A measure of oversaturation is developed and validated with field data. A control policy for oversaturated intersections is developed. The results of this work have applicability wherever intersection oversaturation is frequently experienced. The primary objective of this control policy is to delay or eliminate intersection blockage, which is the outgrowth of oversaturation. The control policy developed in this paper (queue-actuated control) is of a tactical nature. It responds to queues (maximum queues allowed at both approaches of an intersection) similar to the way in which the current traffic-actuated control responds to demand. The new control policy can be applied to isolated intersections as well as to coordinated ones. The specific aspects addressed are field validation of the new measure of oversaturation, development of a delay model for queue-actuated control, and effects of queue-actuated control on adjacent intersections. Validation of the delay model and evaluation of queue-actuated control in alleviating the problems of oversaturation are performed with the UTCS-1 simulator.

Congestion on urban street networks has become a familiar occurrence in central business districts across the country and abroad. Urban street grids are commonly used for a variety of often conflicting purposes. They serve through traffic and circulate traffic with destinations on or adjacent to the street grid. Land access is provided to abutting properties. Pedestrians as well as vehicles are serviced. Conflicting vehicular and pedestrian movements are interfaced at grade, with little or no separation.

This system of conflicting movements, modes, and purposes is controlled by a variety of common traffic control devices and techniques, principally signals. The emphasis of current traffic control schemes is directed toward light traffic flows in which flexible coordination of signals will enable traffic platoons to progress along green waves and toward moderate congestion in which the capacity of bottlenecks will be maximized.

When volumes are excessively high, most of the present concepts of optimization become ineffective or invalid (1). As traffic flows increase, the system or points within the system are subject to breakdown or congestion. Saturation begins principally at intersections, the points of maximum conflict and, if unchecked, spreads to other parts of the system.

In saturated conditions, congestion is unavoidable, and therefore the control policy should be aimed at postponing the onset or minimizing the severity of secondary con-
gestion, which is caused by the blockage of intersections that are not the originators of congestion.

This paper covers two major areas of discussion (2, 3) related to signal control of oversaturated flows:

1. Definition of oversaturation and development of measures of oversaturation and
2. Development of a signal control concept better suited to oversaturated flows.

OVERSATURATION

Necessary Definitions

In addition to standard traffic engineering terminologies, the following definitions are used in this paper.

Critical intersection

A critical intersection (CI) is one whose capacity is the limiting value of the capacity of a segment of roadway or an entire system. Thus, there may be more than one critical intersection in a network, each having a different value of capacity. As traffic flow increases through a network, congestion first develops at critical intersections and then spreads to other intersections, depending on the duration and magnitude of the increased flow and the physical characteristics of the roadway.

Tactical control measure

A tactical control measure is one that is responsive to traffic and is usually applied at one or a few critical intersections.

Strategic control measure

A strategic control measure is a control measure that is applied systemwide. It may or may not be responsive to traffic.

Characterization of Oversaturation

Oversaturation occurs when queues fill a significant portion of a block and interfere with the performance of an adjacent upstream intersection. Oversaturation may be experienced at single points, along major arterials, or throughout entire subsections of a network. On a system operating in an oversaturated mode, the problem appears to be areawide. However, the onset of the oversaturated state occurs at one or perhaps a few critical locations at which congestion first begins to develop. Thus, the problem of oversaturation begins as a localized problem.

Definitions of traffic performance states and the characterization of oversaturation were developed as part of the research under an NCHRP project (2).

Definitions of Traffic Performance States

It is not unusual to find the terms congested, saturated, and oversaturated all applied to the same situation, i.e., where one or more vehicles remained to be served at the end of the green phase. A more exact and nonredundant set of terms needed to be developed to characterize the various conditions of traffic performance that could occur. No new terms were envisioned, only a more specific definition of the existing ones. Thus the terms uncongested, congested, saturated, and oversaturated are used throughout this paper, but they have been defined somewhat differently as shown in Figure 1.

Specifically, the definitions have been constructed around a queue formation mech-
Figure 1. Traffic performance states.

