

# Dynamic Vehicle Traffic Control Using Deep Reinforcement Learning in Automated Material Handling System

Younkook Kang, Sungwon Lyu, Jeeyung Kim, Bongjoon Park, Sungzoon Cho\*

Department of Industrial Engineering, Seoul National University

[ykkang](mailto:ykkang@dm.snu.ac.kr), [lyusungwon](mailto:lyusungwon@dm.snu.ac.kr), [jeeyung.kim](mailto:jeeyung.kim@dm.snu.ac.kr), [bongjoon](mailto:bongjoon@dm.snu.ac.kr), [zoon@snu.ac.kr](mailto:zoon@snu.ac.kr)

## Abstract

In automated material handling systems (AMHS), delivery time is an important issue directly associated with the production cost and the quality of the product. In this paper, we propose a dynamic routing strategy to shorten delivery time and delay. We set the target of control by analyzing traffic flows and selecting the region with the highest flow rate and congestion frequency. Then, we impose a routing cost in order to dynamically reflect the real-time changes of traffic states. Our deep reinforcement learning model consists of a Q-learning step and a recurrent neural network, through which traffic states and action values are predicted. Experiment results show that the proposed method decreases manufacturing costs while increasing productivity. Additionally, we find evidence the reinforcement learning structure proposed in this study can autonomously and dynamically adjust to the changes in traffic patterns.

## Introduction

In the semiconductor manufacturing environment, an automated material handling system (AMHS) plays a major role in raising productivity. There are two main components in a typical AMHS: vehicles and vehicle controllers. Transfer robots serve as vehicles of the system by moving materials between manufacturing equipment. The vehicle controllers, on the other hand, are in charge of dispatching and routing these vehicles. As is in a transportation system, predicting traffic flow and congestion is one of the main issues addressed by AMHS researchers and experts, yet it differs from the typical transportation system problems in a sense that it requires optimizing traffic in the entire manufacturing system rather than optimizing a specific path or a vehicle trajectory. In this paper, we aim to shorten the average delivery time and delay by considering traffic states on a real-time basis when selecting routes.

In our experiment setting, the subject factory consists of 500 vehicles scattered through a layout area larger than an

archetypal soccer field. These vehicles receive delivery orders every 1 to 2 seconds and carry materials to the designated destination in 3 to 4 minutes on average. The challenge in the prediction of traffic flow comes from the frequent, yet seemingly erratic, breakouts of congestions.

In this paper, we proposed a deep reinforcement learning model in order to minimize delivery time and delay given (real-time) changes in the routing costs at key paths. We first determine paths with the greatest impact on traffic and select them as the key paths of prediction. Then, we adjust traffic weights using these key paths in our deep reinforcement learning model which utilizes Q-learning and a recurrent neural network. Experiment results show that: (1) RNN structure works well with predicting traffic patterns on a real-time basis, and; (2) reinforcement learning component well predicts the state and the action values at a given traffic state. Our proposed model contributes to the existing literature by developing an autonomous traffic adjustment system flexibly responding to the real-time changes in the line traffic while requiring minimum human interference.

## Related Work

There has been a myriad of studies focusing on solving the vehicle routing problems via optimization, for example, the time window approach (Smolic-Rocak et al., 2010). Another branch of work includes congestion-awareness based dynamic routing, such as the study by Bartlett et al., 2014. However, these studies are limited in terms of: (1) the size of lines considered in the experiment; (2) simplification of circumstantial conditions, or; (3) fixed traffic patterns.

Recent rise of the deep learning methods such as convolutional neural networks (CNN) or recurrent neural networks (RNN) have resulted in the utilization of these techniques when tackling prediction of traffic congestions (Wang et al., 2016, Fouladgar et al., 2017). Most of these studies, however, focus only on prediction. Moreover, there exist specific conditions, peculiar only to the manu-

facturing environment, that must be considered when solving the congestion problems under AHMS setting, which most of the related studies fail to concern.

