

# A Dynamic OHT Routing Algorithm in Automated Material Handling Systems

Bonggwon Kang\* · Byeong Min Kang\*\* · Soondo Hong\*\*\*†

\*Major in Industrial Data Science & Engineering, Department of Industrial Engineering,  
Pusan National University

\*\*Digital Strategy Department, Busan Bank

\*\*\*Department of Industrial Engineering, Pusan National University

## 자동화 물류시스템 내 차량 혼잡도를 고려한 무인운반차량의 동적 경로 결정 알고리즘

강봉권\* · 강병민\*\* · 홍순도\*\*\*†

\*부산대학교 산업공학과 산업데이터공학융합전공

\*\*부산은행 디지털전략부

\*\*\*부산대학교 산업공학과

An automated material handling system (AMHS) has been emerging as an important factor in the semiconductor wafer manufacturing industry. In general, an automated guided vehicle (AGV) in the Fab's AMHS travels hundreds of miles on guided paths to transport a lot through hundreds of operations. The AMHS aims to transfer wafers while ensuring a short delivery time and high operational reliability. Many linear and analytic approaches have evaluated and improved the performance of the AMHS under a deterministic environment. However, the analytic approaches cannot consider a non-linear, non-convex, and black-box performance measurement of the AMHS owing to the AMHS's complexity and uncertainty. Unexpected vehicle congestion increases the delivery time and deteriorates the Fab's production efficiency. In this study, we propose a Q-Learning based dynamic routing algorithm considering vehicle congestion to reduce the delivery time. The proposed algorithm captures time-variant vehicle traffic and decreases vehicle congestion. Through simulation experiments, we confirm that the proposed algorithm finds an efficient path for the vehicles compared to benchmark algorithms with a reduced mean and decreased standard deviation of the delivery time in the Fab's AMHS.

Keywords : Dynamic Routing, Q-Learning, Automated Material Handling System, Simulation

### 1. 서 론

대규모 고가의 제조설비들로 구성되어있는 반도체 제

조라인(Fab: Semiconductor Fabrication)은 자재 운반의 자동화 및 최적화를 위해 자동화 물류시스템(Automated material handling system, AMHS)을 사용하고 있다. Fab 내 AMHS는 고가 설비의 가동률을 최대화하고자 설비에 지연 없이 웨이퍼를 공급하는 것을 목표로 한다[12].

수백 번의 공정을 거쳐 완성품으로 생산되는 웨이퍼는

Received 20 July 2022; Finally Revised 14 September 2022;  
Accepted 15 September 2022

† Corresponding Author : soondo.hong@pusan.ac.kr

FOUP(Front-opening unified pod)에 실려 하나의 로트로 취급된다[6, 11]. OHT(Overhead hoist transport)는 이러한 로트를 설비에 공급하는 역할을 한다. 유힘 상태인 OHT는 반송 명령에 따라 지정한 로트를 출발 설비에서 적재한 후 목적 설비에 적하하는 일련의 반송작업을 수행한다[10].

하지만, Fab 내 AMHS 용량(Capacity) 부족으로 설비가 유힘 상태임에도 불구하고 적시에 FOUP을 공급하지 못하는 Fab 생산성 저하 현상이 발생하고 있다[17]. OHT 간에 발생하는 정체(Congestion) 현상은 이상적 AMHS 용량에 도달하기 어렵게 한다. 정체현상은 원인에 따라 크게 블로킹 현상과 간섭 현상으로 구분할 수 있다. 블로킹 현상은 OHT가 적재 및 적하작업이 아닌 상태에서 다른 OHT에 의하여 정지한 상태이다. 간섭 현상은 OHT가 주행 중 다른 OHT와 충돌 방지를 위하여 감속하는 상태를 의미한다.

AMHS의 생산성은 FOUP의 반송시간과 반송 횟수로 평가할 수 있다. FOUP의 반송시간은 FOUP이 출발 설비에서 반송을 요청한 뒤에 목적 설비에 적하 완료하는 데까지 소요된 시간이다. FOUP 반송 횟수는 일정 시간당 적하 완료된 FOUP의 수이다. FOUP의 반송 시간과 반송 횟수를 개선하기 위하여 Fab은 크게 3가지 차량 관리 방법을 고려할 수 있다[4]: (a) 차량경로 결정(Vehicle routing rule); (b) 차량배차(Vehicle dispatching rule); (c): 유힘차량관리(Idle vehicle management rule). 본 연구는 시간에 따라 변동하는 정체량을 고려하여 차량 경로를 동적으로 설정하는 Q-Learning 기반 동적 차량 경로 결정 문제(Dynamic Vehicle Routing Problem)를 다룬다.

Q-Learning은 강화학습의 한 종류로서 특정 환경(Environment)에서 에이전트가 현재의 상태(State)를 고려하여 선택할 수 있는 의사결정(Action) 중 보상(Reward)을 최대화 혹은 페널티를 최소화하는 의사결정을 도출하는 것을 목적으로 한다. Q-Learning은 확률적 시행착오를 통하여 효과적인 의사결정을 학습하게 된다[2, 20]. 하지만, Q-Learning 기반 의사결정의 효과성은 상태모델링에 따라 영향받기 때문에 시스템 복잡도와 의사결정 목적에 부합하는 상태모델링이 요구된다. 상태 조합 수가 증가하는 경우 모든 상태에 대한 충분한 학습이 어렵고, 반대로 상태 조합 수가 감소하는 경우는 충분한 학습이 이루어져도 정밀한 의사결정이 어렵다[13]. 본 연구에서는 AMHS 내 OHT 간 정체 패턴을 고려하기 위한 상태모델링을 제안하고, 시뮬레이션 실험을 통해 효과성을 확인하였다.

본 연구의 2장에서는 관련 선행 연구를 소개한다. 3장에서는 Fab 내 AMHS 및 시뮬레이션을 소개한다. 4장에서는 정체 경로를 우회하기 위한 동적 경로 결정 알고리즘을 소개하고, 5장과 6장에서는 시뮬레이션 실험 결과

와 결론을 설명한다.

