Improving model predictive control's optimization for urban traffic

Ilyas Khelafa, Abdenaceur Baghdad, Mohamed El Hachimi
Laboratory of Electronics, Energy, Automatics and Data Processing, Department of Electrical Engineering, Faculty of Sciences and Techniques of Mohammedia, Hassan II University of Casablanca, Casablanca, Morocco

ABSTRACT

When it comes to decreasing traffic congestion and enhancing mobility, traffic forecasting is critical. However, due to the complicated spatio-temporal dynamics of urban transportation networks, which are difficult to describe, this task is tough. Using a model predictive controller (MPC) provides the control of a traffic network's architecture as well as errors in its operations. Based on a real-time simulation, a novel, accurate prediction controller for urban traffic was presented in this study to estimate the number of cars at junctions and their waiting duration. Different optimization approaches were employed and evaluated to improve the MPC's performance. Simulation results demonstrated that the fmincon was very robust and could effectively reduce the number of vehicles in the link, in comparison with other algorithms. This study also includes an in-depth analysis of the characteristics of various prediction horizon sets in an MPC. By increasing the prediction horizon, the amplitude of fluctuation became more important, but when Np=4, the fluctuations reduced.

This is an open access article under the CC BY-SA license.

Corresponding Author:
Ilyas Khelafa
Laboratory of Electronics, Energy, Automatics and Data Processing
Department of Electrical Engineering Faculty of Sciences and Techniques of Mohammedia
Hassan II University of Casablanca
Casablanca, Morocco
Email: ilyas.khelafa@etu.fstm.ac.ma

1. INTRODUCTION

Communities are quickly expanding and changing as a result of population expansion, economy, social, and cultural changes [1], [2]. The number of cars has increased as a result of urbanization, particularly in major cities, and issues concerning these vehicles have developed [3]. Despite the fact that building new roads every year is highly expensive, traffic is still an issue that creates congestion, especially during rush hours [4]. Traffic peaks have become one of the most serious issues in metropolitan areas across the world [2], [5], increasing traffic flow and imposing enormous costs on society at all levels. Due to its tight link with issues such as air pollution, fuel consumption, and security concerns, the relevance of traffic control has grown in recent years. In the form of intelligent transport systems, the smart cities concept offers a potential answer to all of these difficulties [3], [5]-[10]. These are usually made up of several connected intersections and rely on automated traffic control [1], [11].

Until now, traffic management procedures have been based on traditional statistical control approaches. However, due to the complicated traffic management system, these approaches are unable to provide an acceptable response. As a result, the development of smart approaches can aid in the resolution and improvement of urban management issues [8], [12]-[14].
In the planning of transportation infrastructure, traffic modeling and simulations are important [11], [15], [16]. Many research projects focus on designing road networks and junctions, evaluating traffic circumstances to minimize congestion, decreasing vehicle delays, and increasing road safety. Microsimulation traffic models have the potential to greatly change the quality of urban road network planning and design [17]-[19]. They may be utilized as part of the decision-making process when comparing and analyzing alternative options, when considering changes in traffic regulation, or when optimizing traffic conditions for an urban road network in general. The VISSIM microsimulation model was utilized in this study [20], [21]. VISSIM is a micro-simulation model for traffic analysis that is stochastic and discrete.

Different traffic control dynamical models have been discussed. To improve vehicle stopping times, traffic capacity, and vehicle delay, Silva et al. developed a multi-objective signal-timing optimization model [13]. The model optimized using such approaches increased traffic capacity, according to the testing data. The time constant of an intersection control technique was either local or coordinated, and the implementation of their method was configured as offline during specific hours of the day for optimization [22], [23]. Traffic lights at junctions were regarded one of the most successful ways of regulating urban traffic [11], [15], [23]. Travel times and vehicle queues can be reduced along a route with traffic-reduction and better-optimized control systems [7], [24], [25]. Fixed-time and smart techniques, can be used to regulate traffic [26]. For that, new dynamic multifunctional optimization systems for the predictive control of existing network signals have been created and presented as a result of these advancements [7], [27].

The purpose of this work is to develop new dynamic multifunctional optimization algorithms for predictive traffic control in cities utilizing real-time simulation. This research also aims to show how prediction horizon affects model predictive control (MPC). The remainder of the paper is as follows: first, a brief explanation of the urban traffic model used during this work is presented. Then, the control algorithm and the model predictive control are detailed, with a description of the optimization methods proposed. Section 3 presents the simulation results, and the effect of prediction horizon on the MPC. Finally, we conclude this paper in the conclusion section.

