



HAL
open science

A protocol to evaluate and compare traffic light systems

Jules Bompard, Philippe Mathieu, Antoine Nongaillard

► **To cite this version:**

Jules Bompard, Philippe Mathieu, Antoine Nongaillard. A protocol to evaluate and compare traffic light systems. 23rd Driving Simulation and Virtual Reality Conference & Exhibition, Sep 2024, Strasbourg, France. pp.161-168. hal-04783188

HAL Id: hal-04783188

<https://hal.science/hal-04783188v1>

Submitted on 14 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A protocol to evaluate and compare traffic light systems

Jules Bompard¹, Philippe Mathieu¹, Antoine Nongaillard¹

(1) Univ. Lille, CNRS, Centrale Lille, UMR 9189 CRIStAL, F-59000 Lille, France

Abstract - Urban areas are subject to heavy road traffic, creating congestion that has a negative impact on the economy and greenhouse gas emissions. To alleviate this problem, traffic light systems (TLS) have been developed. Existing evaluations of these systems suffer from several biases, reducing the credibility of the results. In this paper, we develop a stochastic protocol for evaluating TLS. The protocol consists of simulating each system a large number of times on a single road network, randomly varying vehicle flows and system parameters. On the basis of average travel times and CO2 emissions per travel, we develop two new metrics for analyzing the potential and the difficulty of setting of the systems. Based on an example of the use of this protocol, we demonstrate that the use of stochastic simulations makes it possible to obtain objective results, unlike existing protocols. We believe that this test method will help road infrastructure managers to select a traffic light management system with greater confidence.

Keywords: Evaluation, Strategies, Simulation, Road traffic

Introduction

The expansion and densification of urban areas has led to exponential growth in the transport sector, mainly road transport, in cities. Well-functioning road networks are associated with better economic outcomes (Graham, 2007), but traffic congestion can make them counter-productive. In economic terms, traffic congestion lead to financial losses, as in the European Union, where the costs are estimated at 1% of GDP each year (Mačiulis, Vasiliauskas, and Jakubauskas, 2009). Road traffic congestion also leads to an increase in greenhouse gas emissions (Bharadwaj, et al., 2017), as well as toxic gases for humans, which increase the risk of respiratory and cardiovascular diseases, and pregnancy and child-birth problems (Krzyzanowski, Kuna-Dibbert, and Schneider, 2005). It is therefore necessary to reduce the road traffic congestion in order to limit these negative effects.

Two solutions have been envisaged to achieve this: building new infrastructures to distribute the vehicle load more evenly, or optimizing existing infrastructures. The former involves significant ecological and economic costs, so research has focused on the latter. Intersections are one of the infrastructures that need to be optimized. In urban areas, they are usually managed using non-adaptive traffic lights, with a fixed time allocated to each phase. This management method is not the most efficient, and a number of recent papers present traffic light systems (TLS) which aim to optimize traffic light management (Chen, et al., 2020; Cools, Gershenson, and D'Hooghe, 2013; Varaiya, 2013; Wei, et al., 2018).

While many of these systems show interesting results for a specific road network topology and traffic flows, it is difficult to generalize their performance to every possible situations. Most of the time, these sys-

tems are evaluated on the basis of a dozen or so simulations, i.e. when flows evolve, for a single network. These experimental conditions do not allow the results to be generalized on any scale whatsoever, and sometimes lead to conflicting results, as Zhang, et al. (2022) shows that the Max Pressure system (Varaiya, 2013) is better than the MPLight system, whereas Chen, et al. (2020) shows the opposite. In addition, the parameters of the various systems studied are rarely transmitted, which makes it impossible to reproduce the simulations. In order to generalize the results of TLS evaluations, it is necessary to develop a test protocol that takes these specific features into account.

In this article, we present a stochastic protocol for the evaluation of TLS. This protocol consists of simulating a certain number of times a system on a given network typology with randomly selected flow and system parameters, in order to estimate the strengths and weaknesses of each one in every condition. The random values are selected in ranges of values based on the real traffic measurements, ensuring that each simulation is representative of real traffic conditions. This way, our protocol have two properties that differentiate it from previous protocols: 1) the results of the experiments are generalizable to every possible traffic conditions for a given network, and 2) the sets of possible values for the system and flow parameters can be easily shared to ensure the reproducibility of the experiments.

Related work

Traffic light systems (TLS)

As explained previously, a number of TLS systems have been proposed to overcome the shortcomings

of fixed-time signals (Koonce and Rodegerdts, 2008). These systems use sensors placed at the intersection to collect and use data from their local environment, and select a traffic light phase, i.e. a red or green light assignment for each lane. They can be classified into three categories: rules-based systems, optimization-based systems and RL-based systems. We briefly explain how the systems in each category work, as it impacts the comparison.

