A market-inspired approach to reservation-based urban road traffic management

Matteo Vasirani
Centre for Intelligent Information Technology
University Rey Juan Carlos
Madrid, Spain
matteo.vasirani@urjc.es

Sascha Ossowski
Centre for Intelligent Information Technology
University Rey Juan Carlos
Madrid, Spain
sascha.ossowski@urjc.es

ABSTRACT

Urban road traffic management is an example of a socially relevant problem that can be modelled as a large-scale, open, distributed system, composed of many autonomous interacting agents, which need to be controlled in a decentralized manner. Most models for urban road traffic management rely on control elements that act on traffic *flows*. Dresner and Stone have put forward the idea of an advanced urban road traffic infrastructure that allows for cars to *individually* reserve space and time at an intersection so as to be able to safely cross it.

In this paper we extend Dresner and Stone's approach to *networks* of intersections. For this purpose, we draw upon market-inspired control methods as a paradigm for urban road traffic management. We conceive the system as a computational economy, where driver agents trade with infrastructure agents in a virtual marketplace, purchasing reservations to cross intersections when commuting through the city. We show that in situations of similar traffic load, an increase of the infrastructure's monetary benefit usually implies a decrease of the drivers' average travel times.

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Coherence and coordination, intelligent agents, multiagent systems*

General Terms

Algorithms, Design, Experimentation

Keywords

market-based coordination mechanisms, reinforcement learning, traffic and transportation

1. INTRODUCTION

The control of a large-scale, open, distributed system, composed of many autonomous interacting agents, with the aim of instilling some desired global properties, is not a trivial task. Being too complex for centralized decision making, the only feasible way is delegating and distributing power and control. Unfortunately, decomposing and distributing the problem generates other issues to cope with.

Consider (future) traffic control systems: intelligent traffic infrastructures, provided with sensors and computing power, aiming at resolving congestions and speeding up the traffic flow; millions of drivers commuting every day from their homes to their respective workplaces and back, making autonomous decisions about route assignment and departure time selection, learning from their past experiences and influencing each other in both positive and negative ways. Here the infrastructure components are faced with a very complex problem, since they aim at controlling a system that is partially observable (e.g. a component cannot access directly the degree of satisfaction of a driver), and without "powerful actuators" (e.g. it is unable to directly intervene in the drivers' behaviour).

In this paper we draw upon market-based control methods [6, 12] as a paradigm for the management of an urban road traffic system, which would otherwise be very difficult to control and maintain. We conceive the system as a computational economy, where driver agents trade with the infrastructure agents in a virtual marketplace, purchasing reservations to cross intersections when commuting through the city. We design the market rules with the aim of aligning the "global profit" (revenues from the infrastructure use) with the "social welfare" (e.g. average travel time), in a way that, in situations of similar traffic load, an increase of the infrastructures' monetary benefits usually implies a decrease of the drivers' average travel times.

The paper is structured as follows: in section 2 we introduce the context and the motivation of our work; in section 3 we present our model of traffic as a computational economy; section 4 shows the experimental results; finally we conclude in section 5.

2. PREVIOUS WORK

In recent years, there is a growing interest in applying agent-based techniques for traffic management [4, 5]. Urban road traffic control appears to be a particularly promising application area for agent technology. To this respect, many approaches aim at optimizing the use of existing traffic infrastructures, by providing adequate coordination policies that are either designed off-line or learned at run-time. For example, in [10] intelligent traffic light agents create "green waves" in a particular direction, while in [11] the traffic lights learn in a coordinated way the best signal plans. Still, in these approaches just the intersections are modelled as agents, while drivers are only considered insofar as they are

Cite as: Title, Author(s), *Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009)*, Decker, Sichman, Sierra and Castelfranchi (eds.), May, 10–15, 2009, Budapest, Hungary, pp. XXX-XXX.

a part of the traffic flow through the road network.

Other approaches conceive the drivers as the agents whose behaviour is to be modelled [15] (e.g. for simulation purposes). In this context, it is particularly interesting to study mechanisms that influence driver behaviour so as to improve their local utility (e.g. to reduce travel time) and/or to enhance the global system performance (e.g. to reduce number and size of congestions). Variable message signs and onboard driver information systems, for instance, may help agents avoid congested road sections [8], but may cause new problems when used by a large population of drivers [2]. Toll-based approaches dynamically adjust the price of using certain road sections so as to achieve an adequate traffic flow distribution within the road network [19]. However, a tight integration with the aforementioned approaches is difficult, since existing urban road traffic management infrastructures based on traffic light controlled intersections affect traffic *flows*, but cannot act on *individual* cars.

The situation is different in air traffic management. In [17], air traffic control agents manage the "fixes" that airplanes transit through, and learn the optimal delay to introduce between flights, in order to minimize travel times and congestions. These infrastructure agents control the system in a distributed way, acting directly upon the individual entities (planes) that compose the (air) traffic flow, a possibility that, as we have argued, does not hold for today's urban road traffic management.

