Predictive vehicle dispatching method for overhead hoist transport systems in semiconductor fabs

Presented by Byun Min Kim

Introduction

Purpose of this paper

- Minimize the average lead time of jobs by assigning it to appropriate vehicle

OHT control system

- OHT : The vehicle that moves FOUP along the rails installed on the ceiling
- OCS(OHT control system) : System that manages and coordinates the movement and operation of OHT within FAB

Importance of creating efficient OHT control system

- Inefficient operation of OHTs always lead to undesirable delays of wafer transfer hence disrupting the production schedules of tools

Related Works

Static vehicle dispatching

- Job-vehicle assignment decisions are not allowed to change once they are made

Dynamic vehicle dispatching

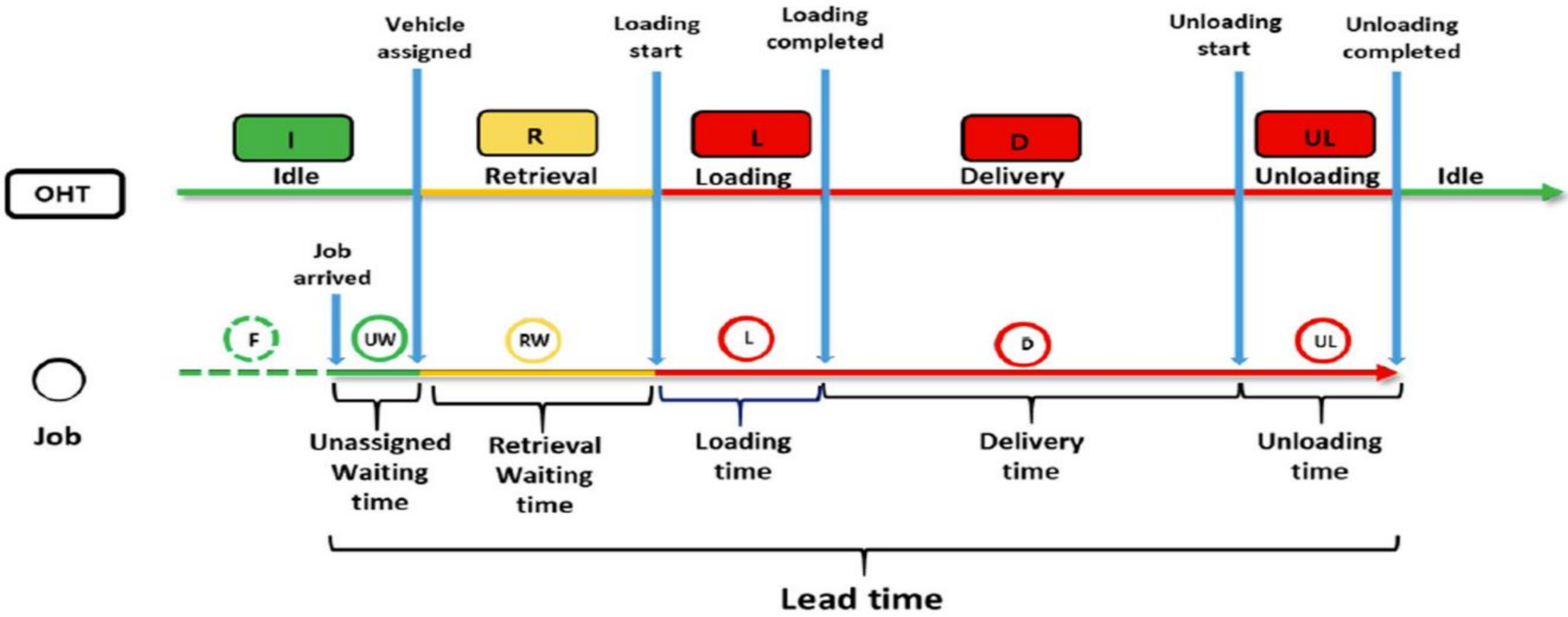
- Allow changing job-vehicle assignment decisions adaptively to capture the constant change of FAB

