Responsive Traffic Signals Designed with Petri nets

Michel dos Santos Soares, Jos Vrancken
Faculty Technology, Policy and Analysis
Delft University of Technology
P.O. Box 5015, 2600 GA
Delft, The Netherlands
email: (m.dossantossoares, j.l.m.vrancken)@tudelft.nl

Abstract—Traffic-responsive techniques make use of real-time measurements acquired with sensors to calculate in real-time suitable settings. Traffic signal systems that react to changing traffic conditions are an important component for improving transportation efficiency. The dynamic behavior of a group of traffic signals controlling a network of intersections is a complex discrete event system that can be modeled using Petri nets. In this paper, one purpose is to design a mechanism based on Petri nets with time interval associated to places to extend the green time to a main road depending on the demand of non-priority roads. Another objective is to try to allow green phases for a sequence of intersections in a small network in order to improve traffic flow for a platoon of vehicles. Modular characteristics of Petri nets are used to address complexity and design models in increasing levels of detail.

KEY WORDS

I. INTRODUCTION
Traffic congestion is a reality in many countries. This leads to economic losses, augments air pollution and has a great impact on the quality of life. One common approach to try to handle traffic congestion is to build more infrastructure, such as roads and viaducts. However, in the past years, it is becoming increasingly more difficult to build more infrastructure. Not only the high cost, but also the lack of space and the environmental damage of building new roads have to be considered.
A different approach is needed, based on applying intelligence in order to manage traffic flow in a more effective and efficient manner. This leads to the research area named Intelligent Transportation Systems (ITS) [1], which is basically concerned with the application of Information and Communication Technologies (ICT) to the planning and control of transportation systems. It is widely accepted by the transportation community that ITS makes existing transportation facilities more efficient with reasonable costs, minimizing the need to build more infrastructure.

Traffic control in urban roads is a major area in which ITS can be applied. Traffic signals are one of the main approaches to control intersections. They regulate, warn and guide transportation with the purpose of improving the safety and efficiency of pedestrians and vehicles. Among the main advantages of traffic signals are the flexibility of the signaling scheme, the ability to provide priority treatment and the feasibility of coordinated control along streets. Nevertheless, when not well-designed, traffic signals may lead to excessive delays when cycle lengths are too long and increase the risk of collisions.

The behavior of a traffic signal can be considered a Discrete Event System (DES). These systems are often large, complex systems that depends on events that occur at specific instants of time. These events are triggered by specific conditions, such as the ending of a state. Typically, DES involves relationship with the environment, conflict for resources and simultaneous occurrence of activities. When designing DES for critical infrastructures, such as transportation and energy networks, system complexity is increased due to the large number of elements and reliability factors, as these systems involve human life. Therefore, is necessary to use methods that are capable of address complexity and provide high reliable solutions. Formal methods fulfill these requirements by allowing formal proofs of desirable behavior and validation of performance characteristics by simulations. In addition, modular design and abstractions that are important to manage complexity are common to several formal methods.

Petri nets and their extensions [2] are often applied to design, simulate and analyze DES [3]. Some of the advantages are the graphical representation, the mathematical formalism that allows properties analysis and the ability to represent DES characteristics such as concurrency, synchronous and asynchronous communication, resource sharing and time.

There are several examples in the literature of the application of Petri nets to design and analyze traffic signals. Lin et al. [4] described the traffic signal control using Petri nets, and then used an approach via programmable logic controller (PLC) to implement the control logic. Tolba et al. [5] presented the evaluation and performance of different traffic signals by means of Petri nets. Julvez and Boel [6] used continuous Petri nets to obtain realistic and compact models for traffic systems. List et al. [7] used Petri nets to model the control of signalized intersections, and evaluated the good properties of the system by means of invariant analysis and simulation. A new model of urban traffic networks with Petri nets and dioid algebra is applied by Correia et al. in [8]. Febbraro and Giglio [9] applied timed Petri nets with the aim of minimizing congestion situations via a traffic-responsive signaling control procedure. Soares and Vrancken [10] proposed an approach based on
the application of Petri nets and theorem proving to formally analyze the good properties of traffic signals controlling one road intersection. Most of these works are based on predetermined intervals for each color phase, represented in timed Petri nets (fixed time associated to transitions or places). Also, the traffic control is generally proposed only for single intersections.

