Available online at www.sciencedirect.com
ScienceDirect
Procedia Social and Behavioral Sciences 2 (2010) 6002–6012
Procedia Social and Behavioral Sciences

The Sixth International Conference on City Logistics

Evaluating city logistics measures using a multi-agent model

Dai Tamagawa[a*], Eiichi Taniguchi[a], Tadashi Yamada[a]

[a]*Department of Urban Management, Kyoto University, Nishikyo-ku, Kyoto, 615-8540, Japan*

Abstract

This paper presents a methodology for evaluating city logistics measures considering the behaviour of several stakeholders associated with urban freight transport using a multi-agent model. The model constructed consists of a learning model and a model for vehicle routing and scheduling problem with time window-forecasted (VRP-TW-F). We used a method of Q-learning, a technique of reinforcement learning, in constructing a learning model. We implemented the model on a test road network representing an urban area. The results indicate that implementing a truck ban directly to environmentally damaged areas and discounting motorway tolls entirely in the urban motorway network together has large environmental effects, and leads to an acceptable environment for all stakeholders.

Keywords: City logistics; urban freight transport; stakeholder; multi-agent model; Q-learning

1. Introduction

It is very important to implement city logistics measures for effective and environmentally-friendly transport as trucks impose large negative impacts on the environment. There are many challenges when addressing urban freight transport problems since these problems are very complicated. One of these challenges is modelling urban freight transport activities considering several stakeholders associated with urban freight transport. There are several stakeholders associated with urban freight transport, thus it is necessary to consider the behaviour of these stakeholders in examining and evaluating city logistics measures (Davidsson et al., 2005, Taniguchi et al., 2007).

In this paper, we examined a methodology for evaluating city logistics measures considering the behaviour of several stakeholders associated with urban freight transport. For establishing the methodology, we constructed a multi-agent model. This model describes the stakeholders as independent agents.

Multi-agent modelling techniques have been used extensively in the transport and logistics area. Ossowski et al. (2005) presented a multi-agent based decision support system for transportation management. Wisetjindawat et al. (2005) proposed a simulation based multi-agent approach for modelling the interactions in freight movement. Jiao et

* Corresponding author. Tel.: +81-75-383-3140
E-mail address: d-tamagawa@kiban.kuciv.kyoto-u.ac.jp.

doi:10.1016/j.sbspro.2010.04.014

al. (2006) applied a multi-agent system to a manufacturing supply chain network. However, there are few studies that implement multi-agent models from the view point of city logistics.

2. Multi-agent Model

2.1. Stakeholders associated with urban freight transport

We considered five stakeholders, freight carriers, shippers, residents, administrators and motorway operators in urban area. We assumed that they had their own objectives and they selected their behaviour to achieve their objectives. When city logistics measures were implemented and the environments were changed, stakeholders would change their behaviour to adapt to their changed environment. Thus, we thought that stakeholders had to evaluate the environments they were found, and assumed that they had their own objectives for evaluating the environments and they selected their behaviour to enhance their satisfaction measured against their own objectives. The objectives and behaviour of stakeholders we assumed are as follows.

The objective of freight carriers is to maximise transport profit, and their behaviour is offering transport charges to shippers and delivering goods of shippers. The objective of shippers is to minimise transport costs paid to freight carriers, and their behaviour is to select freight carriers and request them to deliver the goods. We considered that transport cost consisted of transport charges paid to freight carriers and opportunity costs. Opportunity costs are generated when goods were delivered late for designated time windows. The objective of residents is that NO_x emissions from vehicles are kept under an environmental limit, and their behaviour is to select whether they should make a complaint to administrators or not when the NO_x emissions in their local area exceeded the environmental limit. The objective of administrators is to minimise complaints from residents concerning NO_x emissions, and their behaviour is to determine whether they should implement some city logistics measures or not in the areas that residents make a complaint. The objective of motorway operators is to maximise their toll revenue, and their behaviour is changing the motorway toll. Reflecting these behaviours, interactions among stakeholders could be described as below.