Rules-based systems

Rules-based systems manage an intersection by following a set of rules designed to ensure near-optimal behavior. Among them, the Max Pressure system (Varaiya, 2013) calculates at each time interval the pressure value of each of the intersection's approach lanes, i.e. the number of vehicles on these lanes minus the number of vehicles on the intersection's exit lanes, and selects the phase that gives the green light to the lanes that maximize it. The self-organizing traffic lights (SOTL) system (Cools, Gershenson, and D'Hooghe, 2013) selects a phase that gives the green light to lanes whose number of waiting vehicles is greater than a certain threshold and keeps this phase as long as vehicles are detected during a defined time interval. Other rules-based systems are worth noting, such as the Longest Queue First system of Wunderlich, et al. (2008), which selects the phase whose lights are green for the lanes with the longest tail. These systems are easy to test. As they do not need any training, the data collected in simulation are all representative of their full efficiency, and can be used to compare the systems with other ones.

Optimization-based systems

Optimization-based systems use optimization algorithms to select the best phase from a set of possible ones. Several optimization paradigms have been used to develop TLS. Some systems use fuzzy logic to create a more complex representation of the local environment of an intersection, and use solvers to select the best phase (Azimirad, Pariz, and Sistani, 2010; Collotta, Bello, and Pau, 2015; Mohanaselvi and Shanpriya, 2019). Simulated annealing (Oda, et al., 1997) and evolutionary algorithms (Ceylan and Bell, 2004; Teklu, Sumalee, and Watling, 2007) have also been studied. However, they seem to have been discarded because they are limited by the need for an evaluation function, which can only be obtained by simulation, and then cannot be truly faithful to reality. Nevertheless, a few recent papers have developed systems of this type (Amer, et al., 2016; Mao, et al., 2021) and must be taken into account. Optimization-based systems are as easy to test as the rules-based systems, as they also do not require any training before deployment.

RL-based systems

Systems based on reinforcement learning (RL-based) are systems that learn through continuous experience, by optimizing a reward function that depends on their choices. To date, this is the dominant paradigm for TLS design, with many systems having been developed to improve the representation of the local environment, the calculation of the reward function or the learning method (Chen, et al., 2020; Pol and Oliehoek, 2016; Wei, et al., 2018; Zhang, et al., 2022). Unlike rules-based or optimization-based

systems, RL-based systems require training before deployment. Before testing, several simulations have to be done in order to adjust the policy of each traffic light, making them operate at a suboptimal level. Then, measurements must not be taken during the training phase, but after deployment. An experimental protocol must take this criterion into account to fairly evaluate RL-based, when its policy is supposed to be optimal.

TLS test protocols

When a new TLS is developed, it is customary to test it in simulation in order to measure its performance. To do this, we simulate a road network, either fictitious or real, in which the TLS is implemented, and we generate vehicle flows whose quantity can be arbitrary or based on real-world measures. The use of simulation is essential: to test a dysfunctional system in the real world could generate traffic disruptions that would exacerbate the economic and environmental problems mentioned above. In addition, new traffic simulation software makes it easy to measure metrics that are used to assess the implemented system. We have identified two distinct types of test protocols: validation protocols and comparison protocols. Validation protocols only concern optimization-based and RL-based systems. Their objective is to demonstrate and measure the optimization capabilities of the system, generally by measuring its learning speed or convergence towards the optimal solution (Mao, et al., 2021; Pol and Oliehoek, 2016). Comparison protocols are used to evaluate and rank a set of TLS, thereby proving that one system is superior to the others. The set of TLS is simulated on the same road network and compared using metrics relating to the network as a whole, such as average travel time or average waiting time. They represent the biggest part of the test protocols that are published in recent years, researchers trying to prove that their system is the best of the state of the art.

Identified problems of comparison protocols

Although comparison protocols are the most used to prove the performances of a TLS, we identified several issues that led us to think that their results can be biased or partial. It is therefore not possible to generalize the results of this type of test, and comparative protocols tend to prove the performance of a TLS only under precise conditions. This section lists every issue we identified, and explain why they bias the results of comparison protocols.

Over-optimal networks

If a network is too simple, i.e. its characteristics do not include the most important road traffic fluidity issues, the results provided by an experimental protocol cannot be generalized. Some papers evaluate systems on a network composed of a single intersection (Abdulhai, Pringle, and Karakoulas, 2003; Teo, Kow, and Chin, 2010; Wunderlich, et al., 2008). The complexity of the traffic light optimization problem lies in the fact that two neighboring traffic lights may have conflicting interests. An optimized traffic light may completely block a neighboring traffic light, which would reduce overall performance. A more complete topology is

therefore needed to optimally test a TLS. Recent papers have also evaluated their TLS on a network whose roads have a special lane to turn left (Chen, et al., 2020; Huang, et al., 2021; Wei, et al., 2018). This topology is optimal because a vehicle waiting to turn left cannot slow down or stop the others. As many road networks has only one lane per approach, no generalization can be made from a comparison protocol carried out on 2-lane intersections or more.