Static vehicle dispatching	Dynamic vehicle dispatching
Hu et al. 2020. "Deep Reinforcement Learning Based AGVs Real Time Scheduling with Mixed Rule for Flexible Shop Floor in Industry 4.0."	Liao et al .2006. "Differentiated Preemptive Dispatching for Automatic Materials Handling Services in 300 mm Semiconductor Foundry."
Liao et al. 2002. "Dynamic OHT Allocation and Dispatching in Large-Scaled 300 mm AMHS Management."	Kim et al. 2009. "Effective Overhead Hoist Transport Dispatching Based on the Hungarian Algorithm for a Large Semiconductor FAB.".

Job delivery process

Status of Job : F(Not arrived yet), UW(unassigned waiting job), RW(Retrieval waiting job), L(Loading time), D(Delivery time), UL(Unloading time)

Status of OHT : I(Idle), R(Retrieval), L(Loading), D(Delivery), UL(Unloading)



Problem statement

Assigning available jobs to appropriate vehicles when Job arrived by integer programming

Objective
- Minimize the average lead time of jobs

Decision point
- When Job arrived in storage

Available vehicle's status
- All Vehicles (I(Idle), R(Retrieval), L(Loading), D(Delivery), UL(Unloading))

Available Job's status
- Not arrived yet jobs, retrieval waiting jobs, unassigned waiting jobs

Suppose that
- Retrieval vehicles can change the assigned jobs

Key idea

1. **Considering pre-arrival information (expected arrival time of jobs in the near future and the time needed for occupied vehicles to become idle) that can reduce job waiting time**

Conventional vehicle dispatching

Available job	Available vehicle
Waiting job	Idle vehicle

Proposed vehicle dispatching

Available job	Available vehicle
Waiting job Retrieval waiting job Not arrived yet Job	All Vehicles

2. **Using pre-arrival information while taking into account the possibility of prediction error**

- In reality the arrival time of future jobs usually can not be precisely estimated
- Solve this issue by giving penalizing cost on the assignment of vehicles to far-future jobs

Method (Linear assignment formulation)

Mathematical model

$$\text{Let} \quad x_{jv} = \begin{cases} 1 & \text{if job } j \text{ is assigned to vehicle } v \\ 0 & \text{if not} \end{cases}$$

$$c_{jv} = \text{cost of matching job } j \text{ with vehicle } v \tag{1}$$

$$\text{Minimize} \quad Z = \sum_{j}^{|J|} \sum_{v}^{|V|} c_{jv} x_{jv}$$

Subject to :

$$\sum_{v}^{|V|} x_{jv} = 1, \quad \text{for } j = 1, 2, \ldots, |J| \tag{2}$$

$$\sum_{j}^{|J|} x_{jv} = 1, \quad \text{for } v = 1, 2, \ldots, |V| \tag{3}$$

$$x_{jv} = \text{binary, for all } j \text{ and } v \tag{4}$$

Available job set J

- Waiting jobs + Retrieval waiting jobs + Not arrived yet job

Available vehicle set V

- All vehicles

***Calculated every time when a job arrives* at the storage**

Method (BPD vs CWPD)

BPD (Basic predictive dispatching)

- Cost function defined for the environment without prediction error

VS

CWPD(Certainty Weighted Predictive Dispatching)

- Cost function defined for the environment with prediction error

Method (BPD)

Cost function of BPD(Basic predictive dispatching)

$$c_{jv} = \boxed{JWT_{jv}} + \alpha \cdot RAT_j$$ → Cost function of BPD

$$JWT_{jv} = (TTC_v + RT_{jv} - TTA_j)^+$$ → waiting time of job j when it is matched with vehicle v

$$TTC_v = \begin{cases} 0 & \text{if vehicle } v \text{ is in idle or retrieval status} \\ RLT_v + JT_v^{cl \to dest} + RULT_v & \text{if not} \end{cases}$$

time required for vehicle v to complete its currently assigned job

$$RT_{jv}$$

retrieval time required when job j is assigned to vehicle v

$$TTA_j = (EAT_j - t_{now})^+$$

the time required for job j to arrive at t_{now}

Method (BPD)