2. RESEARCH METHOD

2.1. Modelization

The simplified model is presented in this section (called the S model), it is a discrete-time urban traffic flow model. The representation’s degree of the real system qualifies the S model, which has a trade-off between accuracy and minimal computation. In comparison to previous urban traffic models, the S model may incorporate more complicated scenarios such as the time delay required for vehicles running from the beginning of the link to the end of the queues, as well as the available storage capacity in downstream connections [20].

A collection of intersections J and connections L are making up an urban traffic network. A u intersection is a place where numerous traffic streams collide. A link (u,d) is therefore a stretch of urban road that connects two intersections u and d. The upstream and downstream nodes of a link (u,d) are indicated by the letters u and d. For link (u,d), the input and output node sets are u ∈ I and d ∈ O, respectively (e.g., for the situation of Figure 1 we have I_{u,d} = \{i_1, i_2, i_3\} and O_{u,d} = \{o_1, o_2, o_3\}).

![Figure 1. Structure of urban traffic network](image)

The goal of the S model is to find two states variables at each time step $k$, $n_{u,d}$ and $q_{u,d}$. At each time step $k$, the number of cars in link $(u,d)$ is updated based on the input and output average flow rates over $C$ by:

$$n_{u,d}(k+1) = n_{u,d}(k) - (\alpha_{u,d}^e(k) - \alpha_{u,d}^i(k)).C$$

(1)

at every time step k, the number of cars waiting in the line turning to node o is updated over C by:

$$q_{u,d,o}(k+1) = q_{u,d,o}(k) - (\alpha_{u,d,o}^e(k) - \alpha_{u,d,o}^i(k)).C$$

(2)

and at each time step k, the queue length of cars is:

$$q_{u,d}(k) = \sum_{o \in O_{u,d}} q_{u,d,o}(k)$$

(3)

We have the following:

- $C$: Link's cycle time.
- $n_{u,d}(k)$: Vehicle's number during $[k.C, (k + 1)C)$.
- $q_{u,d}(k)$: Queue length during $[k.C, (k + 1)C)$.
- $\alpha_{u,d}^e(k)$: Average link $(u,d)$ entering flow during $[k.C, (k + 1)C)$.
- $\alpha_{u,d}^i(k)$: Average link $(u,d)$ leaving flow during $[k.C, (k + 1)C)$.
- $\alpha_{u,d}^o(k)$: Average incoming flow at the tail of the queue during $[k.C, (k + 1)C)$.

### 2.2. Controller design

Model predictive control is a rolling horizon approach to implement and apply optimum control [28],[31]. An optimum control problem is solved across a prediction horizon in each control step $k_c$, as shown in Figure 2, but only the first control sample in the optimal control sequence is executed. After that, the horizon is moved one sample, and the optimization is resumed with new measurement data. The optimization is re-done using the process’s prediction model and an estimate of the disturbance inputs.

A process model, cost function, and control rule constitute the predictive control structure, as Figure 3 demonstrates. By minimizing the cost function and applying the first control signal, the future output of a system is predicted over a specified period, using the initial process model of computed control signals in the forecast horizon. After that, within the system, calculate it, and repeat the procedure in the following phases. The major goal is to anticipate how much time cars spend on the road, which will be reduced as future system inputs change at each sample period, therefore improving the system's future behavior, which is the green time light. The length of the vehicle queue and the number of vehicles measured per link are both inputs to this controller.

![Figure 2. The concept of MPC](image1)

![Figure 3. The diagram of MPC](image2)

### 2.3. Optimization

It is important to choose the efficient optimizer that is used with the nonlinear model in order to be able to communicate directly with VISSIM [24], [32], [33]. It's software for simulating multimodal transportation operations. In a single model, it depicts all road users and their interactions. All road users are realistically modeled using scientifically sound motion models. To examine the reaction of the optimizer for this system model, fmincon, nonlinear least-squares solver, particle swarm optimization, and pattern search method were used. Matlab is the simulation program that shines out in all methods. The algorithms are...
explained in the following subsections, along with some suggestions for integrating them into the Matlab environment.