Too many metrics

Another problem with TLS comparison protocols concerns the evaluation metrics. Numerous metrics have been used to date: average travel time (Zang, et al., 2020), throughput (Chen, et al., 2020), waiting time (Cools, Gershenson, and D'Hooghe, 2013), delay (Wei, et al., 2018), and even arbitrary performance indexes (Ceylan and Bell, 2004). This multitude of metrics makes it difficult to compare different studies. Indeed, two papers developing protocols evaluating the same set of TLS with different metrics may obtain different results, accentuating the uncertainty surrounding the choice of the best system. In addition, some studies do not compare their system with the same set of TLS. Some only compare their system with a fixed-time system (Abdulhai, Pringle, and Karakoulas, 2003; Azimirad, Pariz, and Sistani, 2010) or with a limited number of other systems (Cools, Gershenson, and D'Hooghe, 2013; Pol and Oliehoek, 2016), which is not enough to validate the efficiency of a TLS, especially for the fixed-time strategy, which is known to be weak.

Not enough simulations

The number of simulations for some protocols is also problematic. All the papers cited in this article develop protocols based on a small number of simulations, i.e. less than ten on the same network with different vehicle flows. This number raises questions about the generalizability of the results, as only a small number of situations, however realistic, are tested. However, a TLS requires a minimum of flexibility, i.e. that it is not efficient only under certain conditions, but under all the realistic conditions. For example, recent studies have tested their system on the networks of New York, Hangzhou and Jinan, with flows measured in the field, but there is no guarantee that their results are the same in situations other than those tested (Zhang, et al., 2022; Zheng, et al., 2019). In order to test traffic light systems, further simulations are needed to represent the full range of realistic conditions.

TLS settings not supplied

Finally, the major problem with comparison protocols is that the parameters of the systems being tested are not transmitted. A poorly parameterized TLS gives very poor results, when it could potentially give much better performances. So if all the TLS are not configured to their maximum potential, the results observed can be misleading. This shortcoming is related to the small number of simulations mentioned above, which accentuates the phenomenon. The transmission of the parameters of each of the TLS tested is therefore essential to better assess the quality of the results. While the parameters of

the system developed in one paper are logically provided, none of the papers cited in this article provide those of the comparison TLS.

Nowadays, a consensus has emerged around the protocol developed by Zheng, et al. (2019), which has now become a standard (Chen, et al., 2020; Goel, et al., 2023; Zhang, et al., 2022). TLS is tested on real megacities networks and flows, measuring average travel time and vehicle throughput. However, this protocol only involves a small number of simulations, and the parameters of other TLS are never provided. This leads to inconsistencies in the results of some papers, such as this paper which proposes a system that reduces travel time by 95% compared with a fixed-time strategy (Jiang, et al., 2022), or even conflicting results, such as Chen, et al. (2020) and Zhang, et al. (2022) who rank the Max Pressure and the MPLight systems differently.

Methodology

Determining the best TLS globally seems to us to be an impossible task. This is because not all road networks have the same constraints for the traffic flowing on them (3 lanes per road, one-way traffic, etc.), and a TLS that would perform well for a given network could perform poorly if subjected to different constraints. It might be thought that TLS should then be tested on a set of networks that would include all the constraints, but the computational complexity would be such that the experimental protocol would not be feasible. The protocol we develop in this article aims, rather than to determine the best global TLS, to determine the TLS for a given network, i.e. one that minimizes the mean of the metrics collected under different conditions.

Main idea

As the number of simulation and TLS settings, even for a single network, is exponential with the number of parameters, it is impossible to test all possible configurations. However, we have seen that it is necessary to simulate TLS in all possible traffic situations to get an idea of how their performance evolves as a function of conditions. To overcome this problem, we use a stochastic protocol for evaluating TLS: we run n simulations with randomly selected parameters, for each TLS. This choice has two advantages: it limits the number of simulations needed to evaluate a system, and guarantees that its performance are evaluated over a large set of situations, depending on n . Logically, the larger n is, the more reliable the results are, and depend on the size of the exploration space.

Adaptation to RL-based systems

We have identified three types of TLS, including RL-based systems, which operate differently from the others in that they must first be trained before they are functional. This characteristic must be taken into account: the performance of a RL-based system must not be measured during training. In this protocol, RL-based systems are first trained on the same network as the evaluation network, for a fixed duration, and with flows selected in the same intervals and according to the same probability distribution as for the evaluation. Training parameters, such as the

duration of an episode or the number of episodes, must be provided for the reproducibility of the protocol.

Parameter selection

If the parameters of a simulation are randomly selected, it is necessary to define ranges of selectable values for each of them. This is the trickiest part of the protocol, because a range of values must be chosen that is large enough to allow a wide variety of parameters to be tested, but small enough for the empirical results of the n simulations to be representative. Small ranges can only be defined with prior knowledge of how a system operates. If we have no idea how it should be parameterized, a large range must be defined for each parameter to ensure that the optimum calibration is included in the possible ones. In all cases, this range of values has to be shared so that the protocol can be reproduced, and the results analyzed. One of the most important parameters concerns vehicle generation. We believe that only realistic vehicle flows should be simulated. Indeed, if unrealistic traffic scenarios are included, the exploration space is enlarged, which would require more simulations than necessary. If they are known, the vehicle flows should follow the same proportions as those of the network's real routes, with exploration consisting of varying the intensity of these flows.