Cost function of BPD(Basic predictive dispatching)

$c_{jv} = JWT_{jv} + \alpha \cdot RAT_j$ ➡ Cost function of BPD

$RAT_j = EAT_j - \min\{EAT_1, EAT_2, \ldots, EAT_{|J|}\}$ ➡ High RAT_j = Far future job

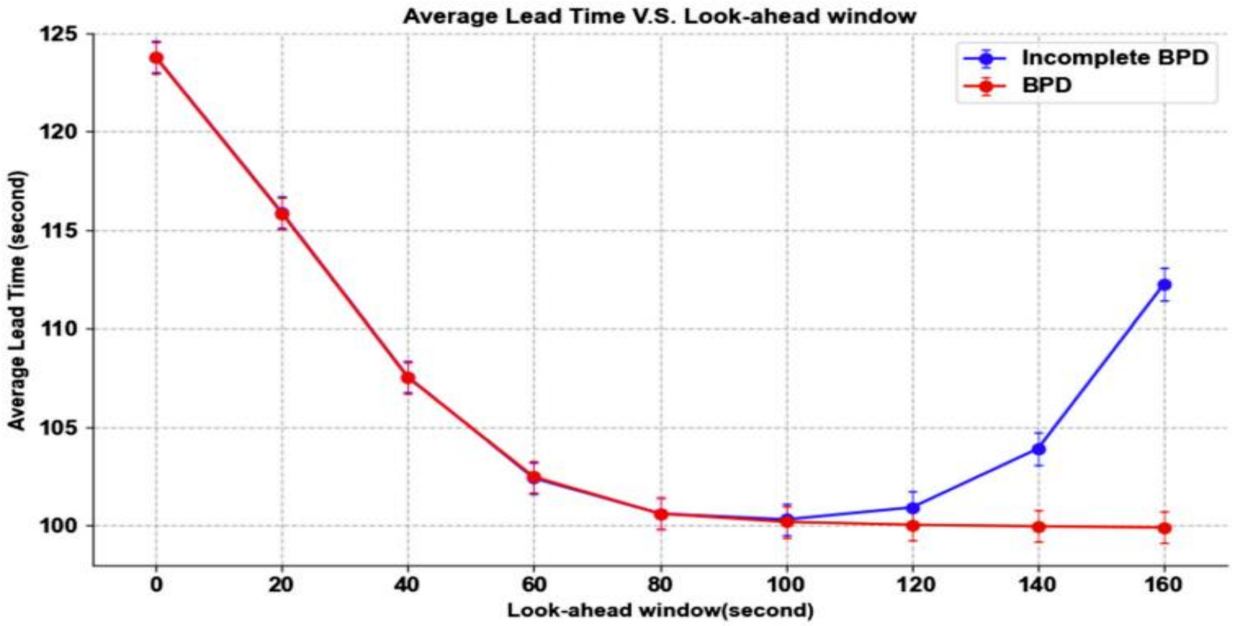
EAT_j = Expected arrival time of job j

Meaning of RAT_j

- If Look ahead window is large (number of Job > number of vehicle) Vehicle have to assign in near future Job rather than far future Job

Look-ahead window = time horizon that the pre-arrival information of jobs can be known

Effectiveness of RAT_j



Method (CWPD)

Cost function of CWPD((Certainty Weighted Predictive Dispatching)

BPD do not consider Prediction error ➡ Job should be differentiated by the certainty of their EAT_j (Expected arrival time of job j)