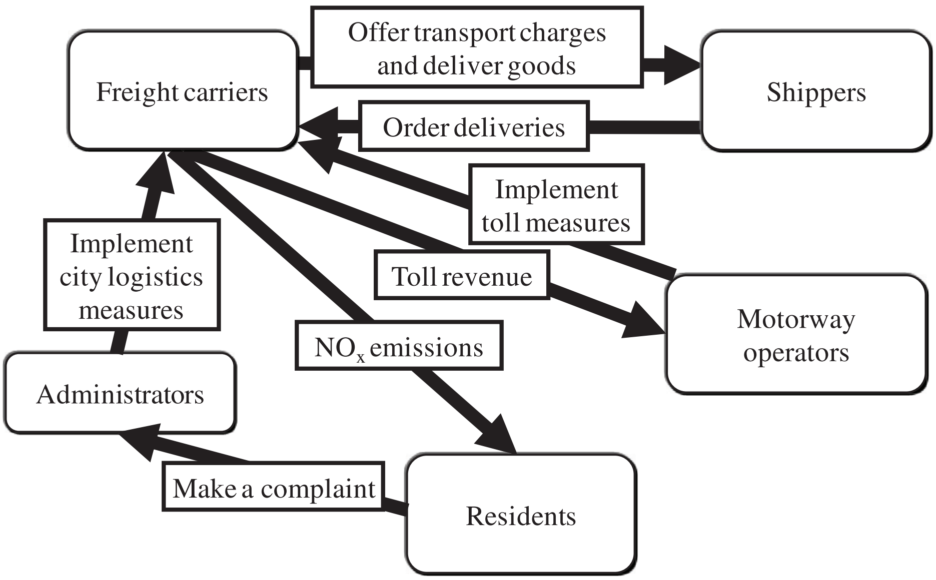

Figure 1 Interactions among stakeholders

2.2. Framework of the model

Figure 2 shows a framework of the model constructed. The model consists of two sub-models. One is the learning model for stakeholders, and the other is the model for vehicle routing and scheduling problem with time window forecasted (VRP-TW-F) (Taniguchi et al., 2001). The learning model evaluates the behaviour of stakeholders and learns the value of them, and selects the behaviour considering the value of them. VRP-TW-F model plans and implements delivery schedules of trucks for each freight carrier. These two models are executed alternately. The flow of calculations of these models is as follows.

(STEP1) Learning model determines and implements the behaviour of stakeholders. In this step, freight carriers offer transport charges to shippers, and shippers select freight carriers to order deliveries considering the charges offered.
(STEP2) VRP-TW-F model plans and implements delivery schedules of trucks for each freight carrier according to the orders from shippers.
(STEP3)The amount of NO_x emissions for each area and toll revenue from motorways are calculated after the deliveries by freight carriers, and the network environment is updated.
(STEP4) Factors of the updated environment are fed-back to the learning model.
(STEP5) The learning model evaluates the behaviour of stakeholders according to the updated environment, and the experiences of stakeholders are updated. After that, go back to (STEP1).

By the iteration between the two sub-models, stakeholders learn their favourable behaviour considering the interactions among stakeholders. Several freight carriers exist in the network and in general that there are several individuals for each type of stakeholder and they learn and behave independently. In this multi-agent model, each individual can learn independently.