- TRAFFIC
  - UNCONGESTED
    - No queue forming
  - CONGESTED
    - SATURATED
      - Queue forming and growing to point where upstream intersection performance adversely affected
    - OVERSATURATED
      - Queue forming and growing to point where upstream intersection performance adversely affected
  - UNSTABLE
    - Queue forming and growing; delay effects still local (transient state may be only of short duration)
  - STABLE
    - Queue forming but not growing; delay effects local

Table 1. Measures of traffic performance.

<table>
<thead>
<tr>
<th>Measures</th>
<th>Descriptors</th>
<th>Descriptors and/or Control Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intersection</td>
<td>Load factor, saturation factor, maximum individual delay per vehicle, number of cycles to clear intersection, number of stops and starts to clear intersection</td>
<td>Queue length, total delay, average delay, link length minus queue length, ratio of queue length to link length, input-output, trapped vehicles, v/c ratio</td>
</tr>
<tr>
<td>System</td>
<td>Total delay, average number of stops per vehicle, density, mean velocity gradient, occupancy, total travel time</td>
<td>Density, occupancy, input-output</td>
</tr>
</tbody>
</table>

Figure 2. Average crossing time for entire approach: (a) site 1 and (b) site 2.
anism. The traffic performance definitions are all described in terms of one approach to a signal. As such they refer to the capacity of one signal's green time for the approach under consideration.

Measures of Oversaturation

Numerous measures of effectiveness (MOEs) are currently used to characterize traffic performance. Some measures describe traffic conditions at points in the system, whereas others describe overall conditions of the system or subsystem. Some of the measures only describe the state of traffic performance, and some have utility as control parameters.

Thus, before any attempt is made to develop a control concept to alleviate the problem of oversaturation, adequate MOEs must be determined. Proper MOEs must describe the conditions of oversaturation and have utility as control parameters if they are to be of any value. They must also be able to predict the onset of oversaturation: the transition from the unstable state to oversaturation. This is where the opportunities for delaying or preventing oversaturation lie.

A number of possible MOEs have been reviewed and evaluated (2) and are given in Table 1.

Point measures

The following measures were selected as the most promising: queue length, ratio of queue length to link length, and link length minus queue length. Each of these point measures of saturation can describe the defined traffic performance states and predict the onset of each state. The latter two measures, those explicitly incorporating link length, are considered to be desirable because they are more general forms.

A limited field study was conducted to determine whether queue-related measures adequately meet evaluation criteria in reality and which of these three point measures best describes the various levels of saturation. The field study and its results are discussed later.

System measures

Evaluation of several candidate system measures revealed that none of them adequately describes the conditions of saturated and oversaturated systems. An alternate measure has been proposed (2).

Data Check of Selected Measures

The primary purpose of a data check of the selected measures is to determine their adequacy in describing or modeling the properties of oversaturation, as well as the onset of oversaturation in reality. Based on an analysis of data collected at two pairs of intersections in Brooklyn, New York, the following conclusions were drawn.

1. The average crossing time (speed) of vehicles leaving the upstream intersection is significantly affected by the queue at the downstream critical intersection. When the clear distance to the end of the queues at the downstream intersection—one of the selected measurements of oversaturation—is less than approximately 230 ft, the quality of operation (expressed in speed) at the upstream intersection starts deteriorating. Data given in Figures 2 and 3 show that the effects are similar for vehicles that did and did not stop at the beginning of green. Queue positions 4 and 3 for sites 1 and 2 respectively both represent a clear distance of approximately 230 ft between the upstream intersections and the end of the queues at the downstream intersections. The clear distance of 230 ft is defined as the critical distance, and a downstream intersection is oversaturated if the distance between the end of the queue and
the upstream intersection becomes less than the critical distance. (The data used here were collected from two pairs of intersections with block lengths of 530 and 810 ft. Further field validation is desirable for blocks shorter than 500 ft.)

2. Although the quality of operation is reduced, the productivity of the upstream intersection is not affected by the queue at the downstream intersection until spillback occurs. The output will, of course, be reduced when spillback occurs. As shown in Figure 4, there are no discernible trends between the average headway and the queue length, for either stopped vehicles or not stopped vehicles. There is no statistically significant difference between the average headways. Because time headway equals space headway divided by speed and because time headways remain the same, space headway decreases as speed decreases. This implies that, as the queue reaches the critical length, the speed and space headways decrease so that the output remains the same. The decrease of space headways results in a more compressed traffic stream and a higher density.