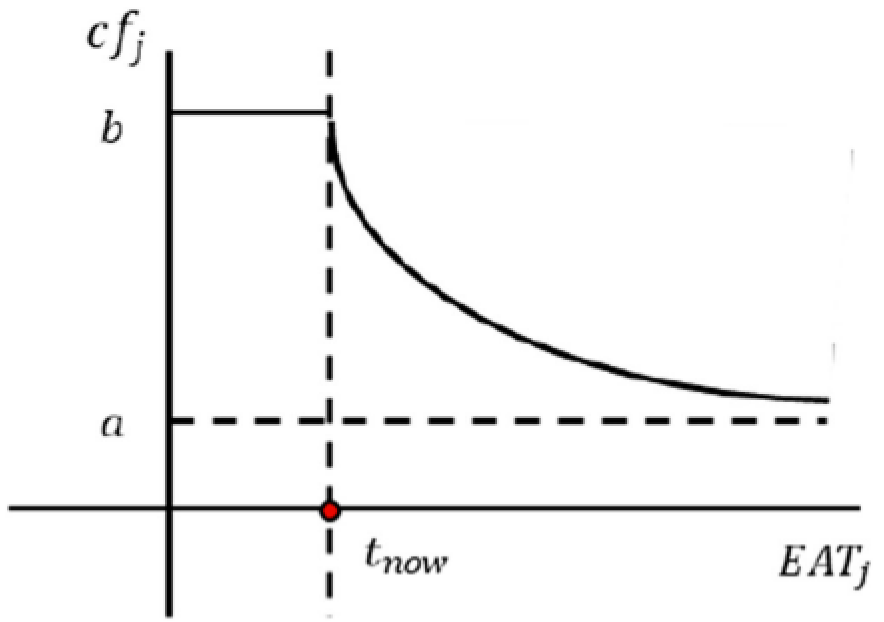
$c_{jv} = JWT_{jv} \times cf_j + \alpha \cdot RAT_j$ ➡ Cost function of CWPD

Meaning of Cf_j(high certainty factor)

$$cf_j = \begin{cases} b & \text{if } AAT_j \leq t_{now} \\ \dfrac{b-a}{1+(EAT_j - t_{now})} + a & \text{if not} \end{cases}$$

- Job with high certainty factor(Cf_j) is more likely to be assigned with the vehicle that minimizes its waiting time

Relationship between Cf_j and EAT_j

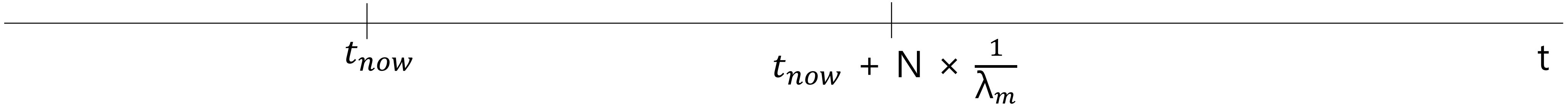


Method (Expected arrival time)

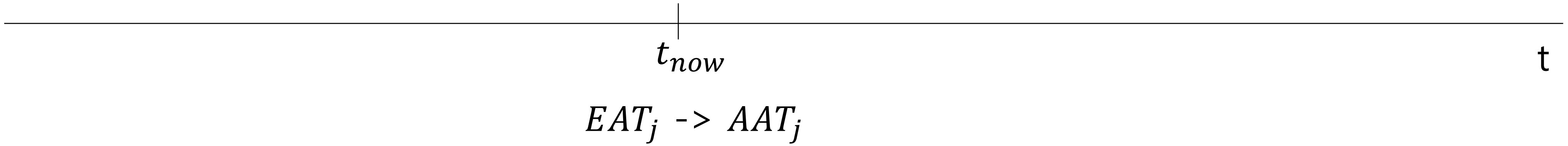
$\frac{1}{\lambda_m}$ = cycle time of bay m
N = Nth Job
EAT_j = Expected arrival time of job j
AAT_j = Actual arrival time of job j

How to define expected arrival time (EAT_j) in prediction error environment

(1) If the job generated at t_{now} is from bay m -> $EAT_j = t_{now} + N \times \frac{1}{\lambda_m}$



(2) For the job j that has already arrived, EAT_j should be updated to AAT_j (Actual arrival time)