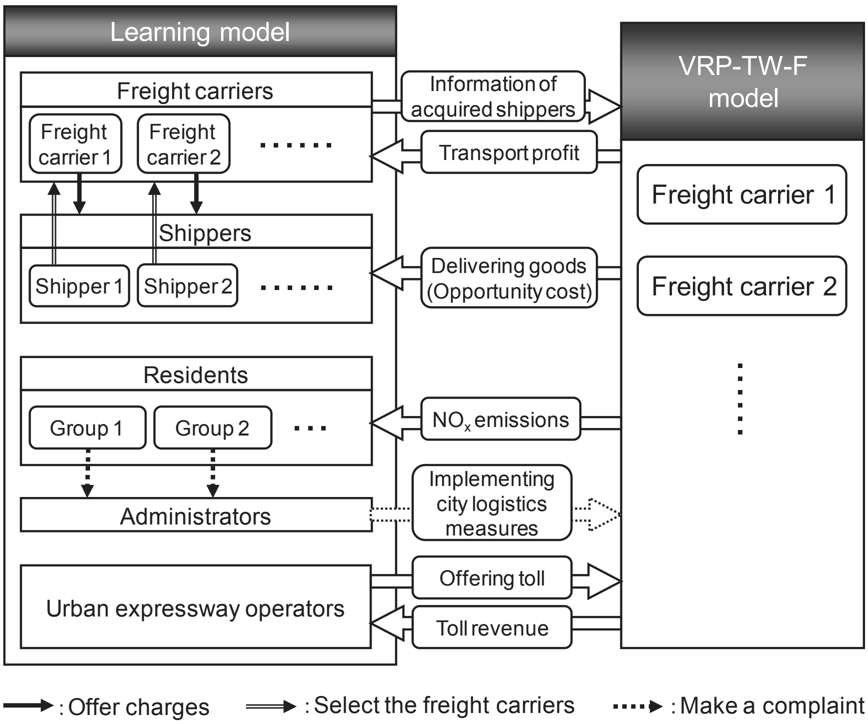

Figure 2 Framework of the multi-agent model

3. VRP-TW-F Model

VRP-TW-F model determines the optimal solution by minimising the total transport cost of freight carriers. The total transport cost comprises three components;

(i) fixed cost of vehicles,
(ii) vehicle operating costs that are proportional to the time travelled, and
(iii) early arrival and delay penalties for designated pickup/delivery time windows at customers.

The model can be formulated as follows:

$$\min \quad C(\boldsymbol{t}_0, \mathbf{X}) = \sum_{l=1}^{m} c_{f,l} \cdot \delta_l(\mathbf{x}_l) + \sum_{l=1}^{m} C_{t,l}(t_{l,0}, \mathbf{X}_l) + \sum_{l=1}^{m} C_{p,l}(t_{l,0}, \mathbf{X}_l) \tag{1}$$

where,

$C(\boldsymbol{t}_0, \mathbf{X})$	: total cost (yen)
$\boldsymbol{t}_0$	: departure time vector for all vehicles from the depot; $\boldsymbol{t}_0 = \{t_{l,0} \mid l = 1, m\}$
$\mathbf{X}$	: assignment and order of visiting customers for all vehicles; $\mathbf{X} = \{X_l \mid l = 1, m\}$
X_l	: assignment and order of visiting customers for vehicle l; $X_l = \{n(i), \mid i = 1, N_l\}$
$n(i)$	: node number of i th customer visited by a vehicle
N_l	: total number of customers visited by vehicle l
m	: maximum number of vehicles available
$c_{f,l}$	: fixed cost for vehicle l (yen)
$\delta_l(\mathbf{x}_l)$	: =1; if vehicle l is used, =0; otherwise
$C_{t,l}(t_{l,0}, \mathbf{X}_l)$	: operating cost for vehicle l (yen)
$C_{p,l}(t_{l,0}, \mathbf{X}_l)$	: penalty cost for vehicle l (yen)

The problem described here is a NP-hard (Non-deterministic Polynomial-hard) combinatorial optimisation problem. Thus, some heuristic algorithms are required to identify good solutions. The model described here uses Genetic Algorithms (GA) to solve the problem.

4. Learning Model

4.1. Construction of the model

If the environment that stakeholders are in is changed by implementing some city logistics measures, they will change their behaviour to adapt to the changed environment. Therefore, their behaviour should be modelled using a learning method. Several methods of learning by agents have been introduced. It could be thought that stakeholders associated with urban freight transport do not know their optimal behaviour. Thus they have to find their optimal behaviour by trial and error. The learning method that agents use to find their optimal behaviour by trial and error is defined as unsupervised learning, and reinforcement learning (Sutton and Barto, 1998). As there are several techniques of reinforcement learning, we selected two techniques of Q-learning (Watkins and Dayan, 1992) and Monte Carlo method. We constructed the learning model using both techniques, and compare the performance of them.

