MULTI-AGENT MODELLING FOR EVALUATING DYNAMIC
VEHICLE ROUTING AND SCHEDULING SYSTEMS

Eiichi TANIGUCHI
Professor
Department of Urban Management
Kyoto University
Kyotodaigaku-katsura, Nishikyo-ku, Kyoto
615-8540 Japan
Fax: +81-75-950-3800
E-mail: taniguchi@kiban.kuciv.kyoto-u.ac.jp

Tadashi YAMADA
Associate Professor
Department of Urban Management
Kyoto University
Kyotodaigaku-katsura, Nishikyo-ku, Kyoto
615-8540 Japan
Fax: +81-75-950-3800
E-mail: t.yamada@kiban.kuciv.kyoto-u.ac.jp

Masayuki OKAMOTO
Hiroshima Port and Airport Development
Office, Chugoku Regional Development
Bureau, Ministry of Land, Infrastructure and
Transport
3-10-28 Ujinakaigan, Minami-ku, Hiroshima
734-0011, Japan
Fax: +81-82-505-0107
E-mail: okamoto-m87pa@pa.cgr.mlit.go.jp

Abstract: This paper presents multi-agent models for evaluating the behaviour and interaction among stakeholders who are involved in urban freight transport systems as well as effects of city logistics measures. Multi-agent simulation on a small test road network demonstrated that the VRPTW-D model which dynamically adjusted vehicle routing planning to the current travel times generated good performance in terms of increasing profits for freight carriers and decreasing costs for shippers. After applying multi-agent models on a large test road network, it was observed that introducing the VRPTW-D model generated a win-win situation by increasing profits for freight carriers and decreasing the costs for shippers. The results also show that the implementation of road pricing can reduce NOx emissions but may increase the costs for shippers. To avoid such effects, introducing co-operative freight transport systems helps shippers to reduce their costs.

Key Words: City logistics, multi-agent model, vehicle routing and scheduling

1. INTRODUCTION

Recently urban freight transport has become an important social issue in terms of the increasing level of traffic congestion, negative impacts on the environment, traffic safety as well as the energy consumption. On the other hand, private companies who are involved in urban freight transport are requested to reduce their inventory and transport costs for surviving in the competitive global market. To cope with these complicated problems, the concepts and measures for city logistics including co-operative delivery systems, public logistics terminals, load factor controls and use of ICT (Information and Communication Technology) have been proposed and implemented (e.g. Taniguchi et al. (2001), Crainic et al. (2004). Anderson et al. (2005)) for sustainable urban freight transport systems.

There are four major stakeholders relating to city logistics; (a) Shippers, (b) Freight carriers, (c) Residents and (d) Administrators. These stakeholders have different objectives and different types of behaviour. Shippers try to minimise their costs in supply chains. Freight
carriers try to meet shippers requests to collect and deliver goods within strict time windows. Residents want quiet, noiseless atmosphere and clean air in their community. Finally administrators hope to activate the vitality of city with sustainable transport systems. Understanding the behaviour of the stakeholders and interaction among them is needed for evaluating city logistics measures before implementing them.

Multi-agent modelling techniques allow complicated urban freight transport systems with multiple actors to be investigated (e.g. Weiss, 1999; Ferber, 1999; Wooldridge, 2002). Multi-agent models generally deal with behaviour and interaction among multiple agents, which are most suitable to understand and study the behaviour of stakeholders in urban freight transport systems and their response to policy measures. Davidsson et al. (2005) provided a survey of existing research on agent-based approaches in freight transport and noted that agent-based approaches seem very suitable for this domain. Duin et al. (1998) showed dynamic actor network analysis for complex logistics problems. Ossowski et al. (2005) presented multi-agent approaches to decision support systems in traffic management. Jiao et al. (2006) presented an agent based framework in the global manufacturing supply chain network. The literature shows a number interesting examples of multi-agent approaches to transport logistics problems but not directly focusing on urban freight transport systems.

This paper tries to present a multi-agent approach using reinforcement learning in urban freight transport systems taking into account the behaviour of multiple stakeholders. In particular, it focuses on the effects of using the VRPTW-D (Vehicle Routing and scheduling Problem with Time Windows-Dynamic) model with real time travel time information on the road network. Moreover, evaluation of city logistics measures such as road pricing and cooperative freight transport systems will be discussed in the multi-agent simulation.