3. Regardless of the queue length at the downstream intersection, the status of the downstream signal indication affects the quality of operation (speed) at the upstream intersection. However, it will not affect the selected MOE (Fig. 5).

4. Several other analyses were made to determine effects of the queue length at the downstream intersection, but are not presented here. These include start-up times, left turn movements, and others.

EVALUATION OF EXISTING CONTROL MEASURES

Developing a control technique better suited to the problem of oversaturation required that the effectiveness of existing control measures be thoroughly evaluated. This evaluation, therefore, included existing control measures that are well-known to practicing traffic engineers as well as advanced concepts that have not been fully tested in the field.

Control Measures Evaluated

In essence, in any form of traffic signal control only three parameters may be manipulated by the traffic engineer: cycle length, allocation of green to competing movements (split), and timing interrelationships between adjacent signals in a system (offset). Therefore, the control parameters evaluated can be classified as follows:

1. Cycle length (strategic) includes short cycle length versus long cycle length and a common cycle length for a system.

2. Split includes strategic control measures, such as uniform bandwidth and gradual change of bandwidth, and tactical control measures, such as Gazis' saturated intersection policy (4, 5) and Longley's control strategy for congested computer-controlled networks (6).

3. Offset (strategic) includes smooth flow theory (7-9), red progression (10), progression versus simultaneous, reverse progression, and combination of progression and simultaneous.

Evaluation Techniques

Inasmuch as the measure of oversaturation is a function of queue length, the primary evaluation criterion of the performance of a given control measure under oversaturated flows should be its ability to limit the queue length to some desired level. Average delay per vehicle is used as a secondary evaluation criterion.

Various existing control measures were tested on fictitious networks as well as at the sites used for MOE evaluation. Extensive simulation was done under oversaturated flow conditions with the UTCS-1 simulator.
Figure 3. Average crossing time for entire approach for vehicles not stopped at beginning of green: (a) site 1 and (b) site 2.

![Figure 3](image)

Figure 4. Headway versus queue position at (a) site 1 and (b) site 2.

![Figure 4](image)

Figure 5. Average crossing time by downstream indication, site 1.

![Figure 5](image)
Results

Based on the results of simulation analysis of the current control measures, the following conclusions are made.

1. Some of the strategic control measures evaluated are very effective in delaying or eliminating oversaturation at critical intersections. They, however, tend to degrade the performance of adjacent intersections in the system. For example, a simultaneous offset plan significantly reduces the queue length at an oversaturated CI, but at the same time increases queue length and delay at other intersections within the system.
2. Strategic control measures generally reduce queues at the CI more than necessary to prevent oversaturation and reduce the productivity of the CI. Control measures such as smooth flow theory and reverse progression usually reduce queue length more than enough to keep a CI from being oversaturated or the upstream intersection from being blocked by extended queues. The productivity of a CI is significantly reduced by a strategic control measure.
3. Tactical control measures do not act positively to prevent the blockage of intersections, which is the prime concern during the period of oversaturation. Under oversaturated flows, a control measure that is more sensitive to intersection blockage is required because of the undesirable effects of such a blockage.
4. A new control concept is desired that not only ensures the prevention of intersection blockage as much as possible but also minimizes the degrading of other intersections in the system. This new control concept should be capable of maintaining high productivity at a critical intersection. Productivity is more important than the quality of operation during a period of oversaturation.

A NEW CONTROL CONCEPT

In oversaturated conditions, congestion at a critical intersection is unavoidable; therefore, the control policy should be aimed at minimizing the effects of queues at the CI on the upstream intersection. As volumes increase further, oversaturation of a CI may also become unavoidable. Then the control policy should be aimed at postponing the onset of oversaturation or preventing the blockage of the upstream intersection. Prevention of intersection blockage becomes more important with arterials or networks where traffic flow is heavy in all directions. In such cases, intersection blockage not only increases the delay through the system but also affects more segments of that system.

Queue-actuated signal control was developed to ensure, as much as possible, the prevention of intersection blockage. It also makes use of all available green time by creating continuous demand (queues) at all approaches. It allows more positive control of queue length at a given intersection.

