PLANNING CONSIDERATIONS FOR TRAFFIC SIGNAL SYSTEMS

M. G. Davis and W. B. Hirsch,
Wilbur Smith and Associates, New York

Computerized traffic signal systems are receiving increasing consideration as a means of alleviating roadway congestion. Many of the systems that are being considered are based on the use of a digital computer. The increased scope and complexity of these signal systems require the application of sound system engineering processes if successful system operation is to be achieved. A summary of the systems engineering process is related, and its application to traffic signal systems is emphasized. Examples are given that examine several traffic control system elements with respect to configuration, reliability, and accuracy.

•IN response to the continual increase in urban traffic congestion, an expanding number of cities are investigating the use of computerized traffic signal systems to achieve optimum use of their existing road networks. Impetus toward the optimal use of existing road systems provided by the federally aided TOPICS program has also resulted in increased interest in computerized traffic control systems. Although the state of system control hardware far exceeds the present state of traffic control technology, justification of a real-time control system using a digital computer can often be reached through consideration of the following factors:

1. For a signal system totaling fifty intersections or more, the cost of a digital computer system is not greatly in excess of the cost of more common system configurations and may be less for equal control capability;
2. The existing state of traffic control technology is not reflected in conventional hardware, and the control algorithms implemented by conventional hardware are not easily altered by equipment modification; and
3. Because a traffic control system represents a substantial investment that should have a useful life of 10 to 20 years, the flexibility and computing power of the digital computer provide the control hadrware that is best able to incorporate advances in control theory.

Implementation of computer-oriented traffic control systems of the scope under consideration can best be achieved by the application of sound systems engineering technology. The systems approach includes the total and continual planning of the system from inception to phaseout. The need for this type of planning and constant evaluation lies in the complexity of the engineering task. The justification for this discipline is derived from the benefits to be gained by successful system implementation and the penalties to be suffered by unsuccessful implementation.

The computer-oriented traffic signal systems currently being planned and instituted differ in several important aspects from the systems of the past two decades. One area of significant difference is the increased size and scope of the equipment serving the new systems and its capacity for expansion and modification. A second difference of great importance is the potential ability of the system to perform optimum control

---

Sponsored by Committee on Traffic Control Devices and presented at the 50th Annual Meeting.

strategies based on modern control theory. Because the degree to which any control system can achieve optimization depends to a large extent on the accuracy and refinement with which the system can measure and evaluate the quantities to be optimized, the accuracy requirements and quantity of detector data can be expected to increase substantially. The increased data communications requirements, number of control and detection elements, severe environment in which control and detection elements must operate, and economic constraints create system problems that must be resolved to achieve reliable system operation.

The advanced state of hardware technology has added a new element of choice to the design of the traffic control system. Historically, the decisions in traffic control systems centered on whether the system was to be completely pretimed or was to be traffic responsive. Once this basic decision was made, the remaining decisions dealt largely with the selection of a manufacturer to supply the equipment. Because the functional differences among the various systems offered by the traffic equipment manufacturers tended to be negligible, the choice of the system supplier had only minimal effect on system planning, system operation, and system cost.

Today, the range of choices available is extremely wide and will probably grow. Previously the system control functions were integral with the hardware and were fixed; now the element of system software has added a completely new dimension to system design. The implications of this added element are extensive and, at times, subtle. Basically, the operation and efficiency of the entire control system is a function of the software. However, before software requirements can be defined, the overall requirements of the system must be determined. Because the choice of possible system requirements and functions has become quite broad and because all elements of the system can be affected by these basic requirements, it follows that a systematic procedure is required to achieve an optimum configuration of hardware and software.

Because software will perform a major role in the implementation of system requirements, the specifying of software becomes a key element in the system specifications. Historically, specifications for traffic signal systems have generally described the control functions of existing hardware or have treated the subject of how control was to be achieved superficially. Because there was little difference among the control algorithms available, these specifications may have been satisfactory. However, the complexity of digital computer control demands that precise specifications be made if the intent of the system requirements is to be met.