Nevertheless, this is likely to change in a future not too far from now. For instance, Dresner and Stone [7] predict an evolution towards advanced "reservation-based" road traffic infrastructures that allow for an agent-centric, instead of a flow-centric, control of intersections. In their model, an intersection is not regulated by traffic lights, but by an intelligent agent that assigns reservations of space and time to each individual vehicle intending to cross the intersection. Their work assumes two types of agents, which are capable of *communicating* with one another:

- *Intersection manager agents* control the space of an intersection and schedule each individual driver's transit through it;
- *Driver agents* autonomously operate their assigned vehicle.

When a vehicle is approaching an intersection, the driver agent requests the intersection manager agent (*driver* and *intersection manager* from now on for short) to reserve the necessary time-space slots to safely cross the intersection. The intersection manager, provided with data such as vehicle ID, vehicle size, arrival time, arrival speed, type of turn, etc., simulates the vehicle's transit through the intersection and informs the driver whether or not its request is in conflict with the already confirmed reservations. If there is no such conflict, the driver stores the reservation details and tries to meet them; otherwise it may try again at a later time. If the driver realizes that the traffic conditions have changed and that it is not able to meet the reservation constraints, it can cancel the reservation and make a new one [7].

3. MARKET-INSPIRED APPROACH

In this paper, we set out from Dresner and Stone's work and assume the existence of an advanced traffic management infrastructure that allows for reservation-based intersection control. Furthermore, we draw upon ideas from toll-based systems and allow intersection managers to sell time-space slots at an intersection, thus generating incentives to prefer or to avoid routes that pass through certain intersections. In addition, we apply distributed learning techniques that have been successfully applied to the air traffic domain, so as to dynamically coordinate the intersection managers' pricing policies. In such a setting, urban road traffic management can be conceived as a computational economy, where drivers trade with the intersection managers in a virtual marketplace, purchasing reservations to cross intersections when commuting through the city.

As with traditional toll-based systems, we would like the "global profit" (revenues from the infrastructure use) and "social welfare" (e.g. average travel time) to be aligned: we would like to build our market in a way that, in situations of similar traffic load, an increase of the infrastructures' monetary benefits usually implies a decrease of the drivers' average travel times. Of course, there is no way to *directly* influence autonomous driver behaviour for this purpose, but we can act on (parts of) the infrastructure, in particular the interaction protocol that driver and intersection managers have to comply with, as well as the agent programs of the intersection managers.

This section is organized as follows: we first present the general characteristics of the urban road traffic management infrastructure that we envision, and describe the protocol by which drivers can purchase reservations from intersections. Section 3.2 puts forward both individual and team learning models by means of which the intersection managers acquire their pricing policies. Finally, in section 3.3 we outline our model of individually rational driver behaviour in this context.

3.1 Interaction model

We assume that the agents in the future urban road management infrastructures will have the following capabilities:

- Intersection managers are able to communicate with each other. This assumption is reasonable, e.g. already existing fibre-optic connections along certain main urban roads could be used.
- Drivers can communicate with intersection managers. Such proximity-based communication is already in use in different elements of today's traffic infrastructures. We assume that a driver is able to communicate with the forthcoming intersection on its route, and also with the neighbours of such intersection (see figure 1).
- Drivers can be provided with the current prices of the intersections in the network. This can be done, for instance, by a price propagation scheme through the intersection network[1].
- A trusted payment system is available, allowing drivers to securely transfer money to intersection managers when required. Such mechanisms are already in use in today's toll roads.

[1]See for example [9], where a gossip-based, adaptive protocol for extremely large and highly dynamic networks is presented. Such protocol has been successfully tested on a network distributed over five continents, whose number of nodes dynamically oscillated between 2500 and 6000.

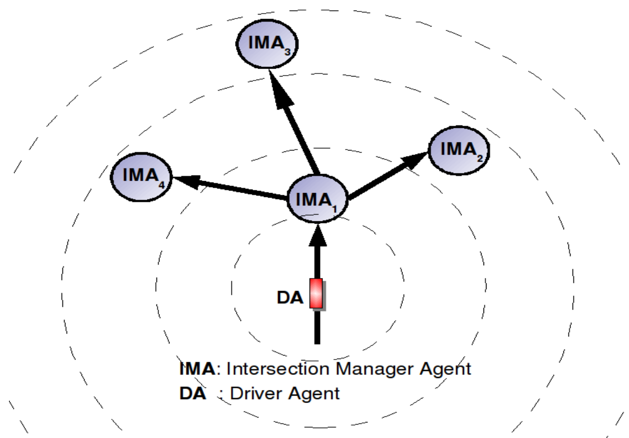

Figure 1: Communication range

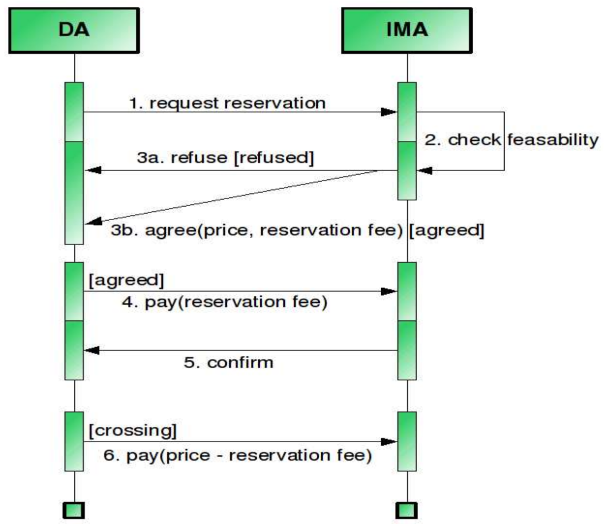

Figure 2: Purchasing protocol

The work of Dresner and Stone, which focuses on engineering a mechanism for single intersection control, cannot be extended directly to this setting. As there is no notion of cost associated to a reservation of space-time slots at intersections, a driver has no incentive to prefer a particular intersection over another. Consequently, when commuting from an origin to a destination, drivers are likely to choose the route with the lowest (estimated) travel time, which will lead to congestion problems in the network in all but low-load situations.