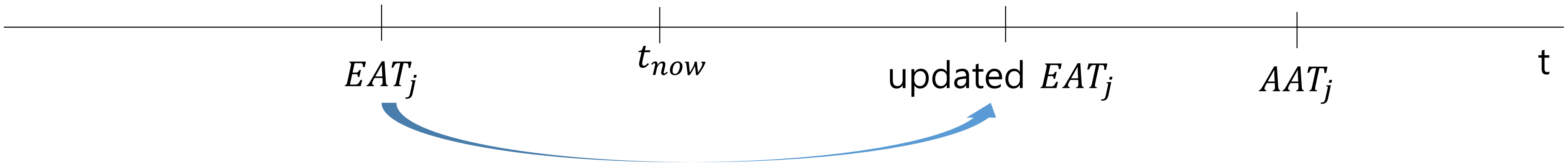


Method (Expected arrival time)

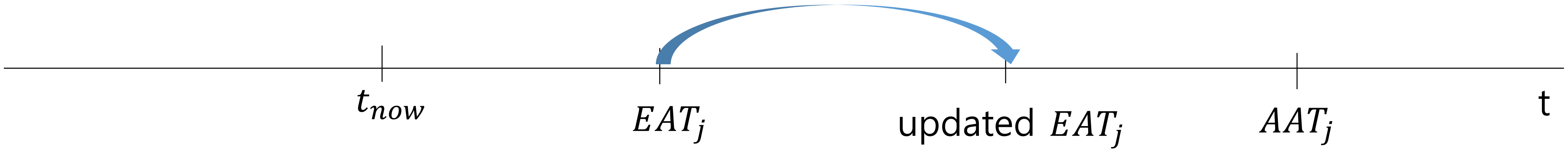
EAT_j = Expected arrival time of job j
AAT_j = Actual arrival time of job j

How to define expected arrival time (EAT_j) in prediction error environment

(3) If job j who has not arrived yet but EAT_j is smaller than t_{now}, Then EAT_j is updated by uniformly sampling from the interval between t_{now} and AAT_j, EAT_j = U(t_{now}, AAT_j)



(4) For the job j who has not arrived yet and EAT_j is also larger than t_{now}, EAT_j is update by uniformly sampling from the interval EAT_j and AAT_j by probability P_j ($P_j = \frac{1}{k \cdot (AAT_j - t_{now}) + 1}$)