2. MULTI-AGENT MODELS

2.1 Overview
Multi-agent models contain multiple and interacting agents within the model, whom take action for solving their own problems. This paper incorporates multi-agent learning models which assume that an agent pursues its own learning goal using their knowledge and communication with other agents. The important feature of multi-agent models is that more emphasis is put on the ability, characteristics and function of individual agents rather than system optimisation. An agent in multi-agent models is an autonomous agent who can perceive the environment and make their own decision to act for satisfying their needs and attaining their objectives. These agents can communicate with each other and the interaction amongst agents is most interesting to observe during the simulation. The multi-agent learning models in this paper assume that each agent takes actions based on their reinforcement learning (Sutton and Barto, 1998) to maximise its action-value function. It is interesting to examine what effects can be given to other agent's action through the interaction of agents.

2.2 Behaviour of Agents
There are assumed to be three types of agents in the multi-agent models adopted in this paper; (a) freight carriers, (b) shippers and (c) administrators. The multi-agent models herewith allow communication between different agents such as shipper to freight carrier, freight carrier to administrator and administrator to shipper, but does not allow communication between same agent such as freight carrier to freight carrier, shipper to shipper and
administrator to administrator. Each agent takes action to maximise its action-value function and does not consider collaboration among agents. Figure 1 shows the interaction among three agents.

![Figure 1 Interaction among agents](image)

2.2.1 Freight carriers
Let us examine the behaviour of each agent. Freight carriers aim at maximising their profits and their policy to propose the price for collecting goods from shippers for maximising their profits. This time only collecting goods is considered. The profits of freight carriers can be equal to the price of collecting goods which is paid by the shipper minus the cost of collecting goods from shippers. The cost of collecting goods is composed of fixed costs of owning trucks, operation costs which depend on the running times of trucks as well as early arrival and delay penalties at customers (shippers). The behaviour of freight carriers can be formulated as follows.

\[
\begin{align*}
\text{maximise} & \quad E[B(f_i)] = f_i \cdot E[n_i(f_i)] - E[C_{c,i}] \\
\text{subject to} & \quad 0 \leq f_i \leq f_{\text{max}} \\
& \quad 0 \leq E[n_i(f_i)] \leq n_{\text{max}}
\end{align*}
\]

where,

- \(E[B(f_i)]\): expected profits of freight carrier \(i\) with the proposed price of \(f_i\)
- \(f_i\): proposed price of freight carrier \(i\)
- \(f_{\text{max}}\): maximum price which freight carrier \(i\) can propose
- \(E[n_i(f_i)]\): expected number of shippers which freight carrier \(i\) can obtain with the proposed price of \(f_i\)
\[ E[C_{c,i}] : \] expected costs of freight carrier \( i \)

\( n_{\text{max}} : \) maximum number of shippers.

Freight carriers are assumed to operate their pickup trucks following optimal routing which can be obtained using two models. One is the VRPTW-F (Vehicle Routing and scheduling Problem with Time Windows - Forecasted) model and the other is VRPTW-D (Vehicle Routing and scheduling Problem with Time Windows- Dynamic) model (Taniguchi and Shimamoto, 2004). The VRPTW-D model is defined as follows, a depot and a number of customers are defined for each freight carrier. A fleet of identical vehicles collects goods from customers and delivers them to the depot or delivers goods to customers from the depot. For each customer a designated time window, specifying the desired time period to be visited is also specified. For example, in the case of collecting goods, vehicles depart from the depot and visit a subset of customers for picking up goods in sequence and return to the depot to unload them. A vehicle is allowed to make multiple trips per day. Each customer must be assigned to exactly one route of a vehicle and all the goods from each customer must be loaded onto the vehicle at the same time. The total weight of the goods in a route must not exceed the capacity of the vehicle. This problem is used to determine the optimal assignment of vehicles to customers and the departure time as well as the order of visiting customers for a freight carrier. The VRPTW-D model explicitly incorporates real time information on travel times for identifying the optimal routes and departure times of vehicles, whereas the VRPTW-F model does not take this into account.

Firstly, the VRPTW-F model needs link travel times which a dynamic traffic simulation provides using the block density method. Then vehicle routing and scheduling planning is undertaken on the previous day to identify the optimal departure time from a depot and visiting order of customers. On the current day vehicle operations are executed without changing the visiting order of customers and the road used.