In this paper, Petri nets with time interval (minimum and maximum) associated to places are applied to the design of a small network of roads with a priority schema. The explicit time representation allows, for instance, to express minimum and maximum green time. Safety and performance rules are considered in the design. The paper shows the representation with Petri nets in levels of increasing detail, and also consider road traffic elements, such as traffic signals and sensors, as part of the system. The idea is to present a responsive model in which green times are extended based on data gathered by sensors. As a matter of fact, priority is given to the main road, and the model has a mechanism to avoid big queues in the non-priority roads. A small subnetwork of two intersections is used as example. The approach can be extended to complete networks, due to its hierarchical and modular design. Finally, formal verification based on Petri nets analysis methods are performed in order to give some results about the reliability of the models.

II. NETWORKS OF TRAFFIC SIGNALS

A road intersection can be defined as the general area where two or more roads join or cross, including the roadway and roadside facilities for traffic management [11]. Traffic signals are an important mechanism applied to solve intersections conflicts and regulate traffic flow. To correctly control the intersection, traffic signal design must take care of safety and security rules, and when possible, of efficiency and speed.

In order to achieve better results, the network level must be evaluated, and problems such as traffic jams solved not only locally, but in a wide area [12]. Otherwise, the risk is not to solve problems, but just move them to another place in the network. Figure 1 shows a network of traffic signals controlling a network of road segments. The network is composed of main roads and arterial roads. Reasoning on a network level increases system complexity. The approach used in this paper is to model and analyze complex systems by hierarchical decomposition into smaller components (divide and conquer) [13]. Thus, the network is divided into subnetworks. Figure 2 shows an example of a subnetwork of roads with two intersections: I1 and I2. The main road has a great flow from B to D, and A and C are non-priority roads.

Some important rules (non-exhaustive) for the proper functioning of traffic signals are [14]:

- the traffic signal must not allow the green state to two conflicting road sections simultaneously;
- each traffic signal must follow a defined sequence of active color lights, normally from green to yellow and red, and then backing to green;
- the right to use an intersection has to be given to all sections.

When reasoning on a network of intersections, three other rules are important:

- any user in the intersection should not wait for more than a maximum service delay, otherwise the user may presume that the traffic signal is not functioning, which can lead to non-secure decisions by the users (red light violations), or the formation of big queues.
- the offset (difference between green initiations times in a network, considering intersections in sequence) is ideally designed in such a way that as the first vehicle just arrives at the next intersection in the network, the signal controlling this intersection turns green.
- the length of green time for each road section can be different, depending on section priority, for instance. This length can be extended to improve traffic flow.

The UML Use Case of figure 3 shows a context diagram for traffic signals controlling a network of intersections and considering the rules above. The use case “Control Phases” depends on actors Vehicle and Sensor. Data detected by sensors (presence and number of vehicles) are sent to the controller. Actor Controller is responsible to select a feasible signalling schema to be sent to actor Traffic Signal. In order to correctly
control the network, safety rules must be applied (include relationship). Whenever possible, which means that safety rules will not be violated, performance rules are considered (extend relationship).

Fig. 3. Traffic Signal Control Use case

Each Use Case is a set of actions performed by the system, which yields an observable result for the involved actors. The dynamic behavior of the Uses Cases are specified with Petri nets in the approach presented in this paper. There are several advantages in using Petri nets in the design of complex DES. The graphical representation is composed of few basic elements, very simple to understand. As a formal tool, Petri nets allow to perform a formal check of desirable properties. The formal analysis of a Petri net model can reveal whether correctness criteria are met. For instance, with reachability analysis it is possible to find out whether an unsafe state that could cause an accident can be reached. Petri nets have mechanisms to provide important aspects for the design of large and complex systems, such as abstraction, hierarchical design and modularity. Transitions and places can represent sub-nets, allowing the modeling in several levels of detail. In addition, it is feasible to construct large models relating smaller Petri nets to each other. An example is the fusion of common places, representing the same resource in two different Petri nets. Place fusion is a simple and effective way to model communication between sub-nets [15].