As outlined previously, in this paper we propose to overcome this problem through the establishment of a computational marketplace, where the drivers (i.e., buyers) must purchase the necessary reservations from the intersection managers (i.e., sellers), in order to cross an intersection. The introduction of the monetary factor provides the drivers with incentives to explore alternatives to the shortest paths, and provides a team of intersection managers a lever to control the system.

Urban road traffic networks can potentially be quite large, so it is not feasible for drivers to make a global "one-shot deal" with the infrastructure as a whole. Instead, we assume that each intersection manager, as a seller, is free to set its desired fare for the reservations that it manages within a certain price range, and that drivers keep purchasing reservations as they travel through the network. Also, in many countries providing a public road infrastructure with a basic quality of service is seen as a state obligation that the citizens cannot be charged for, at least not directly (e.g. there should always be a possibility to reach a destination without having to use toll roads). Therefore, drivers should have the possibility of crossing intersections for free, if they accept a potentially significant increase in their travel times, especially in high-load situations.

In such a setting, in order to design the rules of the marketplace, it is essential to specify the regulations that govern the interactions between a driver and a single intersection manager. Such regulations need to specify how succesful deals are made (i.e. how to make a reservation, and how to use it), and what happens if something goes wrong (e.g. when a reservations needs to be withdrawn, or when a driver arrives at an intersection without a valid reservation).

3.1.1 *Purchasing a reservation when approaching the intersection*

Drivers are able to make reservation requests with intersection managers in its proximity. In order to purchase a reservation, a driver "calls-ahead" the intersection manager and provides the necessary data to simulate its transit through the intersection [7]. A driver is uniquely identified by a vehicle ID, so that at each intersection a driver may hold only one reservation.

If the request cannot be satisfied, due to conflicts with the already confirmed reservations, the intersection manager refuses the reservation request. Otherwise, it notifies the driver with the reservation fare. To actually hold the reservation, the driver must transfer a percentage of the fare as reservation fee, while the intersection manager commits itself to maintain the contracted fare when the driver will pay the rest of the reservation fare during the transit through the intersection. When the driver actually arrives at the intersection, it pays the remaining amount and crosses it safely using the intersections' time-space slots that have been reserved for it (see figure 2).

Notice that this protocol presumes that drivers hold a reservation in order to safely cross the intersection. Although nothing can physically impede the driver to cross the intersection without a reservation, we rely on the assumption that a driver is risk averse and does not take the risk of causing an accident. This assumption also holds for today's intersections regulated by traffic lights or stop signs [7]. Furthermore, a driver has no incentives for taking the risk of crossing an intersection without a reservation for monetary purposes, since it has the possibility of having a reservation for free if it stops at the intersection (see below).

3.1.2 *Receiving a reservation when stopped at the intersection*

If a driver does not hold a valid reservation (either because it did not purchase one, or because it arrives late) when it reaches the edge of the intersection, it must stop. In this case, it is entitled to receive a reservation, when available, for free. The intersection manager is in charge of assigning such a free reservation. Although it may give priority to buyers, it is obliged to eventually grant the reservation, if such request has been denied for a maximum number of times. Furthermore, depending on the intersection topology, it is likely that the bigger the number of stopped cars, the less capacity there is for paying drivers to cross.

To this respect, we assume that the road infrastructure provides intersection managers with a way to actually confirm that a vehicle is stopped at the intersection (e.g. with cameras), in order to avoid that the mechanism is exploited by strategic drivers. While approaching an intersection without a valid reservation is a legal strategy for drivers, it will only be interesting in low-load situations or otherwise to drivers that essentially neglect the criterion of travel-time for their route choice.

3.1.3 Withdrawing a reservation

When a driver purchases a reservation, it tries to meet the reservation constraints, especially the arrival time. If it realizes that these cannot be met, due for example to changing traffic conditions, it can withdraw the reservation. In this case, the driver will lose the reservation fee paid in advance[2].

Notice that, in situations where the travel time to the next intersections can be estimated reasonably well, drivers have incentives for making a reservation as soon as possible. Furthermore they have no incentive to make reservations on intersections lying on alternative paths through the network. If the chance to arrive in time for the reserved time slot is high, the cost involved in making reservations "just in case" on other intersections does not outweigh the potential gains. For instance, in the situation shown in figure 1, where a driver is capable of communicating with the intersections in its proximity, a driver would want to make its reservations as soon as it enters the communication range of the intersection managers on the chosen route.