Secondly, the VRPTW-D model is based on the same vehicle routing and scheduling plan that was undertaken on the previous day. However, plans are revised using real time information of present link travel times, whenever a vehicle arrives at a customer. This real time information is provided by dynamic traffic simulation based on the current conditions of the day. The VRPTW-D model therefore uses more accurate travel time data than the VRPTW-F model. In particular if a road was blocked due to a crash or other traffic impediments, the VRPTW-D model can avoid congested roads to give lower cost vehicle routing and scheduling plans than the VRPTW-F model.

Freight carriers operate their trucks based on the optimal assignment of vehicles to customers and the departure time as well as the order of visiting customers which are identified by the VRPTW-F model or the VRPTW-D model. The only difference of these two models is that the VRPTW-F model only considers the one value of mean travel time on a link, whereas the VRPTW-D model uses the real time travel time on a link.

2.2.2 Shippers

Shippers aim at minimising their costs. Their policy is to choose freight carriers such that minimise the shippers’ costs. It is assumed that shippers’ costs include the fee for collecting goods that they pay to freight carriers as well as opportunity costs of delay arrivals of trucks at shippers. Opportunity costs can be estimated by multiplying the opportunity costs per unit time due to delay arrivals of trucks and the delay times of trucks at shippers. The opportunity
costs of delay arrivals for shippers are different from the delay penalties that are assumed to
be paid by freight carriers to shippers if pickup trucks arrive late at shippers over the end of
time window. The behaviour of shippers can be formulated as follows.

\[
\text{minimise} \quad E[C_{s,j}] = f_i + c_{d,j} \cdot E[t_{d,j}]
\]

where,

- \( E[C_{s,j}] \): expected costs of shipper \( j \)
- \( c_{d,j} \): opportunity costs per unit time due to delay arrivals of trucks of shipper \( j \)
- \( E[t_{d,j}] \): expected delay time of truck of freight carrier \( i \) at shipper \( j \)

2.2.3 Administrators

Administrators aim to reduce negative impacts on the environment. This study assumes that
administrators try to reduce NOx emissions on each link of road network under the required
level. Their policy in this study is to implement three policy measures including (a) road
pricing, (b) providing subsidies to shippers if they choose environmentally friendly freight
carriers and (c) implementing co-operative freight transport systems. Administrators can
estimate NOx emissions using Equation 5, for identifying the link where they implement road
pricing measures and also determining environmentally friendly freight carriers whom
shippers can choose to obtain subsidies.

\[
e_{a,a} = d_a (k_1 \cdot v_m^2 - k_2 \cdot v_m + k_3)
\]

where,

- \( e_{a,a} \): NOx emissions at link \( a \)
- \( d_a \): distance of link \( a \)
- \( v_m \): running velocity of truck
- \( k_1, k_2, k_3 \): constants

2.3 Learning of Agents

Agents in this study are assumed to have a reinforcement learning mechanism. Reinforcement
learning is learning without supervision and agents try to maximise (or minimise) their
expected value which is feedback from an agent’s action based on a policy. The multi-agent
learning herewith adopts the following expected values: (a) For freight carriers; the expected
number of shippers which can be obtained by proposing a price for collecting goods, (b) For
shippers; the expected shipper’s cost corresponding to a proposed price for collecting goods,
(c) For administrators; the expected NOx emissions on a link as well as the expected NOx
emissions of each freight carrier. Distributed learning is adopted in this study, which means
each agent learns independently. Each agent needs to update the expected value after
receiving feedback from an action and iteration will go on.

Learning is based on the Monte Carlo method in this study, which is classified in non-
bootstrapping type learning (Sutton and Barto, 1998). The Monte Carlo method determines
the action-value function based on policy \( \pi \) assigning the action-value function of previous
iteration and the reward given in the previous iteration using learning rate. The formulation
can be given.

\[ Q_\pi(t + 1) = \alpha \cdot r_\pi(t) + (1 - \alpha) \cdot Q_\pi(t) \]  

(6)

where,

- \( Q_\pi(t) \): action-value function of policy \( \pi \) at time \( t \)
- \( r_\pi(t) \): reward of policy \( \pi \) at time \( t \)
- \( \alpha \): learning rate

The reward of policy \( \pi \) at time \( t \) was taken as follows: (a) For freight carriers---the expected profits which is given by equation (1), (b) For shippers---the expected opportunity cost due to delay arrivals of trucks which is given by the second term of equation (4) and (c) For administrators---the estimated NOx emissions which is given by equation (5), respectively. The action-value function can be updated using equation (6), assuming that the action-value function was 0 for \( t = 1 \).