III. ACTUATED CONTROLLERS

A signal phase is a period during which one or more movements concurrently are shown a green indication. Safety considerations dictate that a phase may be shared only by those traffic streams whose paths do not intersect.

To be effective, advanced signal control systems require an accurate current picture of the traffic flow and status on the roadway network. Actuated control uses information about current demand obtained from sensors within the intersection, to alter one or more aspects of the signal timing [16]. An example of a sensor used in road traffic systems is an inductive loop. Actuated controllers may be programmed to accommodate variable green times or phase sequences. The green is terminated in one of two ways: the maximum green time has been achieved, or it is not possible to extend the green due to a high demand on the other road sections.

Traffic-actuated controllers are equipped to receive data of traffic flow patterns from various measuring devices (sensors) at preset time intervals. This information is used to select one of several timing schemes stored in the controller’s memory. The controller then sends information to traffic signals, which influences traffic flow (Figure 4). Traffic-actuated controllers use some sensing devices to alter the length and/or the sequence of signal indicators. They react to arrivals of individual vehicles rather than to changes in patterns. Timing schemes are usually constrained by specified minimum lengths of green indications for various traffic streams that can be marginally extended by vehicle arrivals up to specified maxima. A sequence of intersections along a road could be connected and controlled as a group.

Fig. 4. Traffic schema

Considering figure 2, there are sensors in the main road (from B to D) and in the arterial roads (A and C). Sensors in A and C have the purpose of detecting and counting vehicles. The idea is to control the number of vehicles that are waiting for a green in such a way that big queues are not formed. Sensors on the main road have the purpose of detecting the first of a platoon of vehicles after a green for B in intersection I1 and allow a green in intersection I2. This coordinated signal control tries to allow that platoons of vehicles can proceed through a continuous series of green lights without stopping.

IV. PETRI NETS DESIGN

The mechanism proposed in this paper is useful in scenarios in which it is necessary to give priority to a main road, and vehicle flow in arterial roads is small. Nevertheless, users in these arterial roads should not think that the traffic signal is not working, what may increase the possibility of accidents. Therefore, every time that there is a queue with a predetermined maximum number of vehicles, the controller is alerted and will not allow any green time extensions anymore. This is important to make crossing possible to non-priority roads.

A. Proposed System Architecture

Figure 5 represents the global architecture. The sensors send data to the controller. Based on these data, safety and performance rules, the controller sends reliable and optimal data to the traffic signals. The dynamic behavior of traffic signals is given by Petri nets with time associated to places, as presented in the next subsections.
B. Global p-time Petri Net Model

The formal definition of a Petri net is given as follows:

**Definition 4.1:** A Petri net is a 5-tuple \( N=(P, T, Pre, Post, M_0) \), where:

- \( P = \{p_1, p_2, ..., p_n\} \) is a finite set of places;
- \( T = \{t_1, t_2, ..., t_m\} \) is a finite set of transitions;
- \( Pre : (P \times T) \rightarrow \mathbb{N} \) is an input function that defines directed arcs from places to transitions, and \( \mathbb{N} \) is the set of nonnegative integers;
- \( Post : (P \times T) \rightarrow \mathbb{N} \) is an output function that defines directed arcs from transitions to places; and
- \( M_0 : P \rightarrow \mathbb{N} \) is the initial marking.

If \( Pre(p, t) = 0 \), then there is no directed arc connecting \( p \) to \( t \). In a similar manner, if \( Post(p, t) = 0 \), there is no directed arc connecting \( t \) to \( p \). If \( Pre(p, t) = k \) (the same holds for \( Post(p, t) = k \)), then there are \( k \) parallel arcs connecting \( p \) to \( t \) (transition \( t \) to place \( p \)). Graphically, parallel arcs are represented by a simple arc with a number as label and representing the presence of multiple arcs (weight \( k \)).