## 2. 문헌연구

AMHS의 성능향상을 위해서 AGV(Automated guided vehicle) 물류시스템 설계 문제, 차량 관리 문제(차량 배차 규칙, 유힘 차량 관리, 차량 경로 결정) 등의 여러 운영전략에 걸친 연구가 진행되고 있다. Ganesharajah et al. [3]은 AGV 기반 물류시스템에서 고려할 수 있는 레이어아웃을 소개하고, 최단 시간 반송을 위한 경로 결정 수리모델을 언급하였다. 또한, 차량과 작업을 동시에 고려한 전역 스케줄링 수리모델을 제안하였다. 작업과 차량의 통합 스케줄링은 복잡한 AGV 시스템에 적용엔 제한되므로, 신속한 차량 할당을 위해 사용된 차량배차 규칙들도 다루었다. Maxwell and Muckstadt[14]는 정적인 환경에서 AGV의 최소 투입개수를 결정하기 위한 수리모델을 제시하고, AGV간 정체 및 버퍼의 용량을 고려한 휴리스틱 기반 AGV 배차 규칙을 제시하였다.

Fab 내 AMHS의 반송 능력 개선을 위한 연구도 활발히 진행되고 있다. Kim et al.[9]은 Ganesharajah et al.[3]에서 언급된 차량배차 규칙 중 널리 사용되는 Shortest travel distance first(STDF) 기반의 차량 할당규칙의 한계점을 지적하였다. AMHS 내 차량 대수를 최소한으로 하고, 반송당 이동 시간의 평균 및 반송당 이동 시간의 분산을 줄이는 것을 목적으로 새로운 할당규칙을 제안하였다. 헝가리안 알고리즘을 이용하여 하나의 OHT가 반송작업을 완료할 때마다 당시 유힘 상태인 OHT들이 각각 가장 가까운 웨이퍼를 재할당하는 전략을 제시하였다. 정체가 예상되는 특수한 상황에서의 재할당을 통해 작업당 리드타임, 분산, OHT 유힘율이 개선됨을 보였다. Min and Yun [16]은 반도체 AMHS의 OHT 반송당 이동 시간 산포가 불규칙하고 지연되는 문제를 개선하기 위해 구역에 따른 할당 전략을 제시하였다. 해당 연구에서는 AMHS 내 영역을 구분하고 일정 영역 내의 OHT만 탐색하여 작업을 할당하는 방법을 제안하여 OHT의 공차이동 시간과 FOUP의 대기시간을 줄이고자 하였다. 제안한 방법은 OHT 이동 시간의 평균과 분산을 감소시킬 수 있었다. Bartlett et al.[1]은 OHT 동적 경로 결정을 위하여 Overhead rail들의 방문 비용을 정해진 시간 단위로 업데이트하였다. 방문 비용 주기 및 OHT 작업부하에 따라 제안한 알고리즘의 성능을 비교·분석하였다. OHT의 고장 발생을 고려한 시뮬레이션에서 제안된 알고리즘이 정적 경로 결정 알고리즘보다 우수한 성능을 보임을 확인하였다.

Fab 생산시스템을 고려한 효율적 AMHS 운영 방안

대한 연구도 진행되고 있다. Wan and Shin[19]은 Fab AMHS의 확정된 미래 생산 요구량과 차량의 예상 유휴 전환 시간을 바탕으로 작업 완료시간 최소화를 위해 예방적 차량배차 규칙(Basic Predictive dispatching, BPD)을 제안하였다. 또한, 생산 요구량 예측의 신뢰도를 고려한 신뢰도 가중 예방적 차량배차(Certainty Weighted Predictive Dispatching, CWPD)를 제안하였다. 시물레이션 실험을 통해 생산 요구량 예측의 불확실성 아래 CWPD가 BPD 대비 작업완료 시간을 개선시킬 수 있음을 보였다. Han et al.[5]은 Fab 생산설비의 가동률을 개선하기 위하여 가중치 기반 로트 타겟팅 규칙을 제안하였다. 가중치 기반 로트 타겟팅 규칙의 중요 파라미터인 타겟팅 규칙들의 가중치를 시물레이션을 기반으로 결정하고, 실험을 통해 효과성을 확인하였다.

반도체 물류시스템뿐만 아니라 자동화 컨테이너 터미널에서도 물류시스템의 성능향상을 위해 AGV에 대한 연구들이 진행되고 있다. Kang et al.[8]은 AGV들이 지나가게 될 노드들을 예약하고 이 중 상호연결 관계가 있는 것끼리 묶음으로써 사전에 AGV 간 충돌이 발생할 상황을 파악하고자 하였다. 노드집합을 그룹화함으로써 일부 AGV들이 특정 영역으로 진입하는 것을 일시적으로 제한하였다. 제안된 동적 AGV 유도경로 설계 방법은 AGV 간 충돌을 회피할 수 있었다. 이 전략은 기존의 그리드 제어 방식의 전략과 비교하여 시간당 반송 개수, 평균 정체 횟수, AGV 평균 주행속도, AGV 평균 지연 시간 및 주행 시간이 향상되었음을 보임으로써 제안한 방법이 AGV의 주행효율과 공간 활용률을 개선하였다.

최단 경로 기반 AGV 경로 결정에서 발생하는 정체현상을 경감하기 위해서 Lim et al.[13]은 강화학습(Reinforcement learning) 기반 동적 경로 결정을 제안하였다. 해당 연구에서는 Q-Learning 기반 경로 결정 방법을 제안하고 랜덤하게 생성된 Guided path network 시물레이션을 통해 효과성을 확인하였다. Jeon et al.[7]은 항만터미널 내 안벽영역과 야드영역의 컨테이너 반송을 담당하는 AGV의 Q-Learning 기반 동적 경로 결정 알고리즘을 제안하였다. 시물레이션 실험 결과, 제안한 방법은 안벽크레인의 할당 및 컨테이너 흐름에 따라 STDF 대비 3.2%에서 24.6%까지 개선하였다.