Overall, the framework of the multi-agent simulation is that each agent behaves as an autonomous agent who seeks for its own objective by taking action to maximise or minimise the expected value described above. The interaction among agents can be simulated during the iteration in the time domain by considering the effects of actions taken by other agents.

3. CASE STUDY A

Case study A focuses on validating the multi-agent models in comparison with game theory as well as efficiency of the VRPTW-D model on a small test road network. Figure 2 shows a small test road network where two freight carriers A and B and two shippers 1 and 2 were located. The travel times on solid links in Figure 2 are 40 min. and on dotted links 60 min. without any variation of link travel times at the first stage and then some variation of travel times will be considered in the second stage with the variance of 20 min\(^2\) and 40 min\(^2\). Tables 1 and 2 indicate the costs for freight carriers and shippers. It is assumed that the demand of each shipper can be satisfied by a single visit of truck and a truck can load goods from two shippers.

<table>
<thead>
<tr>
<th>Shipper</th>
<th>Delay penalty (unit/min.)</th>
<th>Fixed cost (unit/vehicle)</th>
<th>Operation cost (unit/min.)</th>
<th>Delay penalty (unit/min.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.3</td>
<td>5</td>
<td>0.05</td>
<td>0.2</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A comparison of game theory and multi-agent modelling was performed. In this game there is no difference in link travel times between freight carriers and shippers. For long time trials, the expected cost of shippers given by equation (4) becomes equal for each freight carrier. Therefore, shippers choose freight carriers only based on the proposed price. On the other hand, each freight carrier can get a shipper’s job by proposing a lower price than that of the competitor.

Table 3 shows the pay-off matrix for the non-cooperative game. Here, freight carriers propose the price for collecting goods at the interval of 6 units (e.g. 60, 54, 48, 42…). The Nash
equilibrium of this game is given: (proposed price of freight carrier A, that of B) = (12, 12) and (18, 18). Figure 3 illustrates the change of proposed price of freight carriers A and B for 5,000-day iterations using multi-agent modelling. This figure indicates that in the case of a learning rate of 0.2, the proposed price did not reach the Nash equilibrium, but in case of learning rate of over 0.4, the proposed price reached the Nash equilibrium of (12, 12). In case of a learning rate of 0.3 it reached Nash equilibrium of (18, 18). Therefore, for the static case without changes in link travel times, results of multi-agent models using reinforcement learning with appropriate learning rate converged into equilibrium point of game theory after a large number of iterations.

![Small test road network A](image)

**Table 3 Pay-off matrix for a non-cooperative game**

<table>
<thead>
<tr>
<th></th>
<th>6</th>
<th>12</th>
<th>18</th>
<th>24</th>
<th>30</th>
<th>36</th>
<th>42</th>
<th>48</th>
<th>54</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed price of freight carrier A (unit)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-8</td>
<td>-8</td>
<td>-8</td>
<td>-8</td>
<td>-8</td>
<td>-8</td>
<td>-8</td>
<td>-8</td>
<td>-8</td>
<td>-8</td>
</tr>
<tr>
<td>12</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>18</td>
<td>-8</td>
<td>-8</td>
<td>-8</td>
<td>-8</td>
<td>-8</td>
<td>-8</td>
<td>-8</td>
<td>-8</td>
<td>-8</td>
<td>-8</td>
</tr>
<tr>
<td>24</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
</tr>
<tr>
<td>30</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
</tr>
<tr>
<td>36</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
</tr>
<tr>
<td>42</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
</tr>
<tr>
<td>48</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
</tr>
<tr>
<td>54</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
</tr>
<tr>
<td>60</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
<td>-5</td>
</tr>
</tbody>
</table>

Figure 2 Small test road network A
In the next stage, cases with variable link travel times were considered. It is assumed that link travel times on the solid link in Figure 2 follows the normal distribution of average, 40 min. and variance, 20 min$^2$, dotted link follows the normal distribution of average, 60 min. and variance, 40 min$^2$. Table 4 shows four cases: (a) In Case A-0 there is no variation in link travel times, (b) In Case A-1 both freight carriers use the VRPTW-F model with variation of link travel times, (b) In Case A-2 only freight carrier B uses the VRPTW-D model with variations in link travel times, (c) In Case A-3 both freight carriers use the VRPTW-D model with variation in link travel times. Multi-agent simulations were performed for 5,000 days.