Q-learning updates a value of action-value functions for agents. In updating, Q-learning uses the maximum value of the action-value function. For example, the updating formula by Q-learning for freight carriers is as follows. An expected transport profit is used as the value of the action-value function for freight carriers.

$$Q_f(s_{f,t}, a_{f,t}) \leftarrow Q_f(s_{f,t}, a_{f,t}) + \alpha_f \left[r_{f,t} + \gamma_f \max_{a_{f,t+1} \in A_f} Q_f(s_{f,t+1}, a_{f,t+1}) - Q_f(s_{f,t}, a_{f,t}) \right] \tag{2}$$

where,

$$r_{f,t} = fee_t \times n_t - lc_t \tag{3}$$

s.t. $\forall s_{f,t} \in S_f$

$\forall a_{f,t} \in A_f$

where,

$Q_f(s_{f,t}, a_{f,t})$: expected total transport profit obtained from state $s_{f,t}$ to the last state when freight carrier selected the behaviour $a_{f,t}$ in state $s_{f,t}$

α_f : learning rate for freight carrier

$r_{f,t}$: actual transport profit obtained in state $s_{f,t}$ when freight carrier selected behaviour $a_{f,t}$ in state $s_{f,t}$

γ_f : discount rate for freight carrier

fee_t : charge that freight carrier offered to shippers in state $s_{f,t}$

n_t : number of obtained shippers when freight carrier selected behaviour $a_{f,t}$ in state $s_{f,t}$

lc_t : transport cost when freight carrier selected behaviour $a_{f,t}$ in state $s_{f,t}$

S_f : set of states for freight carrier

A_f : set of behaviours for freight carrier

The Monte Carlo method also updates the value of action-value functions for agents. When updating, the Monte Carlo method uses a reward that was actually received. The updating formula by the Monte Carlo method for freight carriers is as follows. Meanings of the variables are same as in Q-learning.

$$Q_f(s_{f,t}, a_{f,t}) \leftarrow Q_f(s_{f,t}, a_{f,t}) + \alpha_f \left[R_{f,t} - Q_f(s_{f,t}, a_{f,t}) \right] \tag{4}$$

where,

$$\sum_{\tau=0}^{T-t} \gamma_f^{\tau} \times r_{f,t+\tau} \tag{5}$$

$$r_{f,t+\tau} = fee_{t+\tau} \times n_{t+\tau} - lc_{t+\tau} \tag{6}$$

s.t. $\forall s_{f,t} \in S_f$

$\forall a_{f,t} \in A_f$

where,

$R_{f,t}$: total discounted transport profit from state $s_{f,t}$ to the last state when freight carrier selected behaviour $a_{f,t}$ in state $s_{f,t}$

T : last state in one episode

In constructing the learning model by using reinforcement learning, we have to define, “episode” and “state”. Episode means one term of learning period and consists of several states. In this paper, we defined one state was one day and one episode was one month (30days). Also, not to select limited behaviour and not to generate only an inefficient environment as a result, we considered that agents selected the most valuable behaviour at a rate of 80% and selected randomly at a rate of 20%.

4.2. Comparison between two techniques

We compared the performance of the two techniques to decide which should be adopted as the learning model. In comparing, we considered that only freight carriers and shippers learned and changed their behaviour in order to simplify the problem. The results of the simulation of the case studies using the test road network is presented in next section.

Figure 3 and Figure 4 show the transition of transport profit of freight carriers and transport cost of shippers. Both indicate that the results calculated by Q-learning are lower than the results by the Monte Carlo method. In these models, freight carriers offer the charges at first, and after that, shippers select the freight carriers in considering the offered charges, so it could be considered that shippers have advantages over freight carriers. In consequence, we can mention that more efficient learning is realized using Q-learning because shippers have more benefit when using Q-learning. Thus we decided to use the method of Q-learning for the learning model.