If the full potential of the computerized traffic system is to be realized, a continual program of control evaluation and modification must be established. The system requirements should necessarily reflect this need in both the hardware and the software specifications. In summary, the modern computer-based traffic control system can no longer be viewed as a piece of hardware selected from a catalog; it must be designed to meet the growing transportation problem by implementing the expanding traffic control technology. The complexities of the technology do not permit success to occur without adequate planning.

A system denotes the total resources brought to bear on a problem. These resources include both personnel and hardware and extend beyond equipment performance to include training, maintenance, operation, and evaluation procedures. Figure 1 shows the total involvement of the system engineering discipline. Involvement with the total problem is of itself not the key to successful system engineering. The benefits from systems engineering

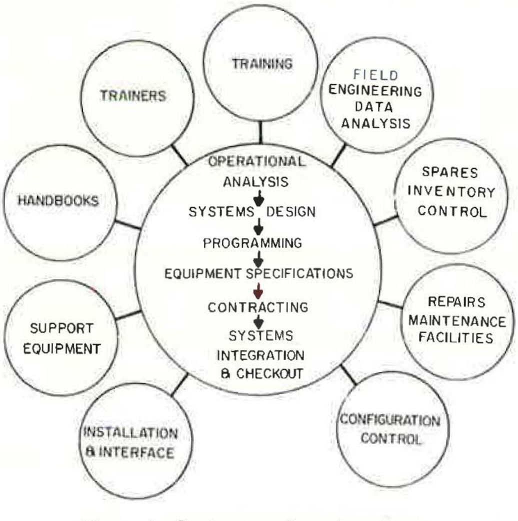


Figure 1. Systems engineering scope.

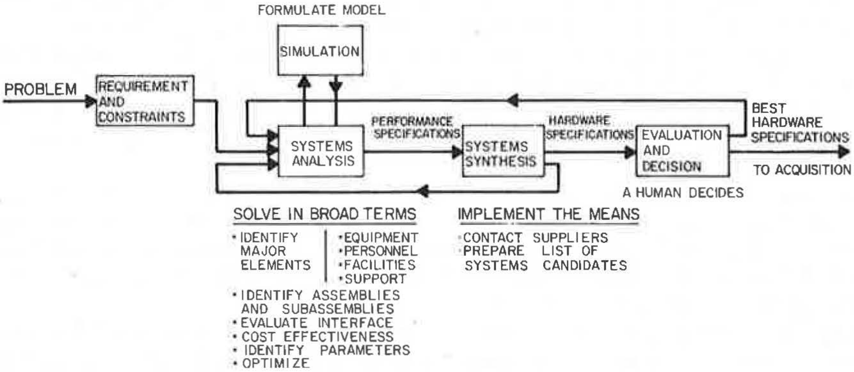


Figure 2. System design procedure.

are not realized until an iterative procedure is established that subjects all aspects of the problems to a systematic "feedback" review. This process is designed to evaluate each component of the problem solution in the context of its relationship both to the total problem and to other components of the solution.

The systems engineering procedure can be divided into five stages: concept formulation, system definition, acquisition, deployment, and phaseout.

In concept formulation the tasks are to define the problem, set objectives, and conceive alternative solutions to the problem. During the concept formulation stage, the technical and economic constraints applicable to the problem are also defined. At this stage all topics are considered at the general level, and continual care is taken not to make specific decisions prematurely. The completion of this phase should result in an accurate definition of the problem and the establishment of requirements to meet the problem.

In the system definition stage, the analysis, simulation, synthesis, evaluation, and selection are performed. At this time the system components (hardware, personnel, facilities, and support) for each alternative plan are defined. These alternatives are then subject to an optimization process as shown in Figure 2. During this process the basic system requirements and constraints are subject to the iterative analyses leading to decisions. The benefit of this stage is derived from the trade-offs affecting each component for the good of the entire system. A secondary benefit of substantial importance that emerges from this stage is the quantification of system elements before making trade-offs. This process, which requires the assignment of numerical values to system elements, also demands that analysis of the system elements be performed to permit accurate quantification. As a consequence no system element is left to change.