Figure 3 Proposed price of freight carriers by multi-agent modelling (Case A-0)
Table 4 Type of VRPTW model used by freight carriers

<table>
<thead>
<tr>
<th>Variation of link travel times</th>
<th>Case A-0</th>
<th>Case A-1</th>
<th>Case A-2</th>
<th>Case A-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>Normal distribution</td>
<td>Normal distribution</td>
<td>Normal distribution</td>
<td></td>
</tr>
<tr>
<td>Freight carrier A</td>
<td>VRPTW-F</td>
<td>VRPTW-F</td>
<td>VRPTW-F</td>
<td>VRPTW-D</td>
</tr>
<tr>
<td>Freight carrier B</td>
<td>VRPTW-F</td>
<td>VRPTW-F</td>
<td>VRPTW-D</td>
<td>VRPTW-D</td>
</tr>
</tbody>
</table>

Figure 4 Proposed price of freight carriers (learning rate = 0.5)

Figure 4 shows the change of the proposed price of freight carriers for four cases. This figure indicates that more dynamic changes of the proposed price can be seen for Cases A-1, A-2 and A-3 with variation of link travel times compared to Case A-0 with fixed link travel times. It is interesting to note that even if link travel times varies in Cases A-1, A-2 and A-3, the proposed price converges into the Nash equilibrium value (18, 18) after a large number of iterations, while in Case A-0, it converges into the Nash equilibrium value (12, 12).

Figures 5 and 6 illustrate the average profits of freight carriers and the average costs of
shippers for the last 2,000 days (from 3,001st to 5,000th day) in case of learning rate of 0.5. Figure 5 indicates that in Case A-2, freight carrier B who uses the VRPTW-D model can obtain larger profits than freight carrier A who uses the VRPTW-F model. Freight carrier B can reduce the delay time at customers by dynamically adjusting vehicle routing planning and it leads to getting more customers. In addition, comparing Cases A-1 and A-3 in Figures 5 and 6 indicates that the use of the VRPTW-D model allows average profits of both freight carriers to increase and average costs of both shippers to decrease. This is actually a win-win situation for two agents, the “freight carriers” and “shippers”, for which the VRPTW-D model was generated.

4. CASE STUDY B

4.1 Test Conditions
The multi-agent models described above were applied to a large test road network of 49 nodes and 168 links as shown in Figure 7. The free running speeds on solid links in Figure 7 are 40km/h and those on dotted links are 30km/h. The dynamic traffic simulation was carried out using a block density method (Taniguchi et al., 2002) to identify the travel times on each link for 90 days. The dynamic traffic simulation requires information on passenger car behaviour. Passenger cars in this study include actual passenger cars and trucks other than those that are considered in the optimal routing and scheduling model. Passenger car Origin-Destination (OD) tables for every hour were estimated using traffic generation rates at each centroid and the probability of O-D choice. The number of passenger cars for each hour was generated using a temporal demand pattern based on the traffic census conducted in Hiroshima City. The total number of passenger cars generated was 500,000 - 600,000 vehicles.

There are ten freight carriers on the large test road network. The depots of freight carriers 1-3, 4-6, 7-8 and 9-10 are located at nodes 1, 13, 7 and 19, respectively. Each freight carrier has two 2-ton trucks. The costs of operating trucks are given in Table 5. There are assumed to be 45 shippers (customers) who are located at 45 nodes other than the nodes where the depots of freight carriers are located. Each shipper has demand of 500 kg and time windows that are randomly determined.
As Table 5 shows, the early arrival penalty is introduced for identifying the optimal routing of the VRPTW-F and the VRPTW-D models. The penalty of early arrival and delay at customers is considered only by freight carriers, since freight carriers need to wait until the starting time of designated time window if they arrive earlier and they have to pay delay penalty if they arrive late. Therefore freight carriers have to pay both the early arrival and delay penalty. On the other hand, as shippers do not care about the early arrival of vehicles at their site, shippers only consider the delay arrival of trucks which leads to generate opportunity cost for shippers. The equation (4) represents such behaviour of shippers who only pay attention to delay arrival of trucks.