Metrics

For each simulation, we keep the average travel time and average CO2 emissions per travel. These metrics are the most relevant, as they are directly linked to the ecological and economic issues mentioned in the introduction. The CO2 emissions of a network per unit of time cannot be used to evaluate the fluidity of a network: vehicles emit less CO2 when they are stationary than when they are moving. Therefore, a traffic jam emit less CO2 per unit of time than moving traffic. Then, we choose the average CO2 emissions per journey, which penalizes traffic jams. The data collected can be plotted in the form of a cloud of points, the abscissa of which is the total load on the network, i.e. the total number of vehicles travelling on the network per hour. This representation provides a visual indication of which TLS works best under different traffic conditions. However, in order to be able to rank the systems, this data needs to be aggregated to form a single value. Conventional forms of aggregation, such as averaging, cannot be used, as all TLS have been tested for random, and therefore different, traffic conditions. Thus, we have developed two new data aggregation methods: potential and variance.

Potential

The potential represents the average performance of a near-optimally tuned TLS. It is calculated by selecting at each load interval the best simulation results, i.e. the lowest average travel time and CO2 emissions, and averaging it. More formally, the potential is described by Eq.1, where m is the maximum possible load value, and k is the selection interval. Each load interval can be seen as a different traffic profile (light, medium, dense, jam), and averaging the best results from these profiles is equivalent to calculating the empirical best performance of a TLS.

$$P = \frac{1}{n/k} \sum_{i=0}^{(n/k)-1} \operatorname{argmin}(ik, ik + k) \quad (1)$$

Variance

The variance represents the average difference in performance between a bad setting and a good setting for a TLS. A well-tuned system may perform very well for a given traffic profile, but perform poorly with another setting. The variance quantifies the average consequence of a poor setting. For each interval of loads, we compute the absolute difference between the simulation results and their average, and then, compute the average of these values. Formally, the variance is defined by Eq.2, where R represents the set of simulation results.

$$D = \frac{1}{n/k} \sum_{i=0}^{(n/k)-1} \left(\frac{1}{\#R} \sum_{r \in R[ik, ik+k]} |r - \bar{R}| \right) \quad (2)$$

The potential and the variance can easily be shared in a table, in order to rank the systems tested. The system that has the best potential, i.e. the lowest value, is possibly the best for a defined network, because it has the capacity to minimize the metrics the better, if it is well calibrated. On the other hand, the variance can mitigate the evaluation of a system with a good potential. If the potential is hard to reach, due to a large number of possible configurations or extreme variations of flows on the network, the consequences are worse if the variance is high. Thus, it is clear that the system with the best potential and the best variance is a best systems than the others tested.

Parameter transmission

One of the aims of this protocol is to facilitate the transmission of simulation parameters and systems under test. With existing protocols, the transmission of parameters requires the transmission of a set of parameters for each simulation, which is not possible. Since we generate simulations randomly from intervals of values for each parameter, it is now sufficient to transmit only these intervals, and the selection probability distribution, in the form of multiple tables for example.

Experiments

To illustrate this new protocol, we carry out a series of experiments in order to compare five different traffic light systems. We simulate traffic flows on an arbitrary network and scenarios. Therefore, the results are not relevant for any real application : we don't try to prove that a system is better than an other on real conditions. The results here only proves the superiority of a system on an unreal network, i.e. a 3x3 grid network. The experiences are conducted on the SUMO simulator (Behrisch, et al., 2011), and more precisely with the python module sumo-experiments[†]. Cityflow (Zhang, et al., 2019) could also have been

[†] <https://github.com/cristal-smac/sumo-experiments>

chosen but the project has not been maintained for years.

Simulation parameters

For this experiment, each of the simulation last 3600 simulation steps, which is equivalent to an hour, and is made 500 times for each traffic light system. This value has been chosen to be high enough to test a big part of possible flows and system calibrations, but low enough to do all simulations in a single week on a personal computer. Network and flows parameters are described below. A summary table of the simulation parameters is provided at the end of the section with Table 1.

Network

The simulations are carried out on 3x3 grid network, where each traffic light is 200 meters away from neighbouring traffic lights. All roads has one lane for each direction. The network has nine entries and nine exits, located on the extremities. Traffic lights only have two phases : one for the west/east (WE) direction and one for the north/south (NS) one. The yellow phase last 3 seconds after each phase. For each intersections, we implement two types of detectors, on each incoming lane : boolean detectors and numerical detectors. The boolean detectors have a range of 20 meters and return the presence of a vehicle as a boolean value. The numerical detectors have a range equal to the lane length, which is 200 meters for each lane here, and returns the number of detected vehicles. We only simulate SUMO default car.