One of the factors of particular importance to traffic control systems is the examination of the interfaces among system elements. The evaluation of each subsystem in relation to its impact and compatibility on other parts of the system is fundamental to the basic system operation. A specific example concerning the interfaces between a computer and its vehicle detectors will be covered later in this paper.

The subject of system reliability may appear in the system design procedure as both a system requirement and a system constraint. It is a universal fact that no element of the system can be considered independent of its reliability considerations. The attainment of reliability goals also strongly influences the basic configuration of the system design and will be discussed later relative to its traffic signal system implications.

The final products of the system definition phase are the hardware and software specifications that will be used for procurement. Accompanying the hardware and software specifications may be specifications or instructions that define the operational aspects of the system implementation, evaluation, maintenance, and personnel requirements.

During the acquisition phase the necessary research, development, and design work is performed leading to production and equipment availability. Depending on the hardware specifications, the acquisition phase could simply result in the procurement of existing hardware. Software development would also proceed during acquisition. As a component of the system development phase, a timetable would have been created that would define the phasing of necessary on-site construction to effect a smooth deployment of the equipment. The construction of required facilities would also proceed during the acquisition phase.

The installation, checkout, and operation of the system occur during the deployment phase. The logistic support, maintenance implementation, and system operation procedures established during system definition would be implemented at this time. Once the system becomes operational, the evaluation concepts would be utilized and the success of the system would be established.

Planning for system phaseout would probably be superficial for a traffic control system. It is necessary, however, to recognize that equipment has a finite useful lifetime and that operation with technically obsolete equipment can be a false economy. Consequently guidelines should be established to aid in determining the point at which system replacement should be considered.

It is important to note that systems engineering is an interdisciplinary activity drawing on the talents of specialists in all aspects of a problem. The economic aspect of procurement of a traffic control system is not to be minimized. The investment in a traffic control system can be a substantial one for the community with the consequence that the economic constraints may be severe. A well-developed analysis combining both economic and engineering talents is needed to ensure that the best system is provided for the available money. In addition, a traffic control system is one of the few public facilities that has the potential of paying for itself by reducing delay to the motorist.

Figure 3 shows the basic configuration of the traffic control system and its data flow. The basic elements shown include detectors, D_n, controllers, C_n, data transmission devices, DC_n, a media for the transmission of data, TM, and the computer.

Data from the detectors, D_n, are converted to a mode suitable for transmission by data convertor DC_{1n}. The data are then transmitted by the transmission medium, TM, to data convertor DC_{4n} where they are converted into a form usable by the data input function of the computer.

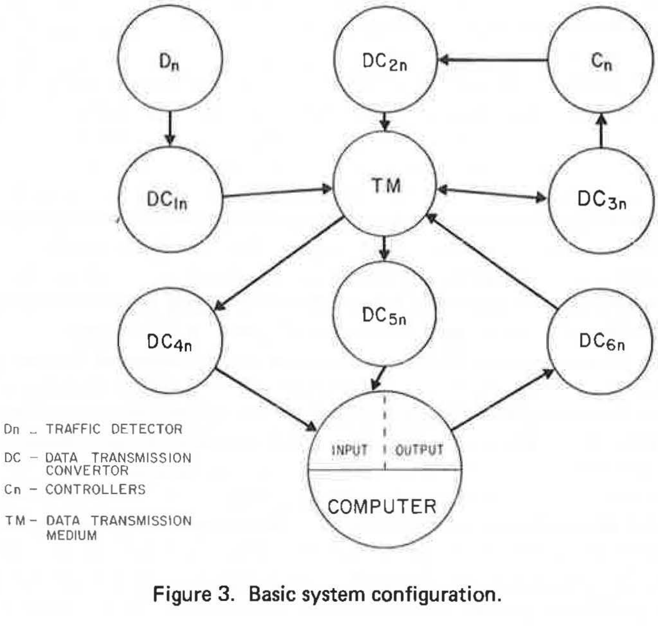


Figure 3. Basic system configuration.