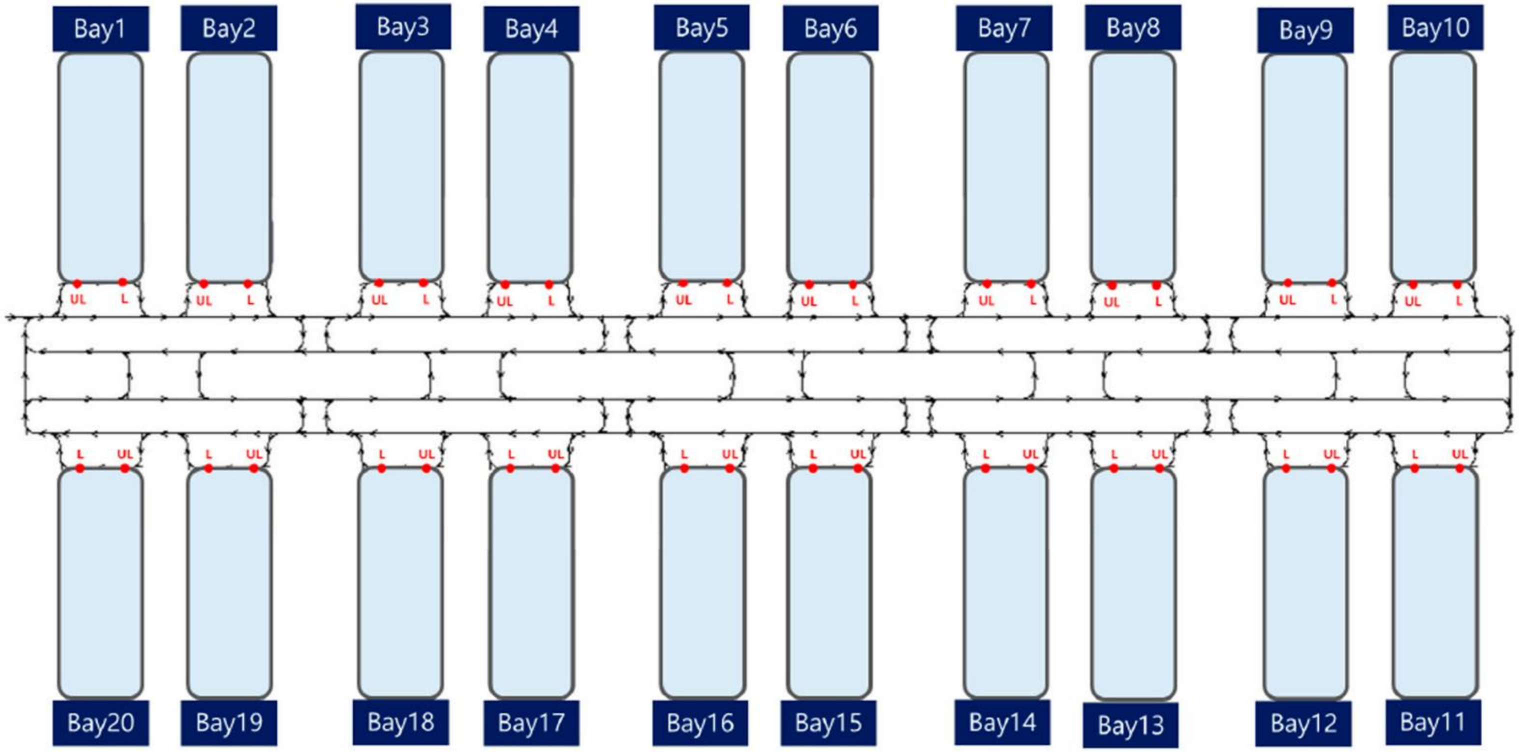


*****As AAT_j gets closer to t_{now}, the probability of EAT_j being updated to be closer to AAT_j increases*****

Experiment

Conditions

- Bay = 20
- Horizontal length = 128m
- Vertical length = 16m
- Number of Vehicles = 28
- Simulation warmup = 24 simulation hours
- Real Simulation = 24 simulation hours
- L= Loading point
- UL = unloading point