Flows

Multiple parameter ranges must be defined to generate traffic over the network. This ranges are defined following the supposedly real traffic flow that we imagine for this network. The network load, i.e. the number of vehicles running, varies from 100 to 5000 vehicles per hour. Flows are unbalanced between the WE entries and the NS entries, with WE flows always bigger than the others. To represent this in simulation, we define an unbalance parameter, ranging from 65 to 80%, that represents the part of the load entering the network by WE entries. The proportional load is then equally distribute over entries of the same direction. Possible vehicle routes are all the entry/exit couples, distributed uniformly across all network load. The unbalance and the load are selected with a uniform distribution over their range. To compute the potential and the variance scores, we select an interval of load k of 500 vehicles, which makes 10 slices of data, of approximately 50 simulations for each TLS.

Traffic light systems

In this experiment, we compare five different traffic light systems, including two rules-based systems and two RL-based systems. All of them are multi-agent systems, meaning that each intersection is managed by an agent, acting independently from the others. We briefly describe each of them, and report their parameters in tables.

Table 1: The experiment simulation parameters

Parameter	Value
Number of simulations n	500
Simulation duration	3600 seconds
Road network	Grid 3x3 - 200m between TL
Numerical detectors range	200 meters
Boolean detectors range	20 meters
Load range	100 - 5000 vehicles/hour
Unbalance	65 - 80 %
Yellow time	3 seconds
Load interval k	500 vehicles/hour
Parameter distribution law	Uniform

Fixed Time system

The fixed time system (Koonce and Rodegerdts, 2008) is the most used system worldwide. Each intersection has an ordered set of phases, and each phase has a duration. When a phase has lasted the agreed time, the intersection moves on to the next phase. A duration selected from a range from 10 to 120 seconds is attributed to each phase. Each traffic light is set with the same settings. The duration of each phase is the only parameter required for this system, and is summarized in Table 2.

Table 2: Parameters of the fixed time system

Parameter	Min	Max	Step	Unit
Phase duration	10	120	1	seconds

Max pressure system

The max pressure system is a rules-based system developed by Varaiya (2013). It defines the concept of pressure for a phase, which is the difference between the number of vehicles moving to the intersection, and the number of vehicle moving to a neighbouring intersection, coming from the intersection. At each period t , the pressure of each phase is computed, and the agent selects the phase with the highest one for the next period. The same phase can be selected multiple times in a row. In simulation, we use the numerical detectors of each intersection to compute the number of incoming vehicles, and compute the pressure, and we variate the period duration from 10 to 120 seconds. This is the only parameter required for this system, summarized in Table 3.

Table 3: Parameters of the max pressure system

Parameter	Min	Max	Step	Unit
Period duration	10	120	1	seconds

Self organizing traffic lights (SOTL)

Like the max pressure system, the SOTL system is a rules-based system, developed by Cools, Gershenson, and D'Hooghe (2013). The agent counts and sums the number of incoming and waiting vehicles on red lanes, i.e. lanes with red traffic lights, at each time interval, here a second. When the sum of vehicles is greater than a threshold x_1 , the traffic light changes its phase, unless there still are vehicle that are detected on green lanes, the lanes with green lights. To prevent the phases taking too long when traffic jams build up, the traffic light also changes phase when the

number of waiting vehicles alone exceeds a threshold x_2 . Like previously, we use SUMO detectors to detect incoming vehicles on each lane. A vehicle outside of the detectors range are not considered. Also, to prevent the system from blinking, when the phases switch to quickly during jams on all lanes, we add a minimum phase time t . The parameters' ranges we used on this experiments are summarized in Table 4.

Table 4: Parameters of the SOTL system

Parameter	Min	Max	Step	Unit
Threshold switch x_1	120	600	1	vehicles
Threshold force x_2	2	x_1	1	vehicles
Minimum phase time t	10	120	1	seconds

Reinforcement learning systems

To test the protocol with systems that require training, we developed two simple RL-based systems, inspired from existing ones. The two systems use a double deep Q-network algorithm to compute the value of a couple state/action, implementing two two-layer network with respectively 128 and 56 relu-activated units. The target network is updated every f training episodes. The action selection is performed by a ϵ -greedy algorithm, where ϵ decrease at each selection by a multiplication with an updater k , in the range 0 to 1, and has a minimum value $\text{min-}\epsilon$ also between 0 and 1. For this experiment, the training phase last six hours, divided in episodes of duration δ , where data are collected every seconds. The value of ϵ start to decrease only when the training phase reach the d -nth episode. During the weights update of the neural network, we use a batch of size s for a number of epochs e , and a discount factor γ , between 0 and 1, for the Bellman (1957) equations. The two systems are evaluated only after the training phase.

The first system (RL1) is inspired from the Intellilight system from Wei, et al. (2018). The state of a traffic light is represented by a vector containing the queue lengths, the vehicle numbers and the total waiting times on each lane, and the current phase. The algorithm performs an action every second : either keep the current phase, or move on to the next. The reward of the algorithm is the mean delay, i.e. the difference between the vehicles speed and the maximum speed of a lane, multiplied by the sum the numbers of vehicles on each lane.