2.3.1. Fmincon
Finds the restricted nonlinear multivariable function’s minimum. The minimum of a problem is defined as:

$$\min_{x} f(x) \text{ with } \begin{cases} \mathbf{1} \leq \mathbf{x} \leq \mathbf{u} \\ y \cdot \mathbf{x} \leq \mathbf{z} \\ \mathbf{y} = \mathbf{z} \\ \mathbf{g}(\mathbf{x}) \leq 0 \\ \mathbf{g}(\mathbf{x}) = 0 \end{cases} \tag{4}$$

with $f(x)$ is a function that returns a scalar. $X$, $\mathbf{l}$, and $\mathbf{u}$ are vectors or matrices that can be passed. $y$ and $y\text{eq}$ are matrices, $z$ and $z\text{eq}$ are vectors, $g(x)$ and $g\text{eq}(x)$ are functions that return vectors.

2.3.2. Lsqnonlin
Solves issues using nonlinear least-squares (nonlinear data-fitting). The nonlinear least-squares curve fitting problems is defined as the form:

$$\min_{x} \|f(x)\|^2 = \min_{x} (f_1(x)^2 + \cdots + f_n(x)^2) \tag{5}$$

the optimization is obtained with a lower and upper bounds $\mathbf{l}$ and $\mathbf{u}$ on the components of $x$. $x$, $\mathbf{l}$, and $\mathbf{u}$ can be vectors or matrices.

2.3.3. Particle swarm
The objective of particle swarm optimization is to create a vector $x$ that achieves a local function minimum, as the form:

$$x = \text{particleswarm}(\text{fun,nv,}\mathbf{l}\mathbf{x},\mathbf{u}\mathbf{x}) \tag{6}$$

where $\text{nv}$ represents the number of design variables, of the function. Particle swarm optimization establishes lower and upper bounds on the design variables, $x$, in order to find a solution within the range:

$$\mathbf{l}\mathbf{x} \leq x \leq \mathbf{u}\mathbf{x}$$

2.3.4. Pattern search
This method aims to find a minimum of function, using pattern search algorithm. A local minimum $x$, to the function that computes the values of the objective function, is found via pattern search, as:

$$x = \text{patternsearch}(\text{fun},x\mathbf{0},y,z,x\text{eq},y\text{eq},\mathbf{l}\mathbf{x},\mathbf{u}\mathbf{x}) \tag{7}$$

$x\mathbf{0}$ represents the vector specifying an initial point for the pattern search algorithm.

3. RESULTS AND DISCUSSION
3.1. Simulation
The numerical experiments finding are presented in this section. The suggested methodology’s performance is evaluated in this part using VISSIM, a real-time simulator linked to Matlab through a communication interface (COM) [20], with 14 connections and two junctions as shown in Figure 1. Links are of the same length: 900 m. All connections are assumed to have three lanes. $V_{\text{free}} = 50$ km/h is the free-flow speed, and the average vehicle length is 7 meters. For network roads, the storage capacity $C(u,d)$ is 386 vehicles.

In each link, the method has been evaluated at various saturation flow rates: 1600 veh/h for left turn, 1500 veh/h for right turn, and 1800 veh/h for through turn. Each road has the same turning rates (0.33 for every direction). Furthermore, as Figure 4 demonstrates, all simulations are conducted with variable traffic demand and producing congested traffic in the simulator.
The four optimization’s methods are simulated on this platform for the performance comparison: fmincon, lsqnonlin, particleswarm and patternsearch. To explore the four optimization algorithm performances, their simulation results are illustrated in Figures 5 and 6. They show the variation of the vehicle’s number and their queue length results when the traffic demand shows a variation with 2 peaks, as the Figure 4 illustrates. It can be seen that fmincon obtains the best final results.

- Fmincon achieves the best performance among the four optimization methods. The best and average results obtained by fmincon are all smaller than the corresponding results obtained by other algorithms, respectively, which demonstrates the strong optimization ability of fmincon in reducing the number of vehicles in the link. We can say that fmincon adapts well to different traffic volumes.
- Lsqnonlin performs similar to fmincon, but before the 1st peak of congested demand. Lsqnonlin lost its stability with increasing demand. As we can see from Figures 5 and 6 that the average results obtained by fmincon and lsqnonlin are close to each other before 1st peak of demand. The worst results of lsqnonlin are generally better than that of particleswarm and patternsearch.
- Patternsearch performs similar to particleswarm, but it is more stable. As we can see from Figures 5 and 6 that the results obtained by the two algorithms patternsearch and particleswarm are close to each other. More than that, They present better results face of fmincon and lsqnonlin, but it was limited by the increasing of vehicle’s demand (1st peak). Furthermore, they keep the worst results even if the demand decreases. Based on the simulation results we can conclude that fmincon is an effective tool for solving traffic signal optimization problems and can significantly reduce the number of vehicles, and that makes the MPC more effective in front of the other controllers, like fixed time controller [20].
3.2. Prediction horizon effect

As shown in Figure 2, the prediction horizon, \( N_p \), is the number of future check intervals that the MPC must anticipate when optimizing its movements to check interval \( k \). Then there's an integer that indicates how long the MPC algorithm should take to get the controlled variables to their reference values. The prediction horizon \( N_p \) is essential both in terms of the amount of data fed to the algorithm and in terms of the optimization problem's numerical viability [34], [35].