3.2 Intersection manager agent model

Intersection managers apply the simple "first-come first-served" algorithm described in [7] to honour reservation requests[3]. Therefore, their decision problem boils down to determining the current reservation fare, and coordinating it with the other intersection managers in the team.

3.2.1 Action space

The action space Z_i of an intersection manager is composed of the prices that it can apply to the reservation fares that it manages. More formally:

$$Z_i = \{p_i \in [p_{min}, p_{max}]\} \tag{1}$$

where p_{min} and p_{max} are the minimum and maximum allowed price for a reservation fare.

3.2.2 Profit function

An intersection manager is characterized by its profit function U_i, defined as the difference between revenues R_i and costs C_i. More formally:

$$U_i = R_i - C_i = \sum_t r_t - C_{max} \cdot e^{-\sum_t d_t} \tag{2}$$

The revenues R_i are calculated as the money earned with the reservations that have been sold over time, r_t. The cost function C_i is a function of the number of drivers that have purchased a reservation through time, d_t. The cost function has a maximum if no drivers have purchased a reservation, and tends to 0 with the increase of drivers (i.e., the costs are amortized).

Such profit function intends to penalize unused intersections, by a mean of few revenues and high costs, as well as congested ones, since vehicles stopped at the intersection do not generate any revenue (recall that a vehicle stopped at the intersection is entitled to receive a reservation for free).

3.2.3 Global profit maximization

Each intersection manager must decide which fare to apply to its intersection in order to maximize the profit. If the fare is too high, the intersection is likely to be avoided by the drivers, generating low profit. On the other hand, if the fare is too low, the intersection will be congested, resulting again in low profit (recall that a vehicle stopped at the intersection is entitled to receive a reservation for free). Furthermore, a fare cannot be defined high or low in absolute terms, but is strictly dependent on the fares applied by the other (possibly neighbouring) intersections.

However, as mentioned previously, intersection managers are part of the infrastructure, so we can program them to work as a team. From the point of view of effective teamwork, the collective of intersection managers is so faced with two tasks: i) discovering the effect of a specific fare scheme and ii) coordinating their fares in order to maximize the global profit. Given equation 2, the global objective of the team of intersection managers can be expressed as a function of the joint action z, $G(z) = \sum_i U_i$. Here the joint action z is defined by the fare scheme of the team of intersection managers, in other words $z = \langle z_1, z_2, \ldots, z_n \rangle \in Z_1 \times Z_2 \times \cdots \times Z_n$, where z_i is the fare applied by the intersection manager i on its intersection. We remark that the functional form of G as a function of z is not known, since it is not known which profit U_i is generated by a specific z.

To learn in a distributed and coordinated fashion which fare scheme leads to the best system performance G we use Q-learning with immediate rewards and ϵ-greedy action selection [18]. After having taken action z_i, the learning agent receives a reward that rates that action, then it updates its action-value function estimation as follows:

$$Q_{t+1}(z_i) = Q_t(z_i) + \alpha \cdot [\mathcal{R}_t(z) - Q_t(z_i)] \tag{3}$$

where α is the learning rate and $\mathcal{R}_t(z)$ is the reward, which depends on the full joint action z. Each intersection manager selects a random action with small probability ϵ, and the *greedy* action (i.e., the action with highest Q-value) with probability $(1 - \epsilon)$.

We evaluated two reward functions, namely the local reward $L_i(z)$ and the (estimated) difference reward $D_i(z)$.

1. Local reward: $L_i(z) = U_i$

   An intersection manager is rewarded with its own profit.

2. Difference reward: $D_i(z) = G(z) - G(z - z_i + c_i)$

[2] A possible extension of such trading interaction could be giving the drivers the possibility of *selling* the reservations that have been purchased but that cannot be used.

[3] If a request for a specific time is not possible to grant, it is rejected by the intersection manager without any further consideration.

The notation $z - z_i + c_i$ refers to a vector where all the components of z affected by agent i are replaced by the constant c_i. Since it is not possible to calculate the term $G(z - z_i + c_i)$ when the functional form of G is not known, we follow [17] for the estimation of the difference reward. If $G(z) = G_f(f(z))$, where G_f has a known functional form, and if $f(z) = \sum_i f_i(z_i)$, where each f_i is an unknown function, the difference reward can be estimated as $\mathbb{E}[\sum_i U_i|z_i] - \mathbb{E}[\sum_i U_i]$. In other words, the difference reward can be estimated as the expected global profit when agent i applies fare z_i minus the expected global profit[4].

3.3 Driver agent model

The deliberation process of a driver is shaped by the fact that it must purchase the reservations to cross the intersections that it encounters during its trip. We model a driver as an individually rational utility maximizer, which aims at choosing the "best" route, accordingly with its utility function. Route choice is the fourth step in the conventional transportation forecasting model [13]. Such utility function, B_i, is a weighted sum of two factors: travel time and costs. More formally:

$$B_i(x) = -[\rho \cdot TT(x) + (1 - \rho) \cdot K(x)] \quad (4)$$

where ρ is a trade-off factor that weights the relative importance of the travel time of route x, $TT(x)$, and the costs implied by such route, $K(x)$. The cost of a route x is the sum of the reservation fares applied by the intersection managers that govern the intersections that lay on route x. Still, special attention deserves the travel time function $TT(x)$. Although this function is unknown in general, we use an optimistic estimation $TT^{est}(x)$ of the travel time, calculated as:

$$TT^{est}(x) = \frac{||x||}{v_{max}} \quad (5)$$

where $||x||$ is the route length and v_{max} is the maximum allowed speed[5].