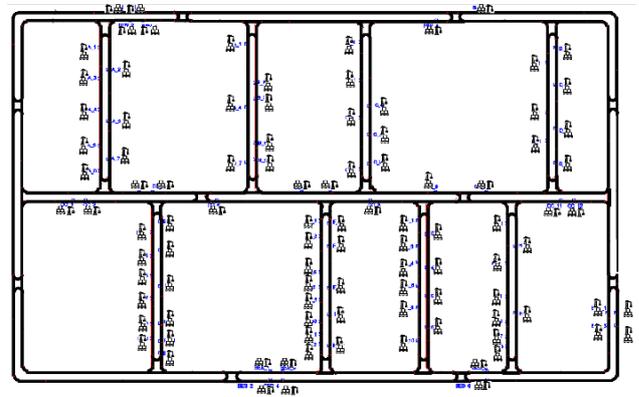
문헌 연구를 통해 파악한 본 연구의 기여는 다음과 같다: (1) 본 연구는 반송 중인 OHT가 정체 현상을 야기할 수 있는 OHT 대수를 고려하여 경로를 동적으로 결정하는 Q-Learning 알고리즘을 제안하였다. 이를 통해 OHT는 목적지까지의 정체 패턴을 고려하여 정체가 빈번한 구역을 우회하였다; (2) 시물레이션 실험을 통해 제안한 방법이 벤치마크 알고리즘 대비 반송 신속성과 안정성을

개선시킴을 확인하였다.

### 3. 반도체 자동화 물류시스템

#### 3.1 Fab AMHS 레이아웃

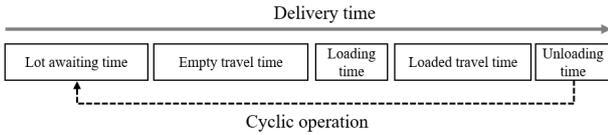
Fab AMHS의 레이아웃은 OHT, Overhead rail, Stocker, 가공설비로 구성되어 있다. 웨이퍼는 수백 번의 가공을 거치므로 Fab은 공정별 복수의 생산설비가 하나의 베이(Bay)에 위치하는 베이 구조를 따른다. Fab AMHS의 Overhead rail는 베이를 기준으로 크게 3가지로 구분할 수 있다: (a) 베이 간에 설비 흐름을 연결하는 인트라베이(Intrabay); (b) 베이 간의 흐름을 연결하는 인터베이(Interbay); (c) 단방향으로 베이를 둘러싸고 있는 외곽루프(Outer Loop). OHT가 반송 명령을 받으면 3가지 Overhead rail을 따라 Stocker 혹은 가공설비에 FOUP을 반송한다. 본 연구에서는 한국시물레이션 학회에서 교육용으로 제공되는 Fab AMHS 프로토타입 레이아웃을 이용하였다[18]. <Figure 1>은 본 연구에 활용된 Fab AMHS의 레이아웃이다.



<Figure 1> Layout of the Prototype Fab's AMHS

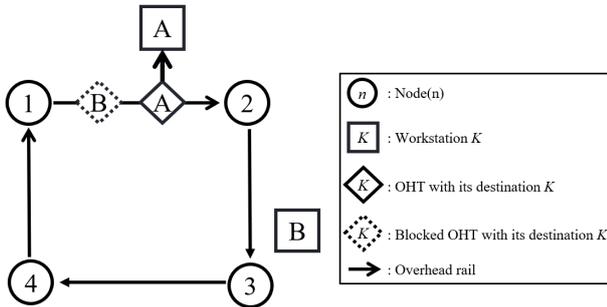
#### 3.2 OHT 반송작업 및 정체현상

<Figure 2>는 단위 반송당 FOUP의 흐름을 나타낸다. FOUP은 호출 가능한 OHT가 나올 때까지 대기하다, 호출 가능한 유휴 OHT가 나오면 호출한다. OHT는 반송요청을 받은 뒤 요청 설비까지 공차 상태로 이동하고, 출발 설비에서 FOUP을 적재한다. 이후 OHT는 FOUP의 최종 목적 설비까지 운반이동하고 FOUP을 적하함으로써 하나의 반송작업을 완료한다. FOUP의 반송시간은 '대기시간', '공차이동시간', '적재시간', '운반시간', '적하시간'의 합이다. 짧은 반송시간은 AMHS의 생산성을 개선할 수 있다.

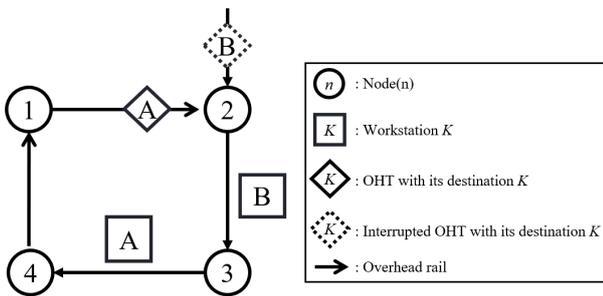


<Figure 2> Material Flow in the Fab's AMHS

OHT 간 정체가 존재하지 않을 때 OHT의 공차이동시간과 운반시간은 자체 흐름을 통해 추정할 수 있으나, 동적으로 변동하는 정체 상황으로 인하여 FOUP의 반송시간은 변동된다. OHT 정체는 블로킹 현상과 간섭 현상으로 구분할 수 있다. 블로킹 현상은 반송 중인 OHT가 적재/적하 등의 이유로 정차 중인 OHT에 의해 정지하게 되는 현상을 말한다. 간섭 현상은 반송 중인 OHT가 주행 중 다른 OHT와 충돌 방지 목적의 안전거리 확보를 위해 감속하는 현상을 말한다. <Figure 3>은 OHT 블로킹 현상의 예시이다. 목적지 설비 B로 반송 중인 OHT는 설비 A에서 적재/적하 작업 중인 다른 OHT에 의하여 정지한 상태이다. 발생한 블로킹으로 인하여 목적 설비 B로 반송 중인 OHT는 설비 A의 적재작업이 완료된 후에 주행이 가능하다.