It is a control policy in which an approach receives green automatically when the queue on that approach becomes equal to or greater than some predetermined length. When the queue at one approach becomes equal to or greater than a given length \(Q_{\text{max}}\), that approach receives green regardless of the conditions on the conflicting approaches, assuming that no other approach reaches \(Q_{\text{max}}\) simultaneously. Thus, how much green one approach receives during a given time interval depends on the link length and the number of lanes, as well as the flow rates on both approaches.

It is obvious that, when one approach has a relatively low flow rate, drivers on that approach suffer long delays because the queue on the approach takes a longer time to reach the predetermined length. When volumes become extremely heavy on both approaches, the effective cycle length becomes very short, because less time is required for the queue to reach a given length. This problem can be reduced somewhat by selecting the proper \(Q_{\text{max}}\). As will be shown later, the effective cycle length is only a function of \(Q_{\text{max}}\), given the flow rates.

Cycle lengths that are too long or too short can be avoided by imposing maximum or
minimum values of green time, which will determine the upper and lower bounds of
cycle length and green time.

One of the advantages of this queue-actuated control is that it is a more positive way
to control queue length and prevent intersection blockage than any other method reviewed.

**Delay and Cycle Length at an Intersection Under Queue-Actuated Control**

Newell (11) and Sagi and Campbell (12) expressed the delay at a signalized intersection under saturated flows in the manner shown in Figure 6.

When queue-actuated signal control is applied to a signalized intersection, the delay at both approaches during the i th cycle can be expressed as shown in Figure 7. Notation used for this and later figures is as follows:

- $Q_1$, $Q_2$ = maximum queue lengths allowed at approaches 1 and 2, in vehicles or feet;
- $Q_{11}$, $Q_{21}$ = queue lengths at approaches 1 and 2 at the end of green, in vehicles or feet;
- $X_1$, $X_2$ = flow rates at approaches 1 and 2, in vehicles per sec ($X_1$ and $X_2$ are assumed to have a Poisson distribution with expected values $E(X_1) = \lambda_1$ and $E(X_2) = \lambda_2$);
- $t_{11}$ = effective red phase at approach 1 during i th cycle, in sec;
- $t_{21}$ = effective green phase at approach 1 during the i th cycle, in sec;
- $t_i$ = effective cycle length during i th cycle ($t_i = t_{11} + t_{21}$), in sec; and
- $S$ = saturation flow (service rate), in vehicles per sec.

$S$ is assumed to have a uniform distribution and to be equal to or greater than $X_1$ and $X_2$. When $S < X_1$, queues become infinite and the system breaks down, because $(S - X_i)$ is the net discharge rate from an intersection.

In addition, it is assumed that lost time due to start-up delay and the amber phase is uniform. Lost time is not included in the delay diagram, but is dealt with separately.

Because of statistical fluctuations of flows, four situations could arise even under oversaturated flow conditions.

**Case 1**

In case 1 (Fig. 8), switching from one phase to another occurs when queues at one approach reach a predetermined length ($Q_{11}$ or $Q_{21}$). Queues at the end of green ($Q_{11}$ or $Q_{21}$) are greater than or equal to 0 at the moment of the termination of green.

**Case 2**

In case 2 (Fig. 9), queues at approach 1 are completely cleared before queues at approach 2 reach $Q_{2}$ and the green phase at approach 1 is terminated.

**Case 3**

In case 3 (Fig. 10), queues at approach 2 are completely cleared before queues at approach 1 reach $Q_{11}$ and the green phase at approach 2 is terminated.

**Case 4**

Because case 4 (Fig. 11) would be likely under light flows, $Q_{11}$ is assumed to equal 0.

**Probability of each case**

Let $P_1$ be the probability of queues being cleared at approach 1 before green is terminated and $P_2$ be the probability of queues being cleared at approach 2 before green is
Figure 6. Delay at one approach of intersection.

For One Approach

Queue

\[ \text{Time} \]

\[ \text{Cycle} \]

Figure 7. Delay at both approaches of intersection with queue-actuated signal control.

Figure 8. Oversaturated flow: case 1.

Figure 9. Oversaturated flow: case 2.