The dynamic behavior (execution) of Petri nets is described by changing markings (from \( M \) to \( M' \)) when an event occurs, which is represented by the firing of an enabled transition. A transition \( t \) is enabled if each place \( P \) that has an input arc to \( t \) contains at least the number of tokens equal to the weight of the arcs connecting \( P \) to \( t \) \((M(p) \geq Pre(p, t) \text{ for any } p \in P)\).

The flow of tokens represent state changing. The firing of an enabled transition \( t \) removes from each input place \( p_i \) the number of tokens equal to the weight of the directed arc connecting \( p_i \) to \( t \), and deposits in each output place \( p_o \) the number of tokens equal to the weight of the directed arc connecting \( t \) to \( p_o \). Therefore, the firing of a transition yields a new marking \( M'(p) = M(p) - Pre(p, t) + Post(p, t) \) for any \( p \in P \).

In order to improve Petri nets applicability, several extensions to the basic theory were proposed. These extensions increase modeling power at the cost of diminishing verification possibilities. In our paper, we use an extension of the Petri net formalism in which time intervals are associated to places in order to formally represent time constraints that are very important in real time systems.

Events that occur in actual systems modeled by a Petri net will possess an event time, which is a value from some discrete or continuous time domain (TD). We may assume that TD is totally ordered and has a minimum element 0 and addition operator +. A possible choice for TD is \( \mathbb{Q}^+ \), the set of nonnegative rational numbers. The p-time Petri net [17] model proposed in this paper adds a time interval to each place, as expressed by the following definition.

**Definition 4.2:** A p-time Petri net is a pair \((N, Ip)\) where \( N \) is a Petri net and \( Ip \) is the application defined as

\[
Ip : P \rightarrow (\mathbb{Q}^+ \cup 0) \times (\mathbb{Q}^+ \cup \infty)
\]

For each place \( p_i \in P \) is associated an interval \( Ip_i = [\alpha_i, \beta_i] \) with \( 0 \leq \alpha_i \leq \beta_i \). The static interval \( Ip_i \) represents the permanency duration of a token in \( p_i \). Before the duration \( \alpha_i \), the token in \( p_i \) is in the non-available state. After \( \alpha_i \) and before \( \beta_i \), the token in \( p_i \) is in the available state for transition firing. After \( \beta_i \), the token in \( p_i \) is again in the non-available state and cannot enable any transition anymore. This token is named “dead token”, and can be seen as a time constraint that was not respected. The dynamic evolution of a p-time Petri net model depends on the marking \( M \) and on the time situation of the tokens (non-available, available, or dead).

Figure 6 represents the global p-time Petri net model of the intersection behavior presented in figure 2. The intersection I1 (place I1) is a shared resource between sections A and B, and the intersection I2 (place I2) is a shared resource between sections C and D. Each state of the traffic signal is represented by a place. Transitions allow state changing, from red to green, green to yellow and back to red for each section. The shaded places RCB, RCD, S1 and S2 are communication places with signals to be sent to road-side equipment.

C. Green time extension schema

The Traffic Detector Handbook [18] recommends a unit extension of 3.5 seconds for roads where speeds are higher than 30 km/h. For the sake of simplicity, it is assumed 4 seconds in this paper and a fixed interval (the same minimal and maximal duration) for each yellow state \((5, 5)\) for instance.

The model of figure 7 is the first level of detail for the green time extension. It has an interval \([45, 57]\) associated to place GB, which corresponds to the minimum and maximum green time allowed.

Sensors detect vehicles and send data to the controller, which will send the decision about the green time extension through the place RCB. The internal behavior of sensors and the controller is specified by Petri nets models. In this paper...
both are abstracted as black boxes receiving data and sending appropriate signals.

Before the end of the minimum green, considering a small amount of time that can be properly defined (preset time intervals), the control will evaluate the presence of vehicles on the minor road. If the queue is smaller than a pre-determined maximum number of vehicles, then the green time can be extended. Otherwise, the green phase ends and the state changes to yellow to allow safety time for vehicle crossing. Green time extension can be given to several small periods until a number of vehicles are waiting in the non-priority section, or until the maximum green time is achieved.