Figure 7 presents the results for urban traffic control with different prediction horizons. The most noticeable point in Figure 7 is that the number of vehicles is almost dependent of prediction horizon. Note that when the prediction horizon increase (from \( N_p=1 \) to \( N_p=3 \)) that yields to vehicle's number increase too. When \( N_p=4 \), there is a change about the number of vehicles, with this prediction horizon, we have better performance front of \( N_p=3 \), but the best results still with \( N_p=1 \).

![Figure 7. The number of vehicles with different optimization algorithms](image)

In Figure 8, some fluctuations can be seen. These fluctuations are very important for \( N_p=3 \). We can remark that the controller loses its performances. For \( N_p=4 \), the MPC takes back the stability, even if it’s different from other prediction horizons (\( N_p=1 \) and \( N_p=2 \)).

For more illustrations, Figure 9 shows control input, which is the green time length. As we can remark, even the output (vehicle’s number) has better results with low prediction horizon, but for control input, it’s remarkable that performs poorly in comparison with \( N_p=4 \) (there is more fluctuations). Furthermore, it is necessary to find a good trade-off between the performance of the system and allocated or available computational complexity.

![Figure 8. The total time spent by vehicles](image) ![Figure 9. The green time length](image)
4. CONCLUSION

In this paper, four modified optimization algorithms, i.e. fmincon, lsqnonlin, patternsearch and particleswarm, were proposed and used to optimize the urban traffic signal. To assess the efficiency of different algorithms, a Vissim-Matlab integrated real-time simulation was realized, in which the former works as a traffic simulator for evaluating candidate solutions and the latter is responsible for the implementation of algorithms. Through this simulation platform, the algorithm optimization performances were tested. Simulation results on the traffic network demonstrated that the fmincon was very robust and could effectively reduce the number of vehicles in the link. In comparison with other algorithms, fmincon achieved the best overall performance in different traffic conditions, like varying traffic volumes.

Another goal of this paper is to present the effect of prediction horizon, by studying a MPC framework. According to the simulation results, when applying different prediction horizon sets in the control of urban traffic, good results are obtained in terms of vehicle’s number and the total time spent by vehicles. By increasing the prediction horizon, the amplitude of fluctuation became more important, since Np=4, the fluctuations reduce, and the controller output is more stable.

REFERENCES


Ilyas Khelafa received a technical university degree in electrical engineering, in 2012 from Higher School of Technology in Fez. In 2015, he graduated from FST Mohammadia with an engineering diploma in electrical and telecommunications. Currently, he is a Ph.D. student in the Electronics, Energy, Automatic and Information Processing (EEA&TI) laboratory at FSTM, Hassan II University of Casablanca. Modeling and control of urban traffic networks are among his research interests. He can be contacted at email: ilyas.khelafa@etu.fstm.ac.ma.

Abdennaceur Baghdad received his doctorate in electronics from Lille University in France in 1992. He is currently a university professor with a PES grade at Hassan II University of Casablanca, Morocco, where he teaches hyper-frequencies, electrical engineering, antennas, and telecommunication networks. He is a member of Hassan II University's EEA&TI laboratory. He supervises and co-supervises a large number of doctoral theses. His studies focused on optical networking, electrical and integrated networks, telecommunication systems, and information technology. He can be contacted at email: nasser_baghdad@yahoo.fr.

Mohamed El Hachimi received his doctorate in electrical engineering from Hassan II University in Morocco in 2020. He is employed at the FST-Mohammed University Hassan II of Casablanca's Laboratory of Electronics, Energy, Automatic, and Data Processing (LEEA&TI). EL HACHIMI's current research focuses on Control Systems Engineering and Electrical Engineering, also on a Model Predictive Controller for an Artificial Pancreas. He can be contacted at email: contact.hachimi@gmail.com.