Figure 10. Oversaturated flow: case 3.
terminated. Then the probability of case 1 is \((1 - P_1)(1 - P_2)\), the probability of case 2 is \(P_1(1 - P_2)\), the probability of case 3 is \((1 - P_1)(P_2)\), and the probability of case 4 is \((P_1)(P_2)\).

Under heavy flow conditions, cases 2, 3, and 4 do not likely occur because (a) every vehicle is expected to stop at least once, and there will not be a period of green without queue; and (b) their occurrence is indication that this control measure is not effective any longer.

The total delay during a given cycle caused by the existence of a queue is the summation of the shaded areas of the delay diagram shown in Figure 12.

The total traffic delay due to lost time \((4, 5)\), such as start-up delay and the amber phase not used, for the \(i\)th cycle can be computed as

\[
D_L = \frac{L}{t_i} \frac{1}{2} \left[ t_1(X_1 + X_2) \right] = \frac{L}{2} (X_1 + X_2)
\]

where \(L = \text{lost time per cycle}\).

Thus, the total delay experienced by traffic on both approaches of an intersection is

\[ D_t = D_Q + D_L. \]

The total delay for each case of four different delay diagrams is summarized in Table 2. The table also includes the effective cycle length, as well as the effective lengths of green and red phases. For simplicity, the maximum queues allowed on both approaches were made the same \(Q_{1\text{a}} = Q_{2\text{a}}\). The expected delay and effective cycle length for all cases are

\[
\text{Average delay} = \sum_{i=1}^{4} P(i)E[D_t(i)]
\]

\[
\text{Average effective cycle length} = \sum_{i=1}^{4} P(i)E[C(i)]
\]

where

- \(P(i)\) = probability of case \(i\),
- \(E[D_t(i)]\) = expected total delay for case \(i\), and
- \(E[C(i)]\) = expected effective cycle length for case \(i\).

The details of the derivation of the equations given in Table 2, evaluation of \(P(i)\), \(E[D_t(i)]\), and \(E[C(i)]\), and a sample application of the delay model are discussed by Lee (3).

Comparison With Other Control Measures

The delay characteristics of an isolated intersection were also compared for three different control measures: fixed time, Longley's control strategy, and queue-actuated control (Fig. 13). The comparison demonstrated that, as volume increases, queue-actuated control results in the least average delay per vehicle, although at low volumes queue-actuated control produces the greatest delay, as expected.
Table 2. Total delay and effective cycle length.

<table>
<thead>
<tr>
<th>Case</th>
<th>Total Delay</th>
<th>Effective Cycle Length, Green and Red Phases</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( D_1 = \left( \frac{1}{2} \right) \left( \frac{S}{X_1} \right) \left( \frac{Q_x - Q_{11}}{X_1} \right) + 2(Q_x + Q_{11} - \frac{S^2 - SX_1 - SX_2}{X_1 X_2} \frac{Q_x - Q_{11}}{S(X_1 + X_2)} + \frac{1}{2} \right) (X_1 + X_2) )</td>
<td>( t_1 = \frac{1}{\left( \frac{X_1}{X_2} \right)} ) ( t_{11} = \left( \frac{Q_x - Q_{11}}{X_1} \right) ) ( t_{21} = \frac{S - X_1}{X_2} ) Same as case 1</td>
</tr>
<tr>
<td>2</td>
<td>( D_2 = \left( \frac{1}{2} \right) \left( \frac{S}{X_1} \right) \left( \frac{Q_x - Q_{11}}{X_1} \right) + 2Q_x + \frac{(S - X_1)(Q_x - Q_{11})}{S} X_2 + \frac{Q_x^2 X_1 X_2}{S(S - X_1)(Q_x - Q_{11})} + \frac{L(X_1 + X_2)}{2} )</td>
<td>( t_1 = \frac{Q_x - Q_{11}}{X_1} + \frac{Q_x}{X_2} ) ( t_{11} = \frac{Q_x}{X_1} ) ( t_{21} = \frac{Q_x}{X_2} )</td>
</tr>
<tr>
<td>3</td>
<td>( D_3 = \frac{Q_x^2}{2X_1} + \frac{2Q_x X_2}{2X_1} + \frac{Q_x^2}{2(S - X_3)} + \frac{Q_x^2}{2X_2} + \frac{L(X_1 + X_2)}{2} )</td>
<td>( t_1 = \frac{Q_x}{X_1} ) ( t_{11} = \frac{Q_x}{X_2} ) ( t_{21} = \frac{Q_x}{X_2} )</td>
</tr>
<tr>
<td>4</td>
<td>( D_4 = \frac{Q_x^2}{2X_1} + \frac{2Q_x X_2}{2(S - X_3)} + \frac{Q_x^2}{2(S - X_3)} )</td>
<td>( t_1 = \frac{Q_x}{X_1} ) ( t_{11} = \frac{Q_x}{X_2} ) ( t_{21} = \frac{Q_x}{X_2} )</td>
</tr>
</tbody>
</table>