Table 5 Costs of operating trucks

<table>
<thead>
<tr>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed cost (Yen/vehicle)</td>
<td>10,141.70</td>
</tr>
<tr>
<td>Operation cost (Yen/min.)</td>
<td>14.2</td>
</tr>
<tr>
<td>Early arrival penalty (Yen/min.)</td>
<td>14.02</td>
</tr>
<tr>
<td>Delay penalty (Yen/min.)</td>
<td>87.7</td>
</tr>
</tbody>
</table>

Table 6(a) Number of freight carriers who use VRPTW-F and VRPTW-D models

<table>
<thead>
<tr>
<th>Case</th>
<th>VRPTW-F</th>
<th>VRPTW-D</th>
</tr>
</thead>
<tbody>
<tr>
<td>B-1</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>B-2</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>B-3</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>B-4</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>B-5</td>
<td>0</td>
<td>10</td>
</tr>
</tbody>
</table>
Table 6(b) Number of freight carriers who use VRPTW-F and VRPTW-D models and city logistics measures

<table>
<thead>
<tr>
<th>Case</th>
<th>VRPTW-F</th>
<th>VRPTW-D</th>
<th>City logistics measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>B-3</td>
<td>7</td>
<td>3</td>
<td>none</td>
</tr>
<tr>
<td>B-6</td>
<td>7</td>
<td>3</td>
<td>Road pricing and subsidies</td>
</tr>
<tr>
<td>B-7</td>
<td>7</td>
<td>3</td>
<td>Case B-6 and co-operative freight transport systems</td>
</tr>
</tbody>
</table>

Table 6 shows the configuration of the cases. For cases B-1 – B-5, two stakeholders, “shippers” and “freight carriers” are involved, whereas for cases B-6 and B-7, three stakeholders, “shippers,” “freight carriers” and “administrators” are involved in multi-agent modelling. In Case B-6 the administrator implements a road pricing scheme such that if NOx emissions on a link exceed the environmentally requested level, the administrator imposes 100 Yen (= about 0.85 US$) whenever a truck travels on the link. In addition, in order to encourage shippers to choose freight carriers who plan vehicle routing and scheduling with less emissions of NOx than the required level (6 kg/day), the administrator provides subsidies of 5,000 Yen (=about 42.7 US$) to shippers. Therefore, the expected costs of shippers represented by equation (4) can be modified as follows.

$$E[C_{s,j}] = f_i + C_{d,j} \cdot E[t_{d,i}] - \delta \cdot w_s$$  

where,

$\delta = 1$, if shipper chooses freight carrier who plan vehicle routing and scheduling with less emissions of NOx than the required level

$= 0$, others

$w_s$: amount of subsidies

In Case B-7 co-operative freight transport systems are introduced. These systems determine the area where each freight carrier covers for collecting goods in advance. If a freight carrier gets a job collecting goods from a shipper outside the responsible area, it can exchange the job with other freight carrier who is responsible for that area under the condition that the proposed price is same.

The value of $C_{d,j}$ was taken as 87.7 yen/min., 263.1 yen/min., 438.5 yen/min., 613.9 yen/min. and 877 yen/min. and it was assumed that each shipper has one of these values. Shippers with one of these values were allocated at all the nodes.

4.2 Results

Figure 8 shows examples of the proposed price of freight carriers. This figure indicates that the proposed price of freight carriers goes down with days and proceeds to the range of 20,000-30,000 Yen and that there is not such a large difference between average value of propose price of freight carriers using the VRPTW-F and the VRPTW-D models.

Figure 9 illustrates the profits of freight carriers and costs of shippers. It can be noticed that freight carriers who used the VRPTW-D model could obtain larger profits than freight carriers who used the VRPTW-F model. This is attributed to that the VRPTW-D model allows freight carriers to reduce the delay penalty at shippers (customers) which leads to getting more shippers due to the reduction of costs of shippers. The profits of freight carriers
who used the VRPTW-D model decreased with the increasing number of freight carriers who used the VRPTW-D model, but the average profit for Case B-5 in which all freight carriers used the VRPTW-D model was a little higher than that of Case B-1, in which all freight carriers used the VRPTW-F model. It is interesting to note that the costs of shippers decreased with the increasing number of freight carriers who used the VRPTW-D model. Therefore, introducing the VRPTW-D model generated a win-win situation by increasing profits of freight carriers and decreasing costs of shippers.