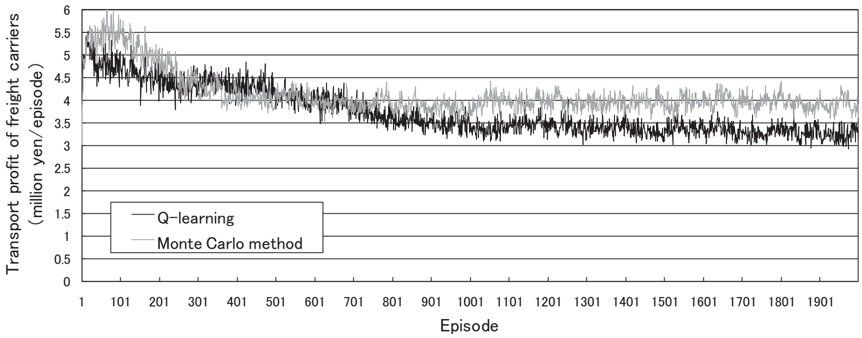

Figure 3 Transition of transport profit of freight carriers

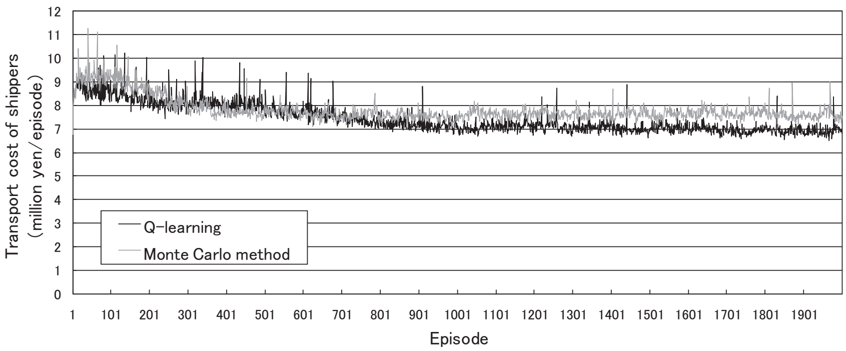

Figure 4 Transition of transport cost of shippers

5. Case Studies

5.1. Test conditions

Figure 5 shows the test road network used in this study. This network is assumed to be an urban area, and consists of 25 nodes and 104 links including 24 motorway links. We divided one day into eight periods of time, and set the travel times of links in each period. There are four freight carriers in the network and they have their own depot. Each freight carrier has three trucks (2 ton truck, 4 ton truck, and 10 ton truck) and can use them flexibly. Freight carriers can select their transport charges from 10,000 yen to 20,000 yen at an interval of 1,000 yen. The location of depots and shippers are also shown in Figure 5. Each shipper has goods of 1,000 kg and has their own time window for arrivals

There is only one kind of administrator and one kind of motorway operator in the network. We set the motorway toll for medium and small trucks (including 2 ton trucks and 4 ton trucks) to be 40 yen per kilometre, and 80 yen per kilometre for large trucks (including 10 ton trucks). We defined an area between two nodes that were connected by one link directly as a zone unit. Residents exist at all zones in the network, thus there are 40 units of residents in the network. We assumed that residents would select whether they make a complaint to administrators or not when NO_x emissions in their zones exceeded the environmental limit.

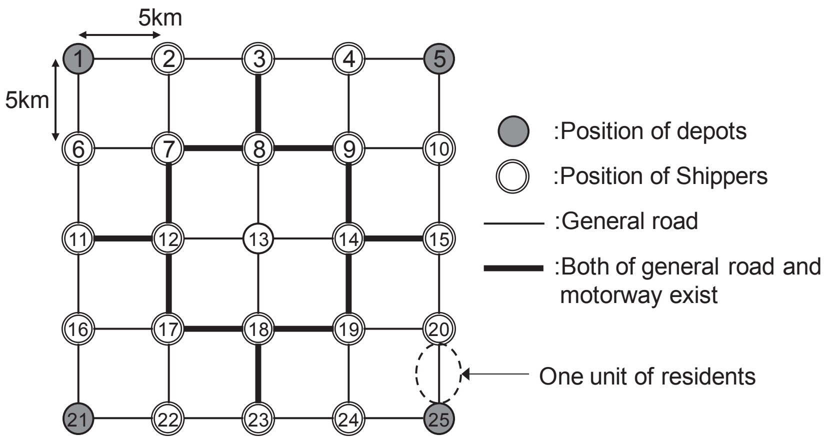

Figure 5 Test road network

5.2. Results

5.2.1. The case only city logistics measures by administrators were implemented

At first, we considered the case where freight carriers, shippers, residents and administrators were learning agents and motorway operators did not learn and change their behaviour (did not change the motorway toll). As city logistics measures by administrators, we considered road pricing for all trucks and truck bans for 10 ton trucks. City logistics measures are implemented only on general roads and not on motorways. We considered that administrators selected whether they would implement these measures or not in the zones that residents had complained. When NO_x emissions in a zone reduced under the environmental limit, residents in that zone would stop making complaints and administrators also would stop implementing the measures.