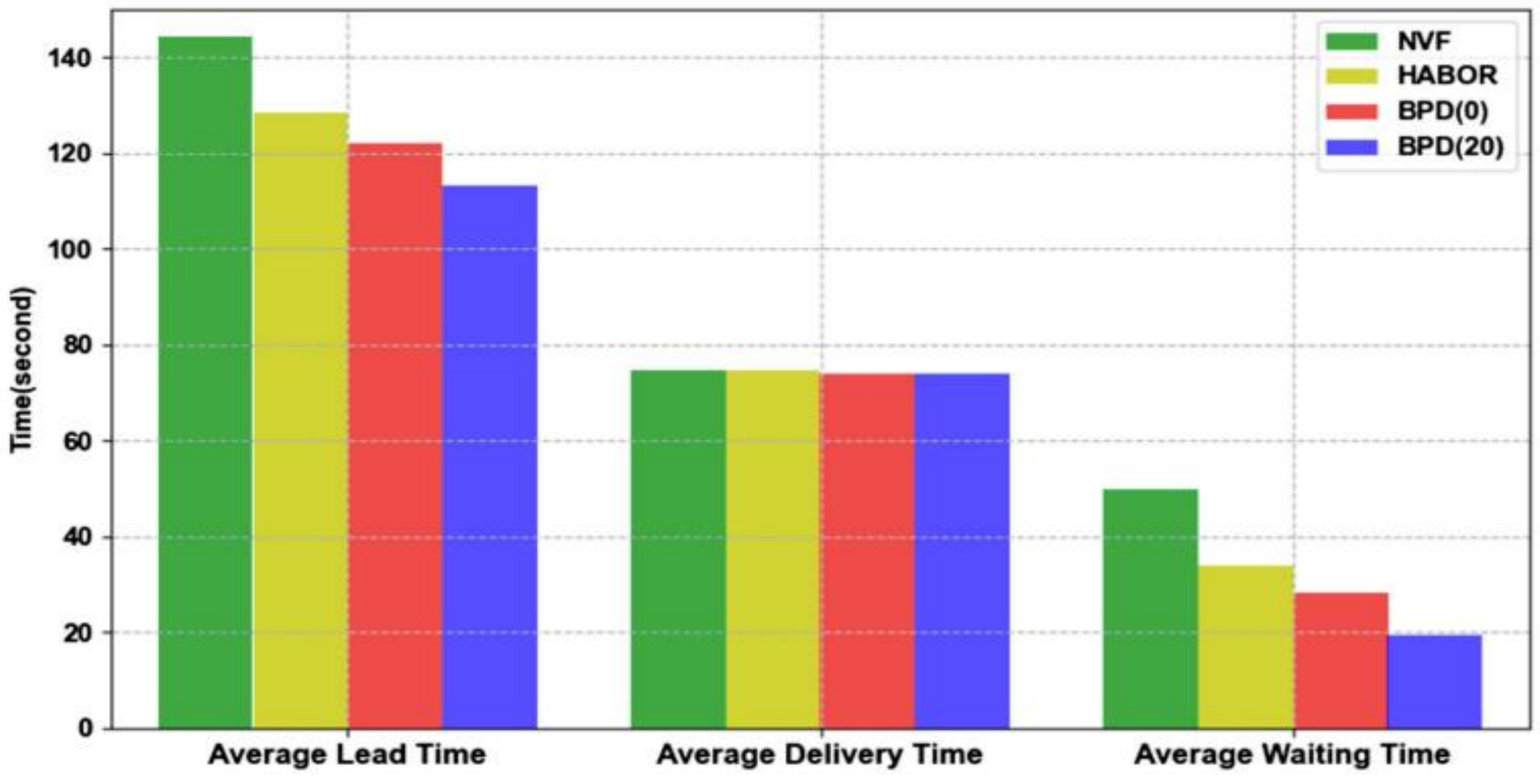


Fab layout

Result

Environment without prediction error (expected arrival time = Actual arrival time)