The computer control decisions are converted from their form at the data output function of the computer to a form suitable for transmission to the controllers by data convertor DC_{6n}. The data then proceed via TM to data convertor DC_{3n} where they are converted to a form usable by controller C_n.

To establish that the desired control action occurs, a confirmation of the controller response is sent to convertor DC_{2n} where it is converted for transmission and proceeds back to the computer input function via TM and DC_{5n}. The only function present in this basic diagram that is not absolutely essential to system operation is the feedback of controller configuration via DC_{2n}, TM, and DC_{5n}. However, the assumption will be made that confirmation of controller synchronization is a basic system requirement.

Figures 4 and 5 show two examples of practical implementation of the basic traffic control system. The system shown in Figure 4 is characterized by a unique data transmission medium and unique data transmission convertors for each communications channel. This is probably the most direct implementation of a traffic control system and will serve as a model to apply examples of system design considerations.

Initially the detector channels will be examined. From the reliability standpoint it may be observed that the failure of any element in the detection channel will result in invalid detector data. It should also be recognized that since the detector channels are independent the failure of an element in one channel will not affect another channel. One option to enhance the reliability of the detector channel would be to provide a redundant channel; however, this raises several questions.

The basic question to be asked is, What is the effect on the total system operation if a single detector channel fails? Before this question can be answered, a determination would have to be made of the possible differences that would occur in the system operation depending on the manner in which the failure occurred. Three possible results of the channel failure can be considered:

1. The channel indicates continuous vehicle presence;
2. The channel indicates no vehicle presence; or
3. The channel shows intermittent vehicle presence with no relation to the actual traffic at the detector.

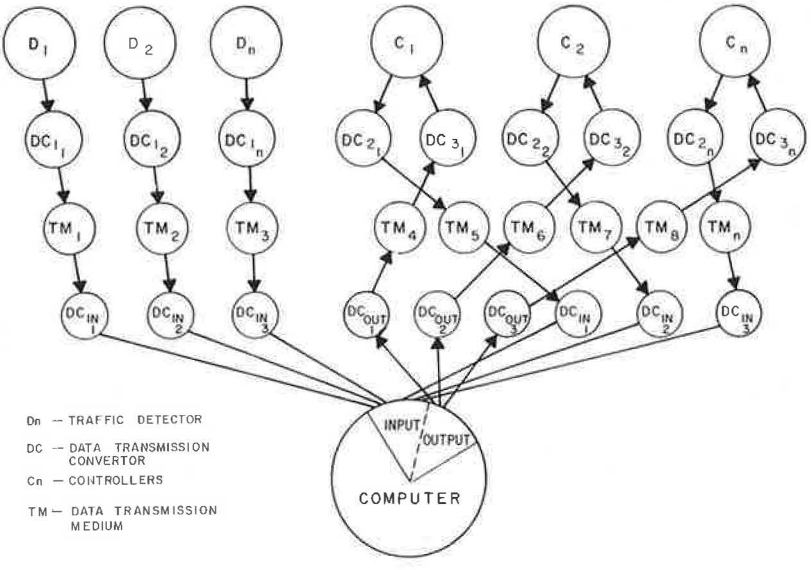
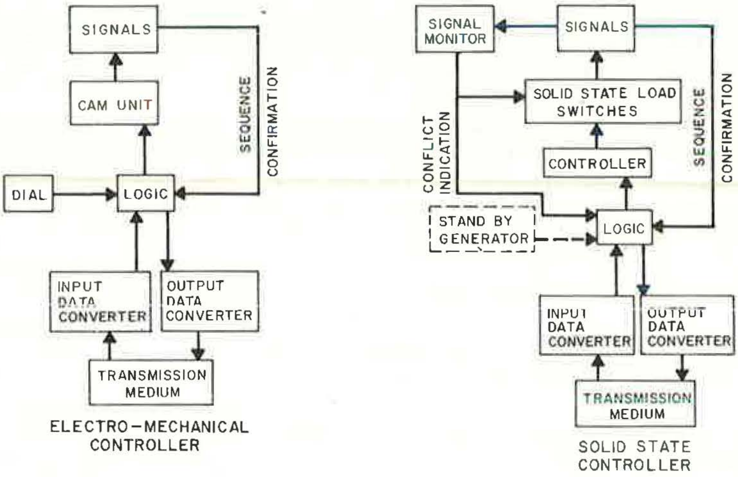