We simulate driver decision making through time as follows: when a driver intends to commute through the road network, firstly it builds its route x, based on the available route and price information. Then it starts commuting following that route, and when it is approaching the first intersection on its path it "calls-ahead" the intersection manager and tries to purchase a reservation (see figure 2). The driver reasons about its route after crossing each intersection; it continuously re-assign its route, choosing the just crossed intersection as new origin, and so reacting to the market fluctuations.

4. EXPERIMENTAL RESULTS

4.1 Simulation environment

[4]Calculating G for a given fare scheme z can be done also in a decentralized way. Similarly to the price propagation scheme, a gossip-based protocol [9] can be used to aggregate the local profit values into a global value.

[5]In other words, the driver optimistically estimates the travel time as the travel time at free flow

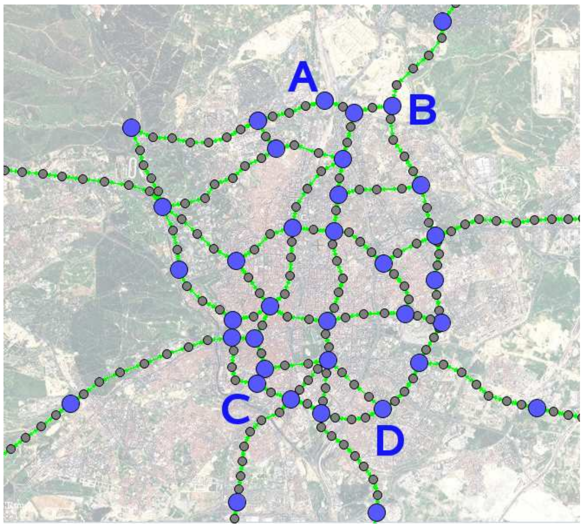

Figure 3: Simulator

To evaluate the approach here presented we implemented a hybrid mesoscopic-microscopic simulator. The traffic flow on the roads is modelled at mesoscopic level, where the dynamics of a vehicle is governed by the average traffic density on the link it traverses rather than the behaviour of other vehicles in the immediate neighbourhood as in microscopic models [16].

Since the mesoscopic model does not offer the necessary level of detail to model a reservation-based intersection, when a vehicle enters an intersection its dynamics switches into a microscopic, cellular-based, simulator, whose update rules follows the Nagel-Schreckenberg [14] model. The cell size is set to 5 meters, and for simplicity we assume that the vehicles cross the intersection at a constant speed, so that any additional tuning of parameters, such as slowdown probability or acceleration/deceleration factors, is not necessary.

The experimental setup is based on an instantiation of the simulation environment based on the city of Madrid (see figure 3, where each big dark vertex is an intersection).

For the intersection manager model the minimum and maximum prices p_{min} and p_{max} were set to 1 and 10 respectively, while the maximum cost C_{max} was to 1. Regarding the learning algorithm, we selected by trial-and-error $\alpha = 0.5$ and $\epsilon = 0.1$, evaluating the local reward and the difference reward. The Q-values are initialized optimistically with the maximum global profit (see section 4.2), in order to guarantee more exploration.

We simulated 2000 drivers (the double of the optimal flow in the simulated road network) commuting along the North-South axis (from A and B to C and D), generated in an interval of 15 minutes.

4.2 Case 1: Single driver agent model

In these experiments we evaluate the system using as trade-off factor for the route choice $\rho = 0.5$ (see eq. 4). In other word, the decision making of the driver is equally affected by the estimated travel time on a particular edge and the fare applied by the intersection manager at the end of that edge.

Figure 4 plots the global profit dynamics during the learning for the two reward functions. The maximum global