The second system (RL2) is inspired from the systems of Wei, et al. (2018) and Chen, et al. (2020). The state representation of the traffic light is the same as the RL1 system, but the action space is different. Every time interval t , the algorithm selects a possible phase for the traffic light, which may be the same as the previous one. The reward to minimize is the pressure of the intersection.

All of the parameters ranges are summarized in Table 5.

Results and discussions

Ranking analysis

The results of the experiment are plotted on Fig 1. Multiple assertions can be made from this first visual analysis. Although the max pressure has the best

Table 5: Parameters of the RL systems

Parameter	Min	Max	Step	Unit
ϵ starting value	1	1	none	none
ϵ -updater k	0.9	0.999	0.001	none
$\text{min-}\epsilon$	0	0.1	0.01	none
Decreasing episode d	20	100	1	none
Episode duration δ	120	600	1	seconds
Batch size s	20	300	1	data
Number of epochs e	20	200	1	epochs
Discount factor γ	0.7	0.99	0.01	none
Target network update f	1	10	1	episodes

mean performances, the RL1 system produce the best results overall, with low traffic flows, but more importantly with very high traffic flows. This phenomenon is the exact reason of why we introduced the notion of potential. If we only computed the mean performances of all systems, RL1 system would have a low evaluation, despite the very good performances with some precise settings. The potential highlights these high performances, even if the majority of the results are bad. As a road network only needs one TLS with one setting, the potential seems to be fairer than the average result. But the disparity in results must not be totally ignored. The reinforcement learning systems and the fixed time strategy have scattered results, which means that their setting has a great impact on their performances. This contrasts with the max pressure system, for which the results are much more compact, and shows that this system is more resilient to variations of flows or settings. The variance highlights this heterogeneity in the results. Potential and variance of each system are shown in Tab 6 in terms of mean travel time and in Tab 7 in terms of mean CO2 emissions per travel.

Table 6: Potential and variance for the TLS, depending on the mean travel time

System	Potential	Variance
Fixed time	369.22	983.73
Max pressure	137.95	363.58
SOTL	193.6	560.66
RL1	85.53	2533.62
RL2	128.58	1221.5

Table 7: Potential and variance for the TLS, depending on the mean CO2 emissions per travel

System	Potential	Variance
Fixed time	0.95	2.54
Max pressure	0.38	0.84
SOTL	0.50	1.35
RL1	0.25	3.94
RL2	0.36	2.56

The interpretation that can be made from the results of this testing protocol is that the RL1 is the best system for a 3x3 network with this type of asymmetric flows, with a potential mean travel time of 85 seconds. However, the system must be very well tuned to be efficient, otherwise the performance can be very poor.

Objective nature of the results

The high number of simulations guarantees a better objectivity of the results than the existing proto-

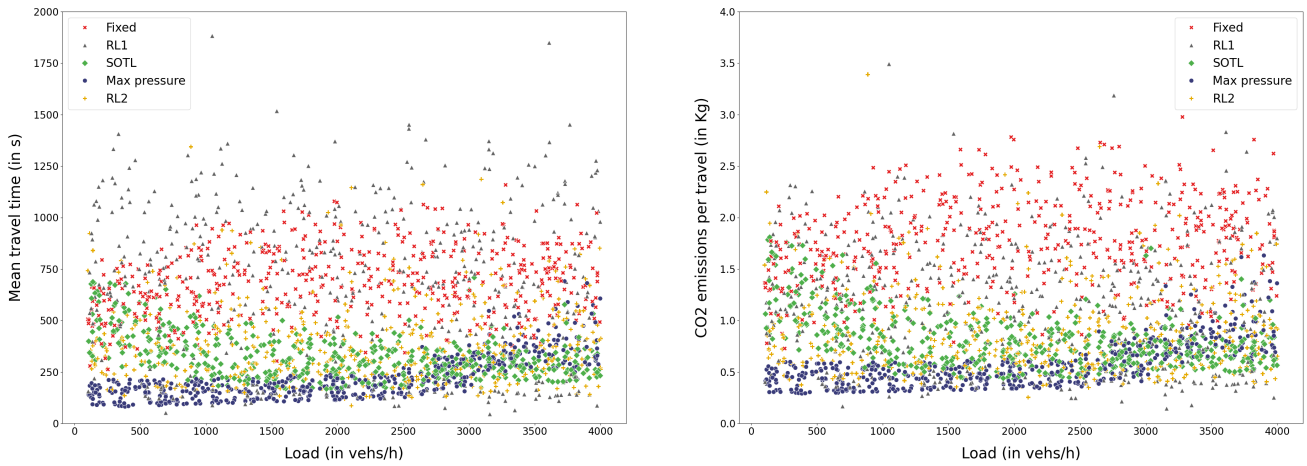


Figure 1: Mean travel time and mean CO2 emissions per travel of the simulations, as a function of the load and the used TLS

cols. In previous protocols, only a few simulations were made, and variations of the performance over the flows or the system settings couldn't be studied. This led to biased results, as shown in the identified problems of existing comparison protocols. We can reproduce this biased with the data in Fig 1. If we select just one form of traffic profile, we can obtain any TLS ranking that we want, as shown in Tab 8.