Figure 4. Example system configuration.



ELECTRO-MECHANICAL CONTROLLER

SOLID STATE CONTROLLER

Figure 5. Pretimed controller configurations.

Because a steady failure can probably be detected and the failed detector's input data disregarded by the computer, the question of the effect on system operation of the absence of the failed detector data must be resolved. If system operation were unacceptably degraded by the absence of a single detector, then a vote might be given to the consideration of a redundant channel, but even here another question emerges: How long can degraded system operation be tolerated compared to the expected time required to repair the failed channel? If the degraded performance cannot be tolerated until the channel is repaired, another vote might be given for redundancy. Still other questions remain: How frequently (what is the mean time between failures) can a detector channel be expected to fail? What is the probability that the failure will be of a continuous type that can be more easily detected? Of course the very basic question of whether automatic detection of a failed detector can be initially provided requires resolution.

Consideration of the detector channel that failed intermittently would require additional analysis and will heavily involve consideration of a means of detecting this phenomenon.

If the evidence weighed in favor of considering redundant detector channels, an entire new analysis process would have to begin that would evaluate whether the entire channel should be redundant, whether just the least reliable element of the channel should be redundant, or whether the channel could be made satisfactorily reliable by redesign or selection of more reliable components.

The intent of the foregoing description is to demonstrate the iterative thought process intrinsic with system design procedures; it was not intended to analyze exhaustively a hypothetical problem.

A second system problem using the same detector channel can be examined based on actual data. This is the problem of detector data accuracy. Because detector data provide the real-time input descriptive of traffic flow, it is reasonable to assume that errors in this input will result in control errors unless the means are provided to compensate for the errors. For example, assume that a loop-detector is the detection element and the parameter to be examined for accuracy is the duration of time a vehicle is within the detector's zone of detection. To develop this example the following assumptions will be made:

1. The loop is 6 ft square and is located in the center of a 12-ft wide traffic lane;
2. The vehicle is 18 ft long;
3. The vehicle is traveling at a speed of 30 mph (44 fps); and
4. The duration of time that the vehicle is over the loop is given by

$$PW = \frac{L + l}{v}$$

where

PW = pulse width in seconds,
L = vehicle length in feet,
l = dimension of loop parallel to vehicle motion in feet, and
v = vehicle speed in feet per second.

Therefore

$$PW = \frac{18 + 6}{44} = 0.545 \text{ sec}$$

Now let us examine the sources of error to this calculated value:

1. Assume a random error of ±2 percent in basic detection accuracy; ±2 percent
2. Assume an error of ±12 percent due to the change in relationship of the vehicle to the loop's electromagnetic field as the relationship of the vehicle's centerline to the loop's centerline changes; ±12 percent
3. Assume an error of ±5 percent due to the recovery time of the loop detector (if two vehicles are following closely the detector generally does not fully recover quickly enough, and the pulse width accorded the following vehicle is decreased); ±5 percent
4. Assume an error due to the drift of the detector with respect to operating temperature of ± 25 percent. (Note: This is a wide variable among loop detectors and varies from a few percent to nearly 100 percent.) ±25 percent

Total accumulate error = ±44 percent

Thus, if tolerances are additive, the actual pulse width produced by the loop detectors may lie between 0.305 and 0.785 sec (i.e., 0.545 sec ± 0.240 sec) and should be carefully evaluated with respect to system requirements. Further, the specifications for this or any other system element should reflect the accuracy requirements of the system.