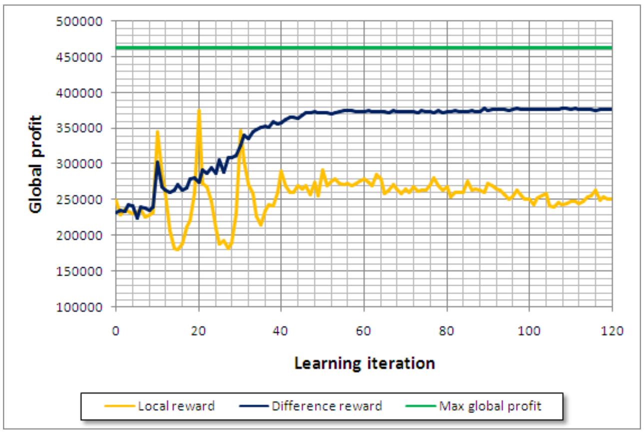

Figure 4: Global profit

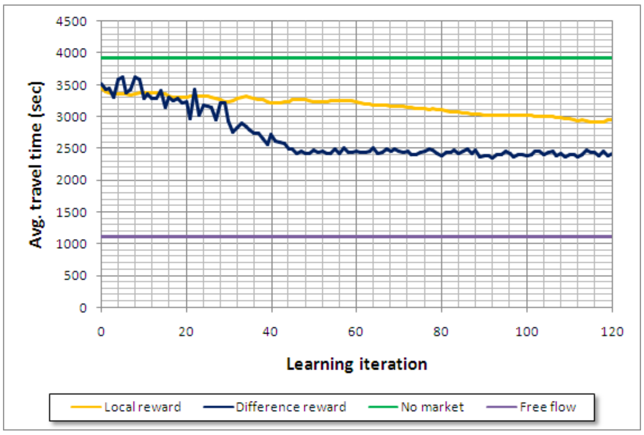

Figure 5: Average travel time

profit, used as baseline, is the global profit that would have been obtained if *all the intersection managers* sold a reservation at the maximum allowed price to *all the drivers* (i.e., every driver transits through all the intersections). Using the difference reward, the intersection managers are able to jointly find a fare scheme that generates high profits, while with the local reward the global profit tends to decrease through time.

Nevertheless it is interesting to see the effect of this profit maximization on the social welfare. We reasonably assume that the lower the average travel time, the higher the social welfare, since every driver decides autonomously how much wealth to allocate on its trip. Figure 5 plots the average travel time of the set of 2000 drivers, using the two different reward functions. The lowest limit is represented by the average travel time at free flow, i.e., when there is no traffic. The highest limit is represented by the average travel time of the 2000 drivers if we remove the market mechanism. In this case the drivers always select the shortest path, generating congestions and, consequently, higher travel times. On the other hand, if the intersection managers try to maximize profit, they indirectly influence the driver decision making and so better allocate the road network resource, generating lower travel times through time. Furthermore, using difference reward the average travel time is lower with respect

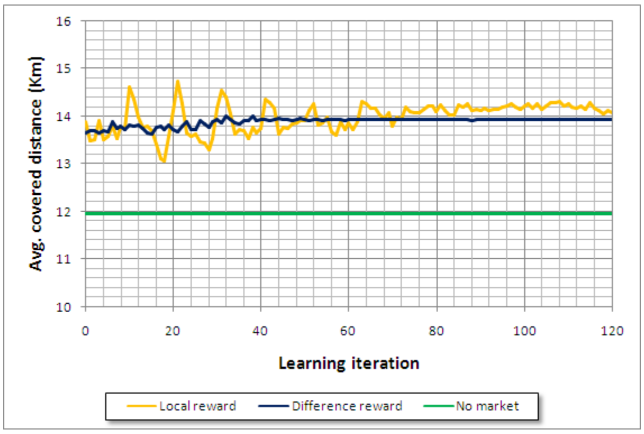

Figure 6: Average covered distance

to using the local reward. This is another clue of the alignment between private profit maximization and social welfare maximization, since to higher global profit corresponds lower average travel time.

This fact is reflected also by the average distance covered by the drivers (see figure 6). Without the market, the covered distance corresponds with the shortest route between the North sources and the South destinations, reflecting the fact that the drivers are congested on the same route. On the other hand, when the agents must participate in the virtual market, the maximization of the global profit performed by the intersection managers make the drivers spread through the network, causing higher covered distances. Nevertheless, an increase of the average covered distance of about the 15% (figure 6) generates a decrease of the average travel time of about the 40% (figure 5).

4.3 Case 2: Different driver agent models

In these experiments we evaluate how the system reacts to two different models of driver. In the first model we set $\rho = 0$ as trade-off factor for the route assignment. This means that the driver's decision making is affected only by the market information, i.e., it prefers routes that transit through "cheap" intersections, no matter the travel time. We call these agents "price-based drivers". Here we are interested in modelling drivers that prefers cheap intersections but still they are willing to pay money for not being stuck at an intersection waiting for receiving a reservation for free. The latter should be modelled as "time-based" drivers (see below) that do not request any reservation when approaching the intersection.

In the second model, 50% of the drivers has $\rho = 0$, and the other 50% has $\rho = 1$ as trade-off factor. The latter type of driver only considers the travel time information, i.e., it prefers the shortest route, no matter the price applied by the intersection managers. We can call these agents "time-based drivers".

We remark that the case where all the drivers have $\rho = 1$ as trade-off factor is not interesting to evaluate the market-based scenario, since they are not affected at all by the market.

Figure 7 plots the global profit and the average travel time when the system is populated by "price-based drivers", using the difference reward and the local reward. The global

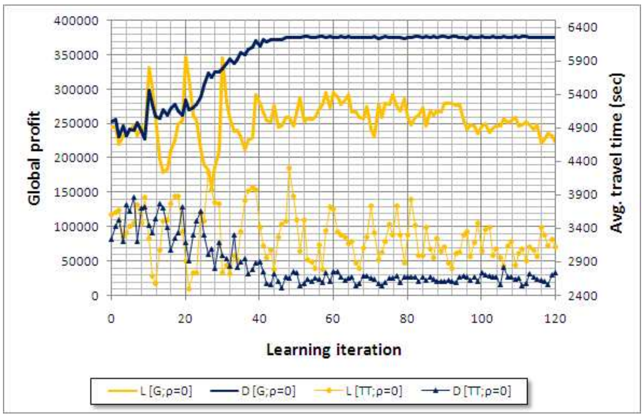

Figure 7: Global profit and avg. travel time with "price-based drivers"