<Figure 3> OHT Blocking Phenomenon



<Figure 4> OHT Interruption Phenomenon

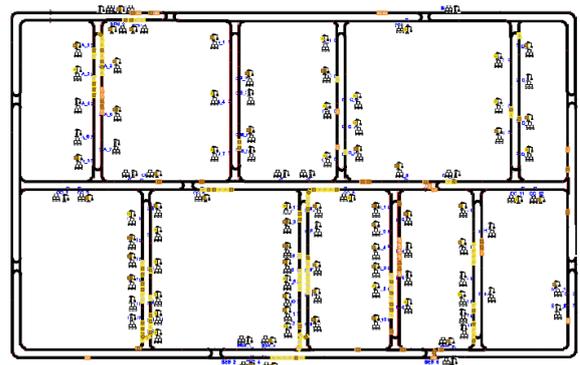
<Figure 4>는 OHT 간섭 현상의 예시이다. 목적지 B로 주행 중인 OHT는 다른 레일에서 진입하는 OHT와 충돌을 방지하기 위하여 감속하게 된다. 두 OHT의 거리가 안전거리만큼 확보될 때까지 감속을 멈추고, 이후 가속을 진행하여 최대 주행속도를 유지하게 된다. 본 연구에

서는 시간의 흐름에 따라 변동하는 블로킹 현상과 간섭 현상을 반영하는 동적 경로 결정 알고리즘을 통하여 AMHS의 생산성을 개선하고자 한다.

### 3.3 Fab AMHS 시뮬레이션 모델링

Fab AMHS의 OHT의 반송 능력은 공차관리 규칙, 차량배차 규칙, 차량의 대수, OHT 감·가속도, OHT 크기, OHT 회전속도 등에 영향을 받는다. 또한, 수백 대의 OHT가 상호 간에 미치는 정체현상은 Fab AMHS의 복잡성과 불확실성을 가중하므로 수리적 추정이 어렵다. 이에 Fab 운영자들은 시뮬레이션을 이용하여 Fab AMHS의 운영을 모사함으로써 생산성을 추정한다.

AMHS 시뮬레이션의 레이아웃은 연구용 프로토타입 레이아웃으로 900여 개의 공정을 요구하는 단일 제품만을 고려하였다. 반송요청은 1초마다 지수분포를 따라 생성하였다. AMHS는 반송작업이 요구될 때마다 반송을 요구한 FOUP과 가장 가까운 유틸 차량을 선정하였다. 적하작업을 마친 유틸 차량은 베이 내 정체 감소를 위하여 다음 작업이 요구될 때까지 외곽루프로 순환시키는 규칙을 사용하였다. OHT 간 블로킹현상을 파악하기 위하여 OHT의 적재/적하 시간은 25초를 확정적으로 설정하였다. OHT의 직선 최대속도와 회전각 최대속도는 각  $4m/s$ 와  $0.8m/s$ 로 설정하였다. 또한, OHT 간 정체현상을 반영하기 위하여 OHT의 감속도는  $2.4m/s^2$ , 가속도는  $1.6m/s^2$ 로 설정하였다. OHT의 길이는 35cm로 설정하였다. <Figure 5>는 본 연구에서 사용된 시뮬레이션의 스냅샷이다.



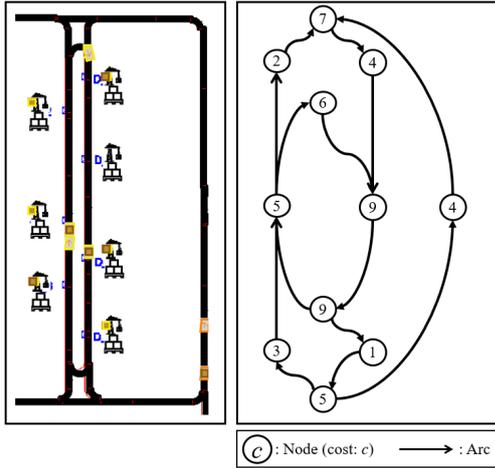
<Figure 5> Snapshot of the Fab's AMHS Simulation

## 4. 정체경로 우회 위한 동적 경로 결정

### 4.1 Fab AMHS 내 OHT 경로 결정 문제

<Figure 6>는 OHT의 차량 경로 결정 문제를 도식화하기

위하여 AMHS를 그래프로 표현한 것이다. AMHS 내 각 레일들은 갈림길을 기준으로 구분하고 노드(Node)로 표현하였다. 아크(Arc)는 노드 간의 연결 관계를 나타낸다. OHT의 경로는 목적 설비가 위치한 노드까지 방문한 노드의 집합으로 구성된다. OHT의 경로 결정은 목적 노드까지의 노드집합(경로)의 비용(반송시간)을 최소화하는 문제이다.

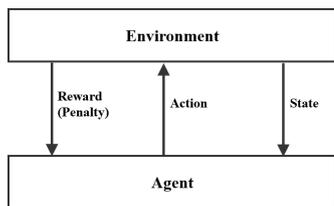


<Figure 6> Graphical Representation of the Fab's AMHS

OHT 경로 결정을 위한 방안에는 정적 경로 결정과 동적 경로 결정이 존재한다. 정적 경로 결정 문제는 노드의 비용이 확정적으로 결정되는 정적 환경에서 경로를 선택하는 문제이며, 동적 경로 결정 문제는 노드에 대한 비용이 시간의 흐름에 따라 변동하는 동적 환경에서 경로를 선택하는 문제이다.

4.2 Q-Learning 기반 동적 경로 결정 알고리즘

Q-Learning은 머신러닝 기법 중 한 종류로 어떤 환경(Environment)에서 에이전트(Agent)가 현재의 상태(State)를 인식하여 모험적 탐색을 통해 의사결정에 대한 보상(Reward)을 학습한다. 이후, 에이전트는 과거 학습을 바탕으로 가능한 의사결정 중 기대 보상을 최대화하거나 페널티를 최소화하는 의사결정을 도출한다[2, 20]. <Figure 7>은 Q-Learning의 구조이다[14].