Figure 6 shows the performance of road pricing for all trucks by administrators. We assumed that the case where no measures were implemented was the base case and examined the performance for each objective by comparing it with the base case. As the evaluation index for residents, we adopted the number of zones that NO_x emissions exceeded the environmental limit in order to evaluate the effects on all residents in the network. Also, as the number of zones that residents made a complaint was similar to the number of zones that NO_x emissions exceeded the environmental limit, we adopted total NO_x emissions in the network instead of the number of zones residents complained. As for the transport cost, total NO_x emissions, and the number of zones that NO_x emissions exceeded the environmental limit, we presented an inverse of the proportion of each case to base case as the performance of each case. On the other hand, the transport profit of freight carriers and toll revenue from the motorway were presented as a proportion of the base case. Therefore, in this figure, the measure is evaluated effective when the value is bigger than 1.0. The results are presented as an average during the period after the network environment became stable. It could be confirmed that road pricing was not so effective for all stakeholders.

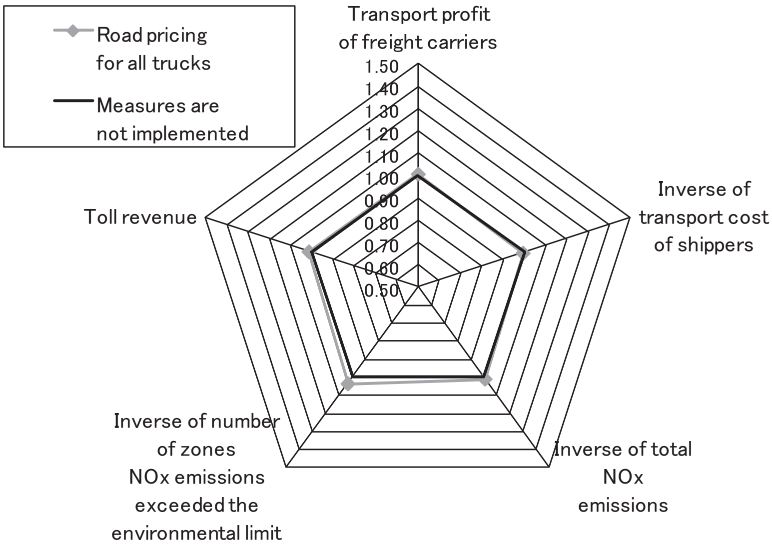

Figure 6 Performance of road pricing for all trucks by administrators

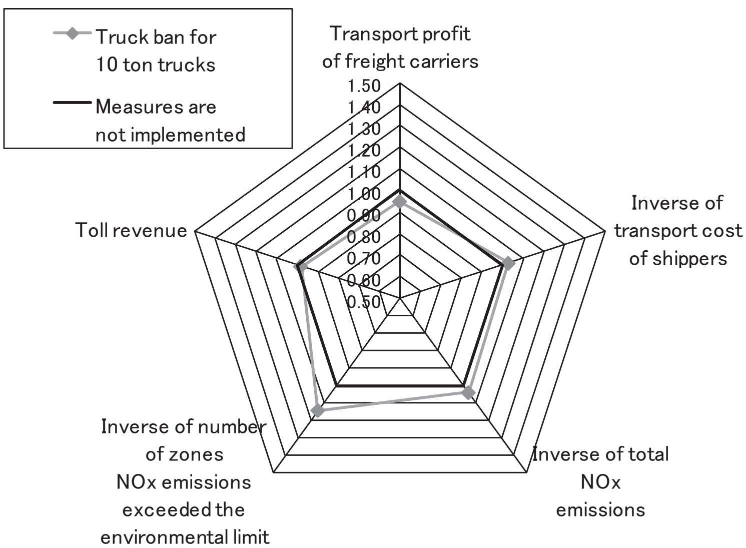

Figure 7 Performance of truck ban for 10 ton trucks by administrators

Figure 7 shows the performance of the truck ban for 10 ton trucks. Total NO_x emissions and the number of zones that NO_x emissions exceeded the environmental limit were reduced. This is because freight carriers refrained from using 10 ton trucks according to implementing truck ban, and NO_x emissions reduced as a result. We can see that the effect of reducing the number of zones that NO_x emissions exceeded the environmental limit is larger than the effect of reducing total NO_x emissions in the network. In this study, the truck ban was implemented only to the zones where NO_x emissions exceeded the environmental limit and residents made a complaint, so this measure was not effective for zones of moderate NO_x emissions. Thus the effect of reducing total NO_x emissions in the network including the zones of moderate NO_x emissions tends to be not so large compared to the effect of reducing the number of zones where NO_x emissions exceeded the environmental limit.