Table 8: Rankings of the different TLS for different traffic loads, in vehicles per hour

Ranking	1768 vehs/h of load	2422	2883	2907
1st (best)	Max pressure (MP)	RL1	SOTL	MP
2nd	RL2	MP	MP	SOTL
3rd	SOTL	RL2	RL1	RL2
4th	RL1	SOTL	RL2	Fixed
5th (worst)	Fixed	Fixed	Fixed	RL1

This contradiction in results proves that existing protocols are not sufficiently relevant for assessing traffic light systems. However, our protocol is not subject to this issue. A single ranking can be produced, and although randomness prevents us from defining it as deterministic, this protocol almost always returns the same ranking for the same set of parameters, provided that n is large.

Parameter influence

Although the results of this protocol are considered to be objective, we must not overlook the impact of certain parameters which can have major consequences on the results of the experiments.

Number of simulations

The number of simulations has a large influence on the final system ranking. As with the existing comparison protocols, a low number of simulations leads to irrelevant results, as any result can be made from individual results. To ensure the objectivity of the protocol, a high number of simulations must be performed, with the idea that the higher the better. The minimum number of simulations depends on the exploration space, i.e. all combinations of flow rates and system settings. In our experience, 100 simulations per TLS

is the bare minimum to ensure the objectivity of the results.

Interval of load

The load interval defines the segmentation of the results space, which is an essential parameter for calculating the potential and the variance. Each interval represents a traffic profile, and categorise the various possible loads on the network. This parameter must therefore be chosen with care. A low value result in a thin segmentation of simulation results. A result slice that is too thin may not contain enough data to objectively calculate the metrics, as is the case with existing protocols. On the other hand, a result slice that is too large does not define precisely the different traffic profiles, and leads to a loss of information. Therefore, the interval of load must be select with a prior knowledge of the network traffic profiles, to ensure an optimal precision of the potential and the variance.

Parameters of the systems

The definition of the ranges of possible values for each system parameter is also a crucial point of the protocol. It defines the possible parameter combinations that are simulated and studied. Among all the configurations, it must find the best possible ones. If this is not the case, the real potential of a system cannot be assessed. If there is uncertainty about the direction in which systems should be parameterised, it is preferable to include as many configurations as possible to avoid the risk of bias.

Conclusion

In this paper, we present a protocol to evaluate and rank traffic light systems. The advantage of this protocol is that it does not suffer from the biases of existing protocols. By introducing two new metrics, potential and variance, as well as a stochastic simulation concept, we are able to evaluate traffic light management systems quantitatively and objectively. We also acknowledge the limitations of our current approach, as the evaluation depends on multiple parameters, and we give some guidelines to set them correctly. In any case, as this protocol makes it easier to transmit each

of the selected parameter, results can be accurately interpreted, and any reviewer can detect anomalies in the test method. Two types of future experiments will be able to implement this work. As part of the development of a new TLS, this protocol can be used to better understand the evolution of performance as a function of its parameters, or changes in traffic flows. On a more technical level, this protocol could be used by road network managers to select the system that works best for their specific network. We believe that this protocol has the potential to become one of the standard evaluation protocols in the future.