<Figure 7> Structure of Q-Learning Algorithm

OHT는 노드마다 학습된 보상을 바탕으로 다음 노드를 선택하게 된다. OHT는 하나의 반송작업이 마칠 때마다 반송 중 방문했던 노드들에 대한 보상 값을 업데이트 한다. 학습이 끝난 뒤 획득한 보상 값을 바탕으로 OHT는 다음 경로를 결정한다. Q-Learning(QL) 알고리즘에 사용된 노테이션은 <Table 1>과 같이 정의하였다[7, 14].

<Table 1> Notations Used in the QL

$t$	The destination node of a current vehicle
$k$	The node where a current vehicle is located
$(k, t)$	The state of the vehicle, which consists of its current node ( $k$ ) and its destination node ( $t$ )
$A(k, t)$	The set of candidates for the next node (action), from which a vehicle in state $(k, t)$ can choose
$a$	An action taken by a vehicle, which is an element of $A(k, t)$
$\gamma$	The discount factor for future penalties
$r[(k, t), a]$	The penalty, which is the travel time of a vehicle, at $(k, t)$ , from the current node to the next node ( $a$ )
$Q[(k, t), a]$	The expected discounted cumulative travel time of a vehicle, which selects action $a$ at state $(k, t)$ from the current node to the destination node

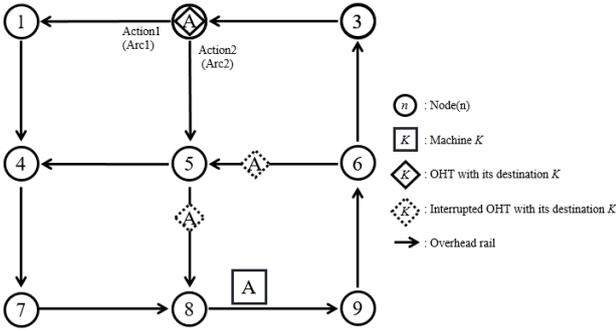
4.3 OHT간 정체패턴을 고려한 상태모델링

효과적인 Q-Learning 기반 의사결정을 위해서는 대상 시스템의 복잡도와 의사결정 목적에 부합한 상태 정의가 요구된다. 상태 조합 수가 증가하는 경우는 모든 상태에 대한 충분한 학습이 어렵고, 반대로 상태 조합 수가 감소하는 경우는 충분한 학습이 이루어져도 정밀한 의사결정이 어렵다[13].

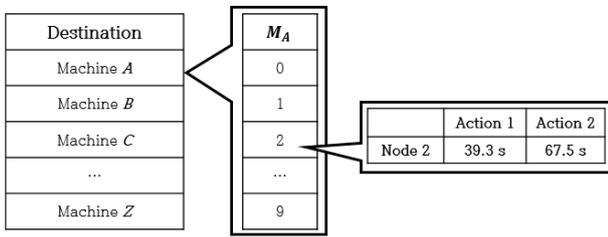
본 연구에서는 AMHS 내 OHT 간 정체 패턴을 고려하기 위한 상태모델링을 제안한다. 블로킹 현상과 간섭 현상을 야기할 수 있는 기대 정체 유발차량 대수( $M_t$ )를 정의하고, 상태모델링에 반영하여 블로킹과 간섭 현상에 대한 영향도를 간접적으로 상태에 반영하고자 하였다.  $M_t$ 은 에이전트(의사결정 주체인 OHT)가 노드를 방문하였을 때 같은 목적지( $t$ )를 가진 OHT 중 더 짧은 잔여 거리를 가진 OHT들의 수이다.

<Figure 8>는  $M_t$  산정에 대한 예시이다. 2번 노드에 진입한 에이전트 OHT(실선)는 같은 목적지를 가진 OHT 중 더 짧은 잔여 거리를 가진 OHT(점선)의 개수를  $M_t$ 로 결정하게 된다. <Figure 9>는 <Figure 8>의 에이전트 OHT에 대한 Q-Table 예시이다.

본 연구에서 제안하는 정체유발 차량대수를 고려한 Q-Learning(Q-Learning with the expected number of congested vehicles, QECV)에 사용된 Notation은 <Table 2>와 같다.



<Figure 8> Example of a Traffic Situation



<Figure 9> Example of a Q-Table in the QECV

<Table 2> Notations used in the QECV

$t$	The destination node of a current vehicle
$k$	The node where a current vehicle is located
$M_t$	The number of vehicles traveling to the destination node $t$ but closer to the destination node than the current vehicle
$(k, t, M_t)$	The state of the vehicle, which consists of its current node ( $k$ ), its destination node ( $t$ ), and the number of vehicles traveling to the destination node ( $t$ ) but closer to the destination node than the current vehicle ( $M_t$ )
$A(k, t, M_t)$	The set of candidates for the next node (action), from which a vehicle in state $(k, t, M_t)$ can choose
$a$	An action taken by a vehicle, which is an element of $A(k, t, M_t)$
$\gamma$	The discount factor for future penalties
$r[(k, t, M_t), a]$	The penalty, which is the travel time of a vehicle, at state $(k, t, M_t)$ , from the current node to the next node ( $a$ )
$Q[(k, t, M_t), a]$	The expected discounted cumulative travel time of a vehicle, which selects action ( $a$ ) at state $(k, t, M_t)$ from the current node to the destination node

OHT는 매 노드에 진입하는 순간 다음 노드에 대한 선택을 Q 값을 바탕으로 결정하도록 하였으며, 반송이 마칠 때마다 Q 값을 아래와 같이 업데이트한다.  $visits_n$ 은 학습기간 동안 OHT가 상태( $t, k, M_t$ )에서 의사결정  $a$ 를 내린 총 횟수이다.

$$\hat{Q}[(k, t, M_t), a] = (1 - \alpha_n) \hat{Q}_{n-1}[(k, t, M_t), a] + \alpha_n \{r[(k, t, M_t), a] + \gamma \min_a \hat{Q}_{n-1}[(a, t, M_t), a']\},$$

where

$$\alpha_n = \frac{1}{1 + visits_n[(k, t, M_t), a]}$$

다음 노드에 대한 선택 확률은 아래와 같이 산출하였다.