5.2.2. The case where city logistics measures by administrators and motorway toll changes were implemented together

In this case, we considered that motorway operators also learned and changed their behaviour as well as other stakeholders. Figure 8 shows the performance of the case that road pricing by administrators and changes of the motorway toll by motorway operators were implemented together. Figure 9 shows the performance of the case that truck ban by administrators and motorway toll change were implemented together. Compared to the base case, toll revenue was increased and total NO_x emissions and the number of zones that NO_x emissions exceeded the environmental limit were reduced. These indexes were improved also compared to the case that only city logistics measures by administrators were implemented. Figure 10 shows the transition of the average rate of change of the motorway toll. The motorway toll was discounted around 15% finally.

From these results, we could consider that motorway operators discounted the toll to increase the toll revenue and freight carriers used motorways more frequently according to the discount of the toll. Also, the average speed of trucks increased according to the increase of motorway use, and as a result, total NO_x emissions and the number of zones where NO_x emissions were exceeded the environmental limit were reduced. Thus we could confirm that effects of city logistics measures implemented directly to the environmentally damaged areas and effects of discounting the motorway toll entirely in the urban motorway network did not compete against each other, and the environmental effects increased.

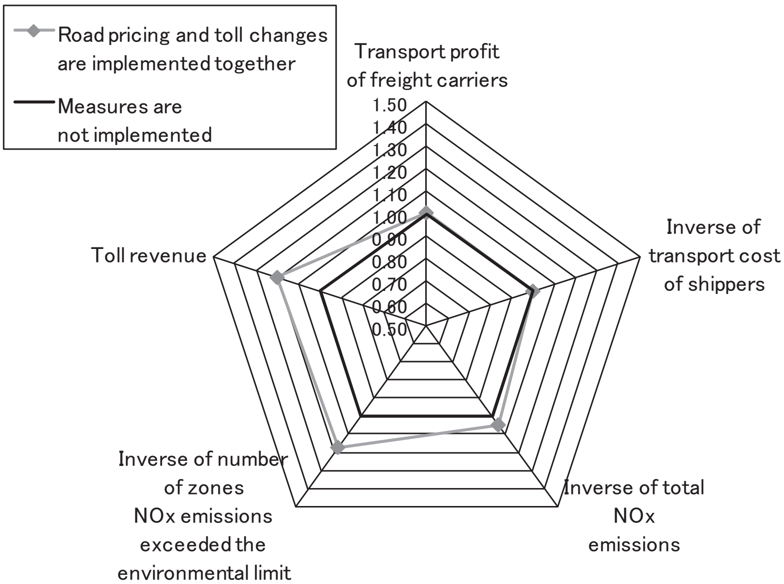

Figure 8 Performance of the case that road pricing for all trucks and changes of the motorway toll were implemented together

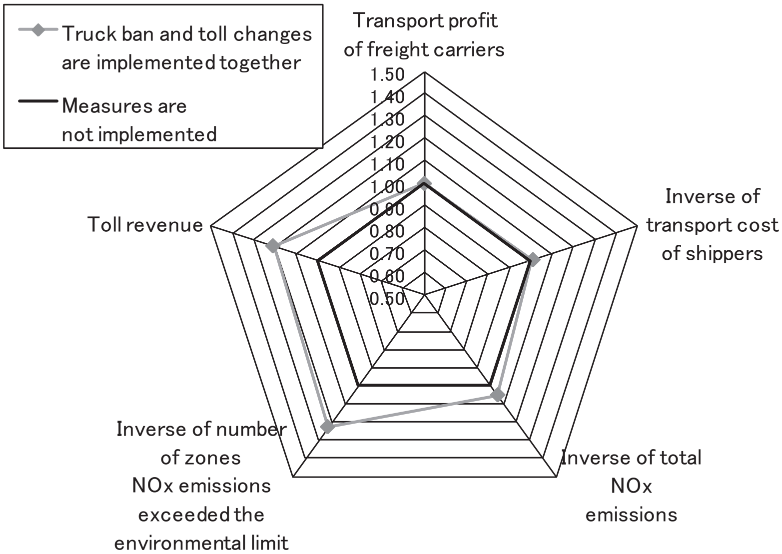

Figure 9 Performance of the case that truck ban for 10 ton trucks and changes of the motorway toll were implemented together