References

- Abdulhai, B., Pringle, R., and Karakoulas, G. J., 2003. Reinforcement learning for true adaptive traffic signal control. *Journal of Transportation Engineering*, 129(3), pp. 278–285.
- Amer, H., Salman, N., Hawes, M., Chaqfeh, M., Mihaylova, L., and Mayfield, M., 2016. An improved simulated annealing technique for enhanced mobility in smart cities. *Sensors*, 16(7), p. 1013.
- Azimirad, E., Pariz, N., and Sistani, M. B. N., 2010. A novel fuzzy model and control of single intersection at urban traffic network. *IEEE Systems Journal*, 4(1), pp. 107–111.
- Behrisch, M., Bieker, L., Erdmann, J., and Krajzewicz, D., 2011. SUMO - simulation of urban mobility: an overview. In: *Proceedings of SIMUL 2011, The Third International Conference on Advances in System Simulation*. ThinkMind.
- Bellman, R., 1957. Dynamic programming. *Press Princeton, New Jersey*, 39.
- Bharadwaj, S., Ballare, S., Rohit, and Chandel, M. K., 2017. Impact of congestion on greenhouse gas emissions for road transport in Mumbai metropolitan region. *Transportation Research Procedia*, 25. World Conference on Transport Research - WCTR 2016 Shanghai. 10-15 July 2016, pp. 3538–3551. ISSN: 2352-1465. <https://doi.org/https://doi.org/10.1016/j.trpro.2017.05.282>.
- Ceylan, H. and Bell, M. G., 2004. Traffic signal timing optimisation based on genetic algorithm approach, including driversâ routing. *Transportation Research Part B: Methodological*, 38(4), pp. 329–342.
- Chen, C., Wei, H., Xu, N., Zheng, G., Yang, M., Xiong, Y., Xu, K., and Li, Z., 2020. Toward a thousand lights: Decentralized deep reinforcement learning for large-scale traffic signal control. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 34. 04, pp. 3414–3421.
- Collotta, M., Bello, L. L., and Pau, G., 2015. A novel approach for dynamic traffic lights management based on Wireless Sensor Networks and multiple fuzzy logic controllers. *Expert Systems with Applications*, 42(13), pp. 5403–5415.
- Cools, S.-B., Gershenson, C., and D'Hooghe, B., 2013. Self-organizing traffic lights: A realistic simulation. *Advances in applied self-organizing systems*, pp. 45–55.
- Goel, H., Zhang, Y., Damani, M., and Sartoretto, G., 2023. Social-Light: Distributed Cooperation Learning towards Network-Wide Traffic Signal Control. *arXiv preprint arXiv:2305.16145*.
- Graham, D. J., 2007. Variable returns to agglomeration and the effect of road traffic congestion. *Journal of Urban Economics*, 62(1), pp. 103–120.
- Huang, H., Hu, Z., Lu, Z., and Wen, X., 2021. Network-scale traffic signal control via multiagent reinforcement learning with deep spatiotemporal attentive network. *IEEE transactions on cybernetics*.
- Jiang, Q., Qin, M., Shi, S., Sun, W., and Zheng, B., 2022. Multi-agent reinforcement learning for traffic signal control through universal communication method. *arXiv preprint arXiv:2204.12190*.
- Koonce, P. and Rodegerdts, L., 2008. *Traffic signal timing manual*. Tech. rep. United States. Federal Highway Administration.
- Krzyzanowski, M., Kuna-Dibbert, B., and Schneider, J., 2005. *Health effects of transport-related air pollution*. WHO Regional Office Europe.
- Mačiulis, A., Vasiliauskas, A. V., and Jakubauskas, G., 2009. The impact of transport on the competitiveness of national economy. *Transport*, 24(2), pp. 93–99.
- Mao, T., Mihäitä, A.-S., Chen, F., and Vu, H. L., 2021. Boosted genetic algorithm using machine learning for traffic control optimization. *IEEE Transactions on Intelligent Transportation Systems*, 23(7), pp. 7112–7141.
- Mohanaseelvi, S. and Shanpriya, B., 2019. Application of fuzzy logic to control traffic signals. In: *AIP Conference Proceedings*. Vol. 2112. 1. AIP Publishing.
- Oda, T., Otokita, T., Tsugui, T., and Mashiyama, Y., 1997. Application of simulated annealing to optimization of traffic signal timings. *IFAC Proceedings Volumes*, 30(8), pp. 733–736.
- Pol, E. Van der and Oliehoek, F. A., 2016. Coordinated deep reinforcement learners for traffic light control. *Proceedings of learning, inference and control of multi-agent systems (at NIPS 2016)*, 8, pp. 21–38.
- Teklu, F., Sumalee, A., and Watling, D., 2007. A genetic algorithm approach for optimizing traffic control signals considering routing. *Computer-Aided Civil and Infrastructure Engineering*, 22(1), pp. 31–43.
- Teo, K. T. K., Kow, W. Y., and Chin, Y., 2010. Optimization of traffic flow within an urban traffic light intersection with genetic algorithm. In: *2010 Second International Conference on Computational Intelligence, Modelling and Simulation*. IEEE, pp. 172–177.
- Varaiya, P., 2013. The max-pressure controller for arbitrary networks of signalized intersections. In: *Advances in dynamic network modeling in complex transportation systems*. Springer, pp. 27–66.
- Wei, H., Zheng, G., Yao, H., and Li, Z., 2018. Intellilight: A reinforcement learning approach for intelligent traffic light control. In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 2496–2505.
- Wunderlich, R., Liu, C., Elhanany, I., and Urbanik, T., 2008. A novel signal-scheduling algorithm with quality-of-service provisioning for an isolated intersection. *IEEE Transactions on intelligent transportation systems*, 9(3), pp. 536–547.
- Zang, X., Yao, H., Zheng, G., Xu, N., Xu, K., and Li, Z., 2020. Meta-light: Value-based meta-reinforcement learning for traffic signal control. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 34. 01, pp. 1153–1160.
- Zhang, H., Feng, S., Liu, C., Ding, Y., Zhu, Y., Zhou, Z., Zhang, W., Yu, Y., Jin, H., and Li, Z., 2019. Cityflow: A multi-agent reinforcement learning environment for large scale city traffic scenario. In: *The world wide web conference*, pp. 3620–3624.
- Zhang, L., Wu, Q., Shen, J., Lü, L., Du, B., and Wu, J., 2022. Expression might be enough: Representing pressure and demand for reinforcement learning based traffic signal control. In: *International Conference on Machine Learning*. PMLR, pp. 26645–26654.
- Zheng, G., Xiong, Y., Zang, X., Feng, J., Wei, H., Zhang, H., Li, Y., Xu, K., and Li, Z., 2019. Learning phase competition for traffic signal control. In: *Proceedings of the 28th ACM international conference on information and knowledge management*, pp. 1963–1972.