Reinforcement learning is a methodology which takes an action in a sequential manner based on the notion of certain cumulative “reward” or “compensation” of the chosen action at a given state of the system. We contribute to the bulk of existing literature by being one of very few studies to apply reinforcement learning framework to solving a real-world problem in the manufacturing industry.

## Proposed Model

### Traffic Flow Analysis

The flow pattern of the line was analyzed in order to determine critical area with large flow and congestion. The 40 path weights among 2000 that enter this area are dynamically adjusted so as to control the entire traffic flow.

### RL modeling

Path weights at selected area are then reflected in the vehicle routing process in order to control the traffic flow. Upon the completion of the time period  $t$ , the reward on the chosen action is fed back to the model in order to update path weights accordingly.

State: flow rates, velocity, velocity deviation, and congestion time of each period line were used variables representative of states. We also consider the time series of these variables.

Action: at a given point in time, an action value that maximizes compensation according to the corresponding state is chosen among 3 candidate value pre-determined. The chosen value is applied to 40 key paths determined by the precedent traffic flow analysis.

Reward: compensation was computed by the linear combination of delivery time and delay given the end goal.

Episode: For each semiconductor line, an event of a transfer occurs back-to-back without interruption. We divided these events by unit time, each of the resulting divisions were then regarded as an individual episode.

### Algorithm

Reinforcement learning algorithm is based on Q-learning and the prediction task is solved by using a RNN structure in order to account for time-series characteristics.

## Experiment

### Environment

Experiments were simulated given the actual operational conditions of an anonymous semiconductor AMHS environment. Experiment environment consists of a historical data of a line with the transfer rate of 8,000 per hour. The AMHS simulator was designed to represent a rail layout with 50 bays and 500 vehicles.

### Data set

A historical data of 24-hour length transfer events was used for training the proposed model. Model performance was tested by using another 8-hour length of events data.

### Result

We measured delivery time and delay count over 30 seconds in a given section to show whether there was an improvement in the transfer process. For comparative analysis, we used static path weights as the baseline. According to Table 1, experiment results show that the proposed model reduces delivery time by 1.03 seconds, and the delay by 23 percent, as compared to optimal static model.

condition		output	
TYPE	Action	Delivery Time	Delay Counts
baseline	None	199.63	2149
Static	Fixed(5)	197.09	1612
Dynamic	Changed(0,5,10)	196.06	1239

. Deliver Time : time from requested to completed  
. Delay Counts : 30second delayed count in total sections

Table 1. experiment result

## Conclusion

In the paper, we propose a dynamic routing model focusing on traffic congestion. Our work contributes to the existing literature by developing a dynamic routing scheme which controls traffic flow adaptively to the changing conditions of the environment. Our experiment shows promising results in shortening delivery time and delay, which is directly associated with a reduction in the production cost and an improvement of facility utilization and productivity. Our work is among the first to apply reinforcement learning in traffic control under a manufacturing environment, and further research on expanding the role of reinforcement learning in various industrial settings is called for.

## References

- Jingyuan Wang et al. 2016. Traffic speed prediction and congestion source exploration: A deep learning method. In *Data Mining (ICDM), 2016 IEEE 16<sup>th</sup> International Conference on*, pages 499–508.
- Kelly Bartlett et al. 2014. Congestion-aware dynamic routing in automated material handling systems. *Computers & Industrial Engineering*, 70:176–182.
- Mnih, V et al. 2013. Playing atari with deep reinforcement learning. *CoRR* abs/1312.5602.
- Mnih, V et al. 2015. Human-level control through deep reinforcement learning. *Nature* 518(7540):529–533.
- Mohammadhani Fouladgar et al. 2017. Scalable deep traffic flow neural networks for urban traffic congestion prediction. *arXiv preprint arXiv:1703.01006*
- Smolic-Rocak et al. 2010. Time windows based dynamic routing in multi-AGV systems. *IEEE Transactions on Automation Science and Engineering*, 7, 151–155.