Two additional sources of error exist in the detector channel—the transmission medium and the data convertors. The source of error due to the transmission medium can generally be disregarded for traffic control purposes.

Two common types of data conversion elements are relays and frequency division multiplex equipment. Relays appear to be the most elementary of devices, but even a simple relay can be a source of error. The relay error in accurately duplicating a pulse width occurs for two reasons. First the drop-out time of a relay is generally greater than the pull-in time. The drop-out time will also be increased by a factor of two to three if the relay coil is shunted by a diode for transient reduction purposes. Second, the response time of a relay changes with respect to its coil resistance, which is a function of temperature. Because the characteristics of relays vary widely with the size and type of relay, the errors due to relay characteristics may be insignificant: however, the errors should be known and evaluated with respect to the system requirements.

Frequency division multiplex equipment can possess both excellent and poor response characteristics depending on its design. In general, frequency division multiplex equipment that has a very narrow bandwidth (100 Hz or less) tends to have very sluggish response times that may disqualify it as an element in a pulse-width-dependent transmission channel.

A third example of system analysis can be shown by comparing two methods of implementing controller elements in a traffic signal system. Figure 5 shows an electromechanical controller and a completely solid-state controller operating under system control.

In both examples the fixed intervals in the signal sequence will be timed by the controller. Pulses generated by the computer will be used to terminate the green intervals, thus providing the means of regulating the cycle, split, and offset of the controller.

A feedback path is also provided in both controllers to verify that the controller is in synchronism with the computer commands.

The logic element interfaces the data convertors with the controller and also switches the controller (or cam unit) to the dial or standby generator in the event of loss of synchronism or upon receipt of a command to begin standby operation. During standby operation the signal timing will be established entirely by the dial unit or the controller and standby generator.

The most noticeable difference between the two configurations is the addition of a signal monitor element and a standby generator element in the solid-state example. The signal monitor is required in response to an assumed system requirement that no component failure should result in dangerous, conflicting signal indications. The mechanical construction of the cam unit makes this requirement intrinsic with the electromechanical controller. However, because the solid-state load-switching devices possess the possibility of failing in either a conducting or nonconducting state, a means must be provided to detect a failure in this element that would result in dangerous signal indications.

If equal reliability of the two controllers is to be achieved, then the combined reliability of the solid-state load switches must equal the reliability of the cam unit. Although solid-state devices are generally more reliable than electromechanical devices, the system engineer cannot be satisfied with broad generalities but must examine all elements and their relationships to the extent that the available data permit.

A similar analysis can be made of the dial unit, the standby generator, and the solid-state controller. In the electromechanical controller the dial unit, by its design, can remain in synchronism with inputs from the computer. The dial unit in conjunction with its cam unit also provides the same function of the solid-state controller, solid-state load switches, standby generator, and signal monitor. Consequently, because there are more elements in the solid-state configuration, their individual probabilities of failure must be higher so that their combined reliability equals the electromechanical system.

It should be noted that the standby generator might be incorporated into the solid-state controller. If this element ceases to be uniquely identifiable, the system reliability consideration becomes one of determining the effect, if any, of the reliability of the controller due to the added standby components.

The intent of this discussion is not to imply that electromechanical devices are more (or less) reliable than solid-state devices. The intent is to emphasize that only by rigorous analysis of all components and elements for a proposed signal system can an optimum design be attained.

REFERENCES

1. Reliability Stress and Failure Rate Data for Electronic Equipment. MIL-HDBK-217A, 1965.
2. Shooman, M. L. Probabilistic Reliability. McGraw-Hill, 1968.
3. Stern, S. Traffic Flow Data Acquisition Using Magnetic-Loop Vehicle Detectors. Highway Research Record 154, 1966, pp. 38-52.