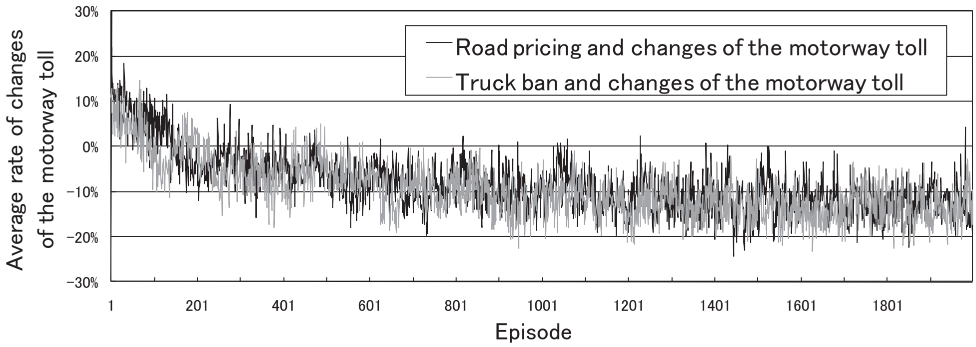

Figure 10 Transition of average rate of changes of the motorway toll

Also, the transport profit of freight carriers and the transport cost of shippers did not change considerably compared to the base case. Therefore, it could be mentioned that these results were acceptable for all stakeholders. However, the increase in toll revenue seemed too large, so we have to consider improving route choice model for truck drivers.

6. Conclusion

In this paper, we presented a methodology for evaluating city logistics measures considering the behaviour of several stakeholders associated with urban freight transport using a multi-agent model. We constructed a multi-agent model that considered each stakeholder as an independent agent. The model consisted of VRP-TW-F model and the learning model. The learning model was constructed by using the method of Q-learning. We applied the model to test road network, and implemented several city logistics measures. The results indicated that effects of city logistics measures implemented directly to the environmentally damaged areas and effects of discounting the motorway toll entirely in the urban motorway network did not compete against each other, and environmental effects increased compared to the case where only city logistics measures were implemented. As a result, an acceptable network environment for all stakeholders was generated. However, the increase in toll revenue by discounting the motorway toll seemed to be too large, thus we have to consider improving the route choice model for truck drivers.

References

Davidsson, P., Henesey, L., Ramstedt, L., Tornquist, J., & Wernstedt, F. (2005). An analysis of agent-based approaches to transport logistics. *Transportation Research Part C,* 13, 255-271.

Jiao, J., You, X., & Kumar, A. (2006). An agent-based framework for collaborative negotiation in the global manufacturing supply chain network. *Robotics and Computer-Integrated Manufacturing,* 22, 239-255.

Ossowski, S., Hernandez, J. Z., Belmonte, M-V., Fernandez, A., Garcia-Serrano, A., Perez-de-la-Cruz, J-L., Serrano, J-M., & Triguero, F. (2005). Decision support for traffic management based on organisational and communicative multiagent abstractions. *Transportation research Part C*, 13, 272-298.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, Massachusetts: The MIT Press.

Taniguchi, E., Thompson, R. G., Yamada, T., & Duin, R. (2001). *City logistics - Network modelling and intelligent transport systems*. Oxford: Pergamon.

Taniguchi, E., Yamada, T., & Okamoto, M. (2007). Multi-agent modelling for evaluating dynamic vehicle routing and scheduling systems. *Journal of the Eastern Asia Society for Transportation Studies*, 7, 933-948.

Watkins, C. J. C. H., & Dayan, P. (1992). Q-Learning. *Machine Learning*, 8, 279-292.

Wisetjindawat, W., Sano, K., & Matsumoto, S. (2005). Supply chain simulation for modeling the interactions in freight movement. *Journal of the Eastern Asia Society for Transportation Studies*, 6, 2991-3004.