Figure 12. Total delay caused by queue.

Figure 13. Delay comparison of different control measures.
Queue-Actuated Control in Coordinated Intersection

The underlying philosophy of coordinated systems is generally one of establishing signal timing to facilitate the uninterrupted movement of through vehicles along a roadway. In saturated and oversaturated operations, however, development of signal timing for such a purpose has little justification in that all vehicles must join queues and stop at least once somewhere upstream of the critical intersections. In practical terms there is no through traffic.

Queue-actuated control in a coordinated system of signals is functionally similar to the more conventional traffic-responsive control equipment, in that it places an essentially cycle-free traffic controller in the midst of a set of fixed-time signals.

It is worthwhile, however, to consider how the queue-actuated control policy performs when implemented at a critical intersection, which itself is in the midst of a coordinated system of signals. Inasmuch as this policy is queue-responsive, i.e., allocation of green time depends on some maximum queue length, the signal at the CI is driven by vehicle inputs from the immediate upstream intersections. Two extreme conditions can exist.

One extreme is when vehicle inputs are all provided by the through link upstream of the intersections upstream of the critical intersection. In this case, green at the CI is initiated as soon as through vehicles reach maximum allowable queue. Thus the relative offset between these two intersections looks like a forward progression. The other extreme is when all vehicles feeding the queue at the CI are turn-ins from the upstream intersection. In such a situation, the maximum allowable queue length will be reached before the initiation of green at the upstream intersection, and the relative offset looks like a reverse progression. For some combination of through and turn-in traffic, the relative offset appears to be simultaneous.

Thus, the actual initiation time and duration of green times at the CI under the queue-actuated control policy depend not only on the predetermined value for maximum allowable queue on the competing approaches and their respective volumes but also on the way in which these volumes input into the competing links.

The situation downstream of an initial intersection operating under this policy is different, in that a fixed-time signal with predetermined green time allocations receives traffic from an essentially cycle-free signal. In such a situation, in theory, a green phase can occur at the CI that will provide more traffic to the next downstream signal than it can handle, and thus create a saturated condition there. In such a situation the problem of saturation will not have been eliminated, only relocated. This can only occur, however, when traffic on the cross street at the critical intersection is very light; thus the main street is allowed an unusually long green time. In practice, however, this is quite unlikely, for one of the inherent features of the critical intersection is the presence of heavy traffic on both of the competing approaches.

The queue-actuated control in the midst of coordinated intersections was extensively tested by simulation on several networks, and its performance was compared with that of the existing control measures discussed previously. From this simulation, it was determined that (a) the adverse effects of the control are limited to the immediate downstream intersection; (b) although delay is increased at the immediate downstream intersection, queue-actuated control produces the least system delay; and (c) queue-actuated control results in the highest output at a critical intersection.

FURTHER RESEARCH

Further work is desirable on the following aspects:

1. The delay model needs further refinement, and more consideration should be given to the simplified assumptions used here.

2. The effects of errors in detection must be studied because of the limited accuracy of present detector systems available and because queue-actuated control must
rely on accurate detection of queue length.

3. The more desirable form of any control measure is one that could be used in oversaturated and undersaturated periods. Therefore the possibility of extending queue-actuated control to the period of undersaturation should be investigated.

4. An analytical model for the effects of queue-actuated control on adjacent intersections would help engineers better understand the effects of this control in the midst of coordinated signals.

REFERENCES