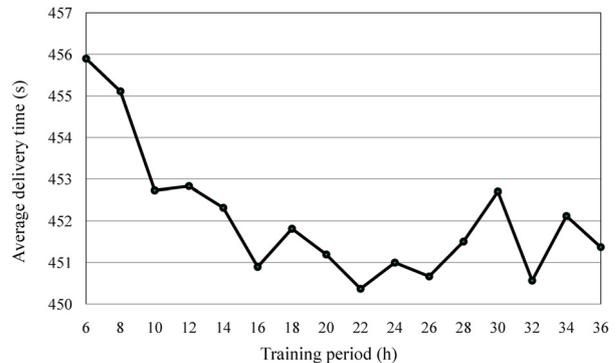
$$p(a | (k, t)) = \frac{\rho^{\hat{Q}[(k, t, M_t), a]^{-1}}}{\sum_{a \in A(k, t, M_t)} \rho^{\hat{Q}[(k, t, M_t), a]^{-1}}}$$

### 5. 시뮬레이션 실험

본 연구에서는 제안한 QECV의 효과성을 평가하기 위하여 정적 경로 결정 알고리즘인 최단거리 알고리즘(STDF)과 4장 2절에 소개된 QL을 이용하였다. OHT 반송 능력을 평가하기 위해서 시간당 반송량, 평균 반송시간, 반송당 이동거리를 산출하였다. 3.3절에 소개된 AMHS 시뮬레이션의 모델링을 위해 Tecnomatix Plant Simulation 12.0(64-bit, Version 12.0.3)을 사용하였다. 시뮬레이션 실행환경 CPU는 Intel(R) Core(TM) i5-11600K @ 3.90GHz, RAM은 32.00GB이다.

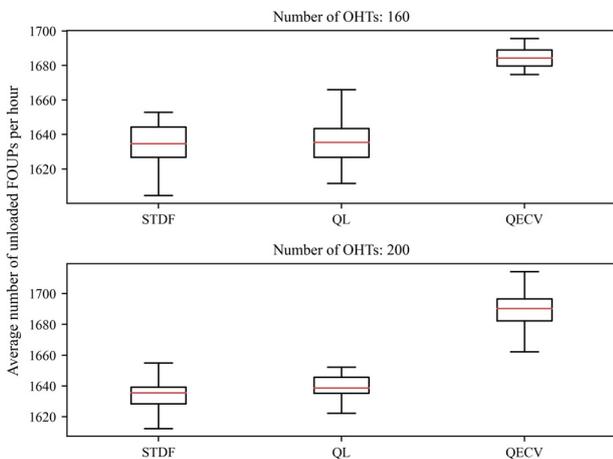
시뮬레이션 Warm-up time은 사전실험을 통해 6시간으로 정하였다. 성과지표는 Warm-up time 이후 12시간 동안 관측하였다. 시뮬레이션 실험의 통계적 유의성 확보를 위해 실험당 30번의 반복실험을 수행하여 평가지표를 분석하였다.

<Figure 10>은 OHT가 160대일 때 QECV 학습시간에 따른 평균 반송시간 추이이다. 학습 시간이 6시간부터 14시간까지 증가하면서 반송시간의 유의미한 감소를 보였다. 반송시간이 16시간 이후부터는 반송시간의 표준편차는 0.70으로 낮은 값을 보였다. 이를 바탕으로 QECV의 학습시간을 24시간으로 고정하여 실험을 진행하였다.



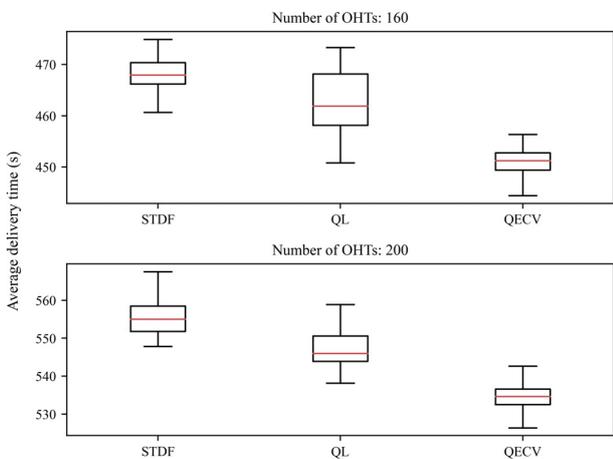
<Figure 10> Average Delivery Time with Different Training Periods

정체 정도에 따른 제안한 방법의 우수성을 평가하고자 OHT 대수를 160대와 200대로 변경하며 성과지표를 분석하였다. <Figure 11>은 OHT 대수에 따른 시간당 반송량이다. OHT가 160대일 때, QL은 정체회피를 통해 STDF 대비 0.14%의 반송량 개선을 이루었지만, 반송량의 표준편차는 2.99% 증가되었다. QECV는 STDF 대비 평균 3.16% 반송량을 개선시켰으며, 반송량의 표준편차는 38.94% 감소시켰다. 제안된 QECV는 기대 정체 유발차량 대수를 상태에 반영함으로써 효과적이고 안정적인 반송량을 보였다. OHT가 200대일 때, QL과 QECV는 STDF 대비 평균 0.13%, 3.22% 반송량 개선을 이루었다.



