in the surveys, and Randall Sonnenberg was responsible for the field work, many calculations, and final figures. The reviewers of this paper are acknowledged with gratitude for their valuable comments and suggestions.

REFERENCES


are still quite representative of the main distribution classes. The limited scope of data in this pilot study did not provide sufficient basis for rigorous statistical testing.

FINDINGS AND CONCLUSIONS

This pilot study indicates that a close link exists between the probabilities of discharge overload and the overload factor. The comparison of the measured and simulated overload factors with the probabilities of discharge overload shown in Figure 7 suggest that for the studied situations, the overload factors may be approximated by the formula for the probability of discharge overload in at least one of two consecutive cycles.

Assuming that the conditions described by the arrival mean and capacity have already lasted for some time, the probability values for oversaturated situations are 1.0, and the calculated values do not apply. This constraint is shown in Figure 8, which illustrates the tentatively proposed system to determine the probability of discharge overload.

An extension of this University of Alberta study will examine the problem in detail on the basis of broad field and simulation data. The results so far provide evidence that it is possible to devise a practical measure that would allow the analyst or designer to determine the seriousness of random overloads. Such an evaluation criterion would be instrumental in understanding suitable signal coordination conditions and assessing the impact of the random delay component or of all three traditional delay elements (uniform, random, and continuous). The newly defined overload factor would represent an easy way of verifying the magnitude of the probability of discharge overload by surveys.


*Publication of this paper sponsored by Committee on Highway Capacity and Quality of Service.*