Let us discuss the cost for collecting travel time information for the VRPTW-F and the VRPTW-D models. In Japan, VICS (Vehicle Information Communication Systems) provides freight carriers with good information of real time travel time on major roads at every 5 minutes. The cost for using VICS is included in the initial cost of setting the on-board equipment of car navigation systems. The travel time information given by VICS can be used both for the VRPTW-F and the VRPTW-D models, because the VRPTW-F model requires the average value of travel times which can be calculated by averaging the real time travel times for a certain period. In this case, the costs for collecting travel time information for the VRPTW-F and the VRPTW-D models are almost same. Therefore, the difference of performance between the VRPTW-F and the VRPTW-D models which is observed in Figure 9 holds without any modification. In case of other countries than Japan, where real time travel times is not given by VICS, costs for collecting real-time travel times would be very high.
compared to costs for collecting average travel times. In that case, the advantage of using the VRPTW-D model shown in Figure 9 may disappear due to high cost of collecting real-time travel times. Note that the results shown in Figure 9 is limited to the countries like Japan where the information of real time travel times is widely given by sophisticated systems such as VICS at very low costs.

![Figure 9 Profits of freight carriers and costs of shippers without any city logistics measures](image)

Figures 10 and 11 compare performance between Cases B-6 and B-7, in which administrators implemented road pricing, subsidies and co-operative freight transport systems were introduced and Case B-3, in which no city logistics measures were implemented. Figure 10 indicates that the profits of freight carriers increased and the costs of shippers increased as well, when administrators introduced road pricing. This was caused by higher proposed price of freight carriers and larger delay times at shippers due to road pricing. But the increase in costs of shippers who chose freight carriers using the VRPTW-D model is relatively smaller than those who chose freight carriers using the VRPTW-F model. Moreover, introducing co-operative freight transport systems improved the situation by increasing the profits of freight carriers and decreasing costs of shippers, no matter what model (VRPTW-F or VRPTW-D) was used. In particular, the costs for shippers who chose freight carriers using the VRPTW-D model was reduced compared with Case B-3.

![Figure 10 Performance comparison between Cases B-6 and B-7](image)

Figure 11 illustrates that NOx emissions in Case B-6, in which administrators implemented road pricing, considerably decreased and the number of links where NOx emissions were over the required limit also decreased. A greater reduction of NOx emissions and the number of links over the required limit was observed in Case B-7, in which co-operative freight transport systems were applied. From these results we can conclude that implementing road pricing can reduce NOx emissions but may increase the costs of shippers. To avoid such effects, introducing co-operative freight transport systems helps shippers to reduce their costs.

![Figure 11 NOx emissions comparison between Cases B-6 and B-7](image)

This demonstrates that multi-agent modelling can contribute to understanding the interaction among stakeholders who are involved in urban freight transport systems and effects of city logistics measures on profits and costs of stakeholders as well as effects on the environment. Multi-agent modelling is also useful to observe the side effects or unexpected effects of city logistics measures. These are benefits of using autonomous agents who make decision by themselves based on reinforcement learning.
5. CONCLUSIONS

This paper presented multi-agent modelling for evaluating the behaviour and interaction among stakeholders who are involved in urban freight transport systems as well as effects of city logistics measures. Comparing multi-agent models and game theory on small test road network showed that results of multi-agent models without any variation of link travel times converged into Nash equilibrium which game theory identified. Multi-agent simulation on a small test road network also demonstrated that the VRPTW-D model which dynamically adjusted vehicle routing planning to the current travel times generated good performance in terms of increasing profits for freight carriers and decreasing costs for shippers. After applying multi-agent models on a large test road network, it was observed that introducing the VRPTW-D model generated a win-win situation by increasing profits for freight carriers and
decreasing the costs for shippers. The results also show that the implementation of road pricing can reduce NOx emissions but may increase the costs for shippers. To avoid such effects, introducing co-operative freight transport systems helps shippers to reduce their costs. Further investigations are required to more precisely model the behaviour of stakeholders and learning processes.

The multi-agent simulation used in this study was validated in a simple small road network compared with results of game theory, but further research should be needed to validate the simulation in more realistic large road network where real actors play a role.

REFERENCES