<Figure 11> Average Number of Unloaded FOUPs with Different Number of OHTs

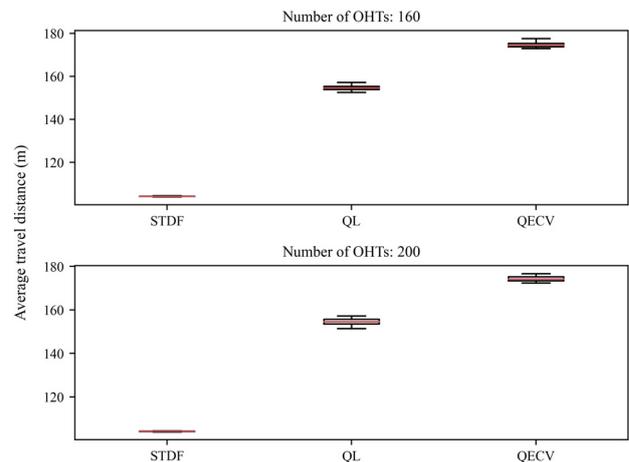
<Figure 12>는 OHT 대수에 따른 두 알고리즘의 평균



<Figure 12> Average Delivery Time with Different Number of OHTs

반송시간이다. OHT가 200대인 환경에서 QL은 STDF 대비 평균 1.35% 반송시간 개선을 보였다. 또한, QL은 STDF 대비 반송시간의 표준편차를 25.99% 증가시켰다. QECV는 STDF 대비 3.54%의 반송시간 개선을 보였고, 표준편차를 40.41% 개선시켰다. OHT가 200대일 때, 제안한 QECV는 반송시간의 평균 및 표준편차를 3.62%, 18.24% 각각 개선할 수 있었다. 이는 기대 정체 유발차량 대수를 상태에 반영함으로써 OHT가 정밀히 정체 상황을 파악한 결과이다. 서로 다른 OHT 대수를 가진 환경에서 제안된 QECV는 STDF 및 QL 대비 짧고 안정적인 반송시간을 보였다.

<Figure 13>은 OHT 대수에 따른 두 알고리즘의 반송당 평균 이동거리이다. OHT가 160대일 때, QL은 STDF 대비 48.36% 증가한 반송당 평균거리를 보였으며, QECV는 STDF 대비 67.58% 증가한 반송당 평균거리를 보였다. OHT가 200대일 때, QECV는 QL 대비 12.86% 긴 반송당 이동거리를 가졌다. STDF는 최단거리만을 이용하였기 때문에 반송거리 표준편차는 0.12로 낮았다. 하지만, QL 및 QECV는 STDF 대비 반송당 표준편차가 8.8배 및 7.2배 상승하였다. 이는 QL과 QECV를 통해 OHT가 정체량에 따라 동적으로 반송경로를 결정한 결과이다. <Figure 11>의 반송시간 결과를 볼 때, QECV는 QL 및 STDF 대비 우회로를 효과적으로 이용하여 장거리 반송거리를 이용하여 단시간 반송을 보였다.



<Figure 13> Average Travel Distance with Different Number of OHTs

<Table 3>은 서로 다른 경로 결정 알고리즘과 OHT 수에 따른 반송시간에 대한 유의성 검정 결과이다. OHT 대수 변경은 모든 실험환경에 유의미한 변화를 발생시켰다. QECV는 OHT가 160대와 200대인 환경에서 STDF 및 QL 대비 유의미한 개선을 이루었다.

&lt;Table 3&gt; The P-values of t-test on Delivery Times

	QL-160	QECV-160	STDF-200	QL-200	QECV-200
STDF-160	4.21E-04	7.28E-26	3.00E-53	1.52E-42	4.80E-54
QL-160	-	1.24E-09	3.23E-48	7.98E-47	3.74E-40
QECV-160	-	-	1.33E-50	3.26E-41	1.23E-55
STDF-200	-	-	-	1.63E-05	5.60E-23
QL-200	-	-	-	-	6.37E-12

&lt;Table 4&gt; The P-values of t-test on Travel Distances

	QL-160	QECV-160	STDF-200	QL-200	QECV-200
STDF-160	3.74E-47	8.25E-54	7.31E-01	1.16E-44	7.21E-53
QL-160	-	3.26E-53	1.01E-47	6.60E-01	5.91E-53
QECV-160	-	-	9.52E-55	9.57E-48	3.67E-01
STDF-200	-	-	-	4.95E-45	1.10E-53
QL-200	-	-	-	-	3.96E-48

<Table 4>는 서로 다른 경로 결정 알고리즘과 OHT 수에 따른 반송거리에 대한 유의성 검정 결과이다. STDF 및 QL은 OHT의 대수가 증가함에도 불구하고 유의미한 경로변경이 일어나지 않았다. 이는 QL이 OHT가 증가함에 따라 가중된 정체패턴을 유의미하게 고려하지 못 했음을 의미한다. 제안된 QECV은 OHT 대수 증가에 따라 유의미한 반송거리 변화를 보였다.

상태를 표현하는 Feature들을 능동적으로 추출하여 상태에 반영할 수 있는 Deep Q Network 등 복잡한 Fab 환경 변화를 고려한 알고리즘이 논의될 수 있다.

## Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government(MSIT) (No. NRF-2020R1A2C2004320).

## 6. 결론

본 연구에서는 Fab AMHS의 생산성 개선을 위한 OHT의 Q-Learning 동적 경로 결정 알고리즘을 제안하였다 (Q-Learning with the expected number of congested vehicles, QECV). 경로를 결정할 에이전트 OHT는 정체를 유발시킬 가능성이 높은 정체 유발차량 대수( $M_t$ )를 파악함으로써 OHT가 주행 중 발생할 정체패턴을 정밀히 대응하고자 하였다.

제안한 QECV의 효과성을 파악하기 위하여 OHT 대수를 변경하며 비교 알고리즘 STDF 및 QL의 성과지표를 분석하였다. 시뮬레이션 실험을 통해 QECV는 OHT가 정체가 적은 우회로를 선택하도록 유도함을 확인하였다. 이에 따라 제안한 알고리즘은 최근거리 알고리즘 대비 67.58% 이상 높은 이동거리를 보였지만, 반송시간 측면에서는 3.54% 개선하였다. 또한, QECV는 반송시간의 표준편차를 33.20% 감소시켜 반송능력의 안정성을 향상시켰다. 이를 통해 제안된 QECV가 Fab AMHS 내 정체패턴을 고려하여 반송 신속성과 반송 안정성을 개선할 수 있음을 확인하였다.