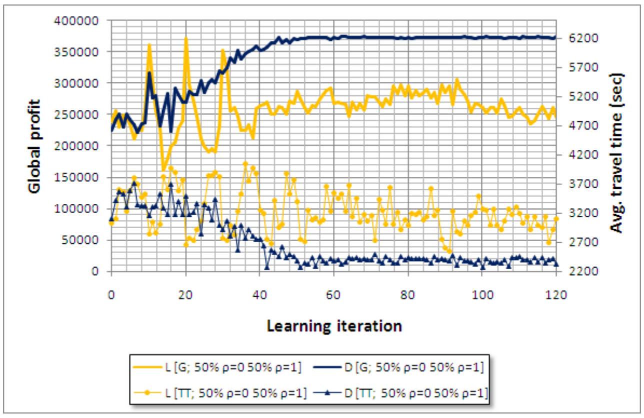

Figure 8: Global profit and avg. travel time with "price-based" and "time-based" drivers

profit reached by the intersection managers is approximately the same of that reached in the experiments of section 4.2. If the drivers are modelled with the same parameters, the maximization of the global profit is a single objective optimization, since all the drivers react in the same way to the market fluctuations, levered by the intersection managers. On the other hand, the average travel time is higher than that of the experiments of section 4.2.

Figure 8 plots the global profit and the average travel time when the system is populated for a half by "price-based drivers" and for a half by "time-based drivers", using the difference reward and the local reward. This the most realistic setting, since it is reasonable that a part of the population of drivers may own more money than time, so preferring shorter and more expensive routes (e.g. business drivers), while the other part of the population of drivers may have more time to spend than money (e.g. leisure drivers). The global profit reached by the intersection managers is slightly lower than that of previous experiments, due to the "noise" added by the "time-based drivers" that are not affected by the market. On the other hand, the average travel time is lower than that of the previous experiments. Furthermore, in all the experiments the difference reward performs better than the local reward, especially when the population of drivers is composed of "price-based" and "time-based" drivers.

5. CONCLUSION

In this paper we studied the application of market-inspired mechanisms for managing future urban road traffic infrastructures, where drivers can individually reserve the necessary space-time slots to safely cross intersections. We put forward an interaction protocol between drivers and intersection managers that aims at efficiently assigning reservations to drivers, based on a pricing model, while reducing the effect of strategic manipulation. At the same time, this protocol maintains the option for drivers to navigate through the road network for free, although with increased travel times – an important characteristic to foster the acceptance of the mechanism. We have put forward a learning model for intersection manager agents, so as to coordinate their pricing policies within a team of intersection managers. A model of individually rational driver behaviour within this context has been outlined based on the notion of route choice. Finally, we have analyzed the system performance with regard to different driver profiles. We showed that, in general, an increase in the global profit of the intersection manager team is aligned with reduced average travel times, and that coordinated policies of intersection managers (as part of the infrastructure) outperform significantly strategies based on the maximization of only local utility. This holds for different types of drivers populations. In summary, besides the "knowledge engineering" work of framing the problem and designing adequate "rules of the game" for a domain of potential social relevance, our contribution consists in an innovative combination, adaptation, and integration of economically inspired and computational learning techniques for a truly open class of multiagent system.

This paper is, of course, related to Dresner and Stone's work that we draw upon [7]. We have extended that work significantly along two lines: (i) we substituted the reservation protocol for time-space slots through a trading model; and (ii) we extended the approach from a single intersection setting to an urban road network with multiple intersections, giving rise to new problems of coordinating the intersections price-based time-space reservation policies.

Unlike traditional learning-based methods for traffic control that consider traffic flows, history-based controllers – just like this work – focus on single vehicles [1]. Vehicles earn credits when they wait at traffic light and pay when passing through intersections. Traffic signals base their decisions on the credits of the various vehicles stopped at the intersection. At the end of its trip, a vehicle communicates its commuting time. Still, collecting the commuting times of all the vehicles, to assess the efficiency of the control, seems quite unrealistic. Furthermore, this approach does not provide incentives to vehicles to refer their commuting time nor to report it truthfully.

Bazzan and Junges [3] study how to affect single vehicle decision making using congestion tolls. In that work, a control centre provides agents with a (noisy) estimation of the cost of choosing a certain route r. The agents periodically update the heuristic information related to the available routes on the basis of the utility received in the past episodes. As in this paper, the aim is to align the private utility with the global optimum, although in [3] this is done in a centralized way, with the control centre predicting the traffic state and providing information to the drivers. In our work the global optimum distributed maximization occurs at the market level, and such alignment is indirectly induced by the market rules.

Regarding the distributed coordination of the pricing policies, our approach is most similar to the work of Tumer [17] related to air traffic control. We exploited the similarities between the two domains to set up an agent-centric model for urban road traffic control. However, in the urban road traffic domain intervening directly on the single entities that compose the traffic flow is not possible and only indirect influence is viable.