The model of figure 7 is refined into the model of figure 8 to allow a maximum of 3 extensions, each one of up to 4 seconds, for the green time of the main road. The communication places RCB1 and RCB2 (refined from place RCB), represented as shaded places in the model of figure 8, send real-time signals to the traffic signal. Before an extension is allowed, the sensor will send data to the controller, which will process and return one of two responses, as communication places: RCB1, allowing the extension, or RCB2, forcing the traffic signal to switch to a yellow state. If the queue of vehicles reaches its maximum during the \([0, 4]\) extended interval, a response is sent to the communication place RCB2, and the control switches immediately to the yellow state. The advantage of this approach is that instead of just identifying one vehicle and allowing the green for the non-priority road, the idea is to wait for a real demand of vehicles.

The 3 possible extensions are shown in figure 8 (places GB_e1, GB_e2 and GB_e3). Other scenarios, in which more than 3 extensions are allowed, can be modeled and used in a similar manner. This mechanism will allow a minimum of 45 seconds and a maximum of 57 seconds for the green time on the main road.

**D. Offset mechanism**

The difference (normally in seconds) between green initiation times in a network, considering intersections in sequence, is referred to as signal offset [16]. Ideally, the offset is designed in such a way that as the first vehicle just arrives at the next intersection in the network, the signal controlling this intersection turns green. This ideal offset is difficult to achieve when pre-determined offsets are defined. The aim of the offset mechanism considered in this paper is to facilitate that a platoon of vehicles going from section B receive a green when reaching section D, improving traffic flow from intersection I1 to I2.

Coordinated traffic signals [19] are very useful in road networks where traffic signals are located in close proximity. The progressing is given in such a way that the number of stops at intersections are minimized. However, coordination has some limitations. Coordination is more effective if it concerns a platoon of vehicles, although there are no guarantees that all vehicles in a platoon will be given green in the network at the required time. For long sections, the effect may not provide the expected result due to several factors, such as different vehicles speed, traffic already congested and difference in phases of the subsequent traffic signals. These factors may limit the improvement in traffic flow.

Another problem that is often neglected is that the success in coordinating traffic signals in main roads may lead to traffic jams in arterial roads (when vehicles wait too long for their green time), which may move to the main road in the near future.

The communication places S1 and S2 and the sensors in the global Petri net of figure 6 acts in practice as a delay. Figure 9 shows the Petri net model for the offset, considering a fusion of these places and sensors as the place Delay. By fusion of these places, communication between sub-nets takes place. The firing of the transition that allows a green for section D now depends on the place Delay, which represents the presence of a vehicle detected by a sensor requesting the green phase.
B. Behavior Analysis

Reachability analysis can find, through exploratory state space, if all states are reachable, and if a determined state that could lead to a property violation can be reached. Reachability graphs were created for each scenario studied, and in everyone all desirable states were reached. In addition, unsafe states were not reached, and the initial state is reachable again (reversibility). The Petri net is 1-bounded (safe), i.e., resources are not created infinitely. This is an important property in a traffic control system, as resources such as intersections are limited. In practice, an unbounded Petri net would allow the creation of resources in the model, which may lead to simultaneous green time for conflicting intersections. The net is also live, meaning that all transitions can be fired and the controller is able to provide green time, passing to yellow and red to all sections, and that all sections can use the intersection. As a matter of fact, the net is deadlock free.

VI. Conclusion

The coordination of adjacent traffic signals allows that vehicles can move through selected portions of the traffic network with less delay. The paper presented a mechanism based on Petri nets with time intervals associated to places to design a small network of intersections with the purpose of giving priority to a main road and trying to avoid big queues in the other non-priority roads. The purpose is to use real-time data to design traffic-responsive signals with an offset mechanism between two intersections. The approach shows how Petri nets can be applied in a hierarchical and modular way, with different levels of abstraction, which may facilitate the scalability to consider the application for complete networks. In addition, the notion of time intervals associated to Petri nets places is used to represent minimum and maximum durations for enabling transitions. Properties analysis were performed using simulation, invariants and reachability graph, which attest models good properties. Future research will apply the approach in larger and more complicated road networks.

VII. Acknowledgements

This work was supported by the Next Generation Infrastructures Foundation and the Research Center Next Generation Infrastructures, both situated in Delft, The Netherlands.

References