향후 연구에서는 생산물류 통합 시뮬레이션을 바탕으로 동적으로 변화하는 Fab 생산설비 상태를 고려한 OHT 동적 경로 결정 알고리즘이 논의될 수 있다. 또한, Q-Learning의

## References

- [1] Bartlett, K., Lee, J., Ahmed, S., Nemhauser, G., Sokol, J., and Na, B., Congestion-aware dynamic routing in automated material handling systems, *Computers & Industrial Engineering*, 2014, Vol. 70, pp. 176-182.
- [2] Farahnakian, F., Ebrahimi, M., Daneshlab, M., Liljeberg, P., and Plosila, J., Q-learning based congestion-aware routing algorithm for on-chip network, *IEEE 2nd International Conference on Networked Embedded Systems for Enterprise Applications*, Perth, Australia, 2011, pp. 1-7.
- [3] Ganesharajah, T., Hall, N.G., and Sriskandarajah, C., Design and operational issues in AGV-served manufacturing systems, *Annals of operations Research*, 1998, Vol. 76, pp. 109-154.
- [4] Ghiani, G., Manni, E., Quaranta, A., and Triki, C., Anticipatory algorithms for same-day courier dispatching, *Transportation Research Part E: Logistics and Transportation Review*, 2009, Vol. 45, No. 1, pp. 96-106.
- [5] Han, G., Kang, B., Kim, H.J., and Hong, S., A GA-based

- Optimization of a Weighted Lot Targeting Rule in a Semiconductor Wafer Fab, *Journal of the Korean Institute of Industrial Engineers*, 2022, Accepted.
- [6] Hong, S., Hwang, I., and Jang, Y., Practical Q-learning-based Route-guidance and Vehicle Assignment for OHT Systems in Semiconductor Fabs, *IEEE Transactions on Semiconductor Manufacturing*, 2022.
- [7] Jeon, S.M., Kim, K.H., and Kopfer, H., Routing automated guided vehicles in container terminals through the Q-learning technique, *Logistics Research*, 2011, Vol. 3, No. 1, pp. 19-27.
- [8] Kang, J., Choi, R., Kang, B., Ryu, K.R., and Kim, K. H., Collision Avoidance and Deadlock Resolution for AGVs in an Automated Container Terminal, *Korea Intelligent Information Systems Society*, 2005, Vol. 11, No. 3, pp. 25-43.
- [9] Kim, B.-I., Shin, J., and Jeong, S., Assignment Problem Based Vehicle Dispatching for a Semiconductor FAB, *Proceedings of the Korean Operations and Management Science Society Conference*, 2006, Seoul, Korea, pp. 525-528.
- [10] Kim, D.S. and Choi, J.Y., An Efficient Algorithm for Improving Detour in OLED FAB, *Journal of Korean Society of Industrial and Systems Engineering*, 2018, Vol. 41, No. 3, pp. 120-128.
- [11] Kim, H.J. and Lee, J., A WIP Allocation Approach for Enhancing On-Time Transfers of Wafer Lots in a Semiconductor Fabrication Facility, *Korean Operations Research And Management Society*, 2021, Vol. 38, No. 1, pp. 47-55.
- [12] Kim, H.J., Lim, D.E., and Lee, S.M., Development of Deep-Learning-Based Scheduling Approach to Improve the Automated Material Handling System Throughput Capacity in a Semiconductor Manufacturing Facility, *Journal of the Korea Management Engineers Society*, 2019, Vol. 24, No. 2, pp. 73-83.
- [13] Kim, H.-W. and Lee, W.-C., Real-Time Path Planning for Mobile Robots Using Q-Learning, *Journal of IKEEE*, 2020, Vol. 24, No. 4, pp. 991-997.
- [14] Lim, J.K., Kim, K.H., Lim, J.M., Yoshimoto, K., and Takahashi, T., Routing Automated Guided Vehicles Using Q-Learning, *Journal of Japan Industrial Management Association*, 2003, Vol. 54, No. 1, pp. 1-10.
- [15] Maxwell, W.L. and Muckstadt, J.A., Design of automatic guided vehicle systems, *IIE Transactions*, 1982, Vol. 14, No. 2, pp. 114-124.
- [16] Min, K.H. and Yun, I.H., A study on Distance Allocation Method of the Faster Move Logistics between the Semiconductor Production Equipment, *Proceedings of the Korean Society of Precision Engineering Conference*, 2014, Jeju, Korea, pp. 1102-1103.
- [17] Siebert, M., Bartlett, K., Kim, H.J., Ahmed, S., Lee, J.-H., Nazzal, D., Nemhauser, G., and Sokol, J., Lot targeting and lot dispatching decision policies for semiconductor manufacturing: Optimisation under uncertainty with simulation validation, *International Journal of Production Research*, 2018, Vol. 56, No. 1-2, pp. 629-641.
- [18] The Korea Society for Simulation, <https://www.simulation.or.kr/html/?pmode=BBBS0009500005&page=1&smode=view&seq=263&searchValue=&searchTitle=strTitle>.
- [19] Wan, J. and Shin, H., Predictive vehicle dispatching method for overhead hoist transport systems in semiconductor fabs, *International Journal of Production Research*, 2022, Vol. 60, No. 10, pp. 3063-3077.
- [20] Watkins, C.J. and Dayan, P., Q-learning, *Machine learning*, 1992, Vol. 8, No. 3, pp. 279-292.

#### ORCID

Bonggwon Kang | <http://orcid.org/0000-0001-7367-9817>

Byeong Min Kang | <http://orcid.org/0000-0002-9859-3734>

Soondo Hong | <http://orcid.org/0000-0001-7817-6776>