The lines along which this work can be extended in the future are manifold. In the above experiments the price of a reservation requested at time t was unique. It is reasonable to assume that the value of a reservation is inversely proportional to how far in advance it is requested: drivers in a hurry may be willing to pay the more the closer the requested time slot. One may also want to take the road network topology into account. For instance, certain "bottleneck intersections" may deserve special attention and may be endowed with a bigger price range to choose among than their less frequented counterparts.

The coordinated fare scheme learnt by the team of intersection managers depends on the demand pattern, i.e., the amount and profile of the drivers that populate the road network. Still, it would be interesting introducing states in the learning problem, and so coping with the dynamic changes of the demand patterns that may occur during a day (e.g. morning peak, noon, etc.).

In the present work we used a simple reservation assignment protocol – there was no explicit negotiation between drivers and intersection managers. A possible extension is letting the drivers and intersection managers reach an agreement on an acceptable price for both agents. Similarly, intersection managers could form coalitions, aiming at improving the coalition's profit rather than the global one. Finally, the market can be enriched with different products offered by the intersection managers, such as discounts for usage at a particular time or daily subscriptions [12].

6. ACKNOWLEDGMENTS

This research was partially supported by the Spanish Ministry of Science and Education through projects "THOMAS" (TIN2006-14630-C03-02) and "AT" (CONSOLIDER CSD2007-0022, INGENIO 2010).

7. REFERENCES

[1] G. Balan and S. Luke. *History-based Traffic Control.* Proceedings of the 5th international joint conference on Autonomous Agents and Multiagent Systems, pp. 616-621, 2006.

[2] A. Bazzan and F. Klügl *Reducing the Effects of the Braess Paradox with Information Manipulation* Application of Agent Technology in Traffic and Transportation, pp. 85-97, Birkhäuser, 2005.

[3] A. Bazzan and R. Junges. *Congestion Tolls as Utility Alignment Between Agent and System Optimum.* Proceedings of the 5th international joint conference on Autonomous Agents and Multiagent Systems, pp. 126-128, 2006.

[4] A. Bazzan and F. Klügl *Multi-agent Architectures for Traffic and Transportation Engineering* IGI-Global, forthcoming, 2008.

[5] A. Bazzan. *Opportunities for Multiagent Systems and Multiagent Reinforcement Learning in Traffic Control* Autonomous Agents and Multiagent Systems, Springer, 2008.

[6] S. H. Clearwater. *Market-based Control: a Paradigm for Distributed Resource Allocation* World Scientific Publishing Co., Inc., 1996.

[7] K. M. Dresner and P. Stone. *A Multiagent Approach to Autonomous Intersection Management.* Journal of Artificial Intelligence Research, 31:591-656, 2008.

[8] J. Hernández and S. Ossowski and A. García-Serrano *Multiagent Architectures for Intelligent Traffic Management Systems.* Transportation Research Part C - Emerging Technologies, Vol. 10, No. 5., pp 473-506, 2002

[9] M. Jelasity and A. Montresor and O. Babaoglu. *Gossip-based Aggregation in Large Dynamic Networks.* ACM Transactions on Computer Systems, volume 23 number 3, pp. 219-252, 2005.

[10] R. Junges and A. Bazzan. *Evaluating the Performance of DCOP Algorithms in a Real World, Dynamic Problem.* Proceedings of the 7th International Joint Conference on Autonomous Agents and Multi-Agent Systems, pp. 599-606, 2008.

[11] L. Kuyer and S. Whiteson and B. Bakker and N. Vlassis. *Multiagent Reinforcement Learning for Urban Traffic Control using Coordination Graphs.* Proceedings of the 19th European Conference on Machine Learning, pp. 656-671, 2008.

[12] http://www.marketbasedcontrol.com/

[13] M. Meyer and E. J. Miller. *Urban Transportation Planning.* McGraw-Hill, 1984.

[14] K. Nagel and M. Schreckenberg. *A Cellular Automaton Model for Freeway Traffic.* J. Phys. I France 2, pp. 2221-2229, 1992.

[15] R. Rossetti and S. Bampi and R. Liu and D. Van Vliet and H. Cybis. *An Agent-based Framework for the Assessment of Drivers' Decision Making.* Proceedings of the 3rd Annual IEEE Conference on Intelligent Transportation Systems, pp. 387-392, 2000.

[16] T. Schwerdtfeger. *DYNEMO: A Model for the Simulation of Traffic Flow in Motorway Networks.* Proceedings of the 9th International Symposium on Transportation and Traffic Theory, pp. 65-87, 1984.

[17] K. Tumer and A. Agogino. *Distributed Agent-Based Air Traffic Flow Management.* Proceedings of the 6th International Joint Conference on Autonomous Agents and Multi-Agent Systems, pp. 330-337, 2007.

[18] C. Watkins. *Learning from Delayed Rewards* PhD Thesis, King's College, Cambridge, UK, 1989.

[19] B. Wie and R. Tobin. *Dynamic Congestion Pricing for General Traffic Networks* Transportation Research Part B, Vol. 32 No. 5, pp. 313-327, 1998.

[20] D. Wolpert and K. Tumer. *Optimal Payoff Functions for Members of Collectives.* Advances in Complex Systems, volume 04, issue 02, pp. 265-279, 2001.