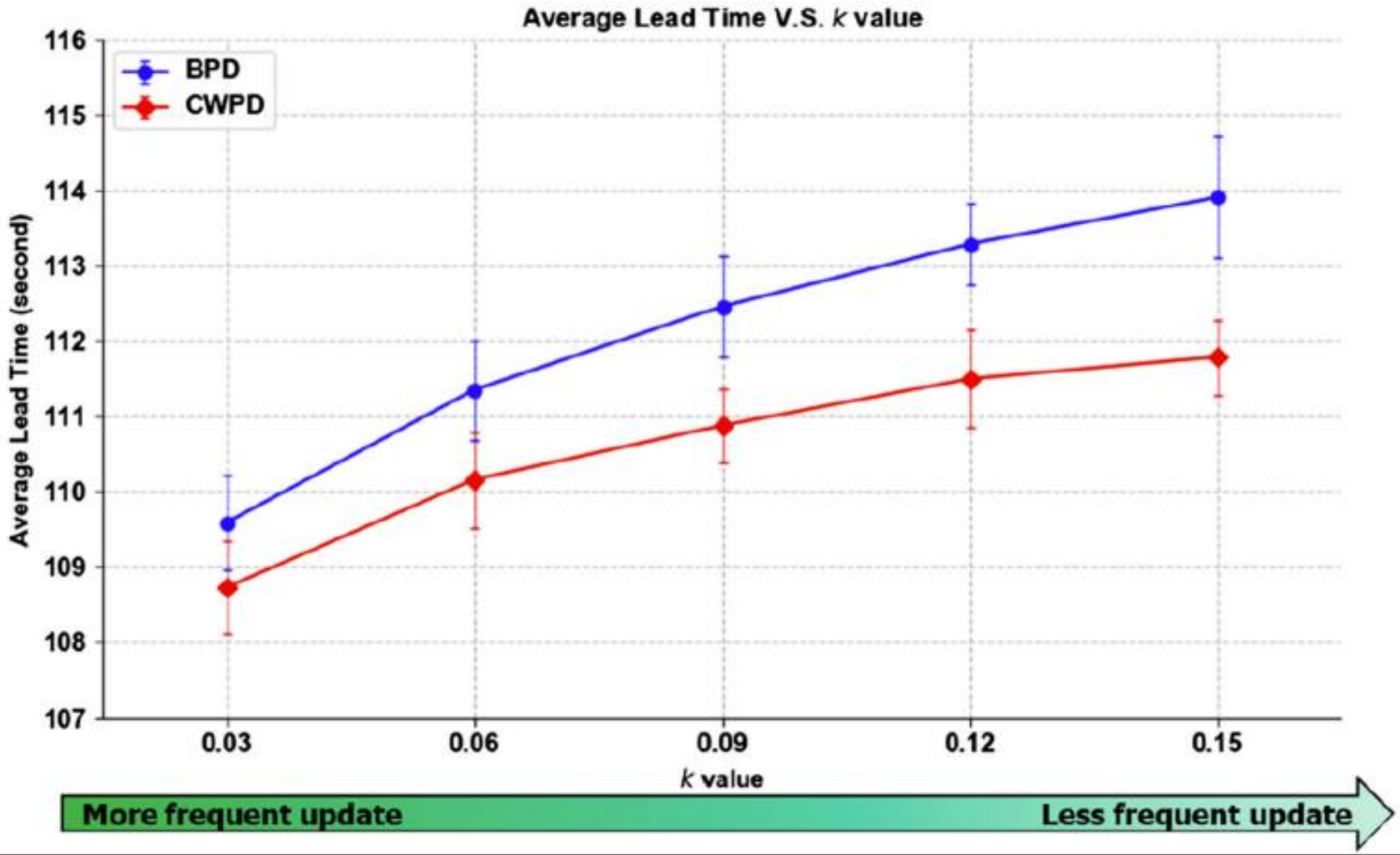
- NVF = Nearest Vehicle first
- HABOR = Formulate as an assignment problem and solved by Hungarian algorithm
- BPD(0), BPD(20) = Proposed BPD method with 0s and 20second time window



Result

Environment with prediction error (expected arrival time =~ Actual arrival time)

- BPD =Cost function defined for the environment without prediction error
- CWPD = Cost function defined for the environment with prediction error



Conclusion

- **Vehicle dispatching using the job state and the vehicle state expected in the near-future can significantly improve AMHS's performance in semiconductor fabs, which can mean a considerable cost saving for semiconductor manufacturers.**

- **The idea of predictive dispatching can also be applied by online ridesharing companies such as Uber or CJ routing problem**

- **Further studies may include the accurate prediction of the future transfer requests and vehicles' travel time because they can directly improve the performance of predictive dispatching**

Appendix

Table 4. Summary of notations.

Symbol	Explanation
J	available job set
V	available vehicle set
λ_{mn}	the arrival rate of transfer request from bay m to bay n
λ_m	the arrival rate of transfer request from bay m
EAT_j	the expected arrival time of job j
AAT_j	the actual arrival time of job j
t_{now}	the moment of dispatching
P_j	the probability of updating EAT_j for job j
TTA_j	time required for job j to arrive
$JT_v^{cl \to dest}$	Journey time from the current location of vehicle v to its destination
RLT_v	remaining loading time of vehicle v
$RULT_v$	remaining unloading time of vehicle v
RAT_j	relative arrival time of job j
TTC_v	time required for vehicle v to become idle
JWT_{jv}	job waiting time when job j is matched with vehicle v
c_{jv}	matching cost when job j is matched with vehicle v
cf_j	certainty factor of job j