

딥러닝을 활용한 반도체 제조 물류 시스템 통행량 예측모델 설계

A Deep Learning-Based Model for Predicting Traffic Congestion in Semiconductor Fabrication

김종명^a, 김옥현^b, 홍성빈^b, 임대은^{b*}

Jong Myeong Kim^a, Ock Hyeon Kim^b, Sung Bin Hong^b, Dae-Eun Lim^{b*}

^a Environmental Technology Division, Korea Testing Laboratory(KTL) (08389) 87, Digital-ro 26-gil, Guro-gu, Seoul, Republic of Korea

^b Division of Architectural, Civil, and Environmental Engineering, College of Engineering, Kangwon National University Chuncheon, Gangwon, 24341, Republic of Korea

Received 20 September 2019; Revised 16 October 2019; Accepted 21 October 2019

Abstract

Semiconductor logistics systems are facing difficulties in increasing production as production processes become more complicated due to the upgrading of fine processes. Therefore, the purpose of the research is to design predictive models that can predict traffic during the pre-planning stage, identify the risk zones that occur during the production process, and prevent them in advance. As a solution, we build FABs using automode simulation to collect data. Then, the traffic prediction model of the areas of interest is constructed using deep learning techniques (keras - multistory conception structure). The design of the predictive model gave an estimate of the traffic in the area of interest with an accuracy of about 87%. The expected effect can be used as an indicator for making decisions by proactively identifying congestion risk areas during the Fab Design or Factory Expansion Planning stage, as the maximum traffic per section is predicted.

Keywords: Semiconductor, prediction model, AutoMod, deep learning

1. 서론

반도체 제조라인(이하 Fab)의 자동화율이 높아지면서 반도체 물류 시스템의 복잡도는 계속적으로 증가하고 있다. 이 때문에 생산량 증대 시 기대 산출량(throughput)에 따라가지 못하는 경우가 발생하는데, 이는 생산량을 높이면 자동화된 물류운반 시스템(automated material handling system, AMHS)에 대한 운반 요구량이 많아지고, 결국 이동량도 많아지기 때문이다. 많은 이동횟수는 AMHS 내 정체(congestion)현상을 발생시켜 전체 생산능력을 떨어뜨릴 수 있다. 즉, 정체현상은 전체 Fab의 생산능력을 떨어뜨리는 원인으로 꼽히고 있다. 이러한 정체현상 방지를 위해서는 AMHS의 처리능력을 평가할 수 있는 지표가 있어야 한다. 하지만 많은 변수가 상호작용하고 반도체 제조공정 자체가 복잡하기 때문에 지표의 개발이 어렵다.

이러한 지표의 개발을 위해서 많은 연구들이 있다. 기존 연구들을 볼 때, 실시간을 가정하고 반도체 AMHS의 처리 능력을 분석하는 연구는 활발히 진행되고 있다. 그 중에서 딥러닝(deep learning)을 활용하여 OHT(Overhead Hoist Transfer, Fig. 1 참고)의 정체현상을 분석한 연구들이 많이 나오고 있다. 딥러닝은 여러 비선형 변환 기법의 조합을 통해 높은 수준의 추상화를 시도하는 기계학습(machine learning)알고리즘의 집합으로 정의된다. 이러한 알고리즘을 통해 주로 속도를 예측하는 연구가 많다. 선행연구에서는 입력 변수로 이전 시점과 인접 구간의 속도 수준, 유량(count)을 이용했고, 예측 변수로는 다음 시점의 관심 구간의 속도 수준을 예측하는 모델이 소개되었다¹⁾.

본 논문에서는 기존 연구들과 달리 Fab이 가동 되기 전을 가정하고, 제조라인을 기획하는 단계에서 한정된 정보만을 이용해 사전에 생산능력을 예측하는 모델을 구축하여 AMHS 능력을 평가하는

* Corresponding author. Tel.: +82-33-250-6282

fax: +82-33-259-5547

E-mail address: del@kangwon.ac.kr (Dae-Eun Lim).

방법을 개발하는 것을 목표로 한다. 여기서 AMHS의 처리능력은 OHT의 최대 통행량 (통과대수)으로 정의하여 연구방향을 설정한다.



Fig. 1 OHT

2. 연구 배경

2장에서는 본 연구의 배경이 되는 Fab과 AMHS를 살펴보고 모델링 방향에 대해 논의한다. 또한 주요 연구방법론인 딥러닝에 대해서도 살펴본다.

2.1 Fab

반도체를 생산하기 위해서는 wafer 가공과정을 거친다. 산화막 형성공정, 포토공정, 식각(etch)공정, 증착공정등의 세부공정으로 나뉘는데 이 공정들을 이루는 제조시스템을 Fab(Fabrication)이라 부른다. Fab은 반도체 전체 공정의 80%를 차지하며 웨이퍼의 표면에 여러 종류의 막을 형성한 후 이미 만든 마스크를 이용해 특정 부분을 선택적으로 깎아 내는 작업을 되풀이함으로써 전자회로를 구성해나가는 전체 과정이다. Fab은 jobshop 방식으로 순서와 상관없이 반복적으로 공정이 진행된다.

Fab을 구성하는 요소로는 장비(equipment)와 AMHS로 구성되는데, AMHS는 다시 자동화된 운반 수단인 OHT와 OHT가 이동하는 레일 (rail) 등으로 이루어져 있다. 장비는 각 공정마다 그 기능이 다르며 각 장비마다 출입구 역할을 하는 포트 (port)가 2개씩 배치되어 있다(Fig. 2).

2.2 다층 퍼셉트론

다층 퍼셉트론(Multi Layer Perceptron, MLP)이란 입력층과 출력층 사이에 하나 이상의 은닉층이 존재하는 신경망(Fig. 3)이다. 다층 퍼셉트론은 각각의 x 값에 적절한 가중치 w 를 부여하고 상황에 맞는 활성화함수(activation function)를 활용하여 y 값을 도출하는 것을 의미한다. 단층 퍼셉트론은 XOR같은 상황에서는 직선하나로 분류가 불가능하다는 한계가 있다. 즉 선형일 때는 가능하지만 비선형일 경우에는 불가능하다. 이를 여러 개의 은닉층을 통해 이 한계를 극복한 모형이 다층퍼셉트론이다. 하지만 너무 많은 은닉층을 사용하게 되면 모형이 복잡해지며 학습시간 또한 급격히 늘어나기 때문에 적절한 층의 수를 정하는 것이 중요하다.

MLP모형의 구조와 역할에 대해 설명하면 크게 입력층, 한 개 이상의 은닉층, 출력층으로 구분되고 각각의 층에는 입력노드, 은닉노드, 출력노드가 존재한다. 선형모형이라고 가정할 때 입력노드는 독립변수이고 출력노드는 종속 변수라고 할 수 있고 은닉층에는 은닉노드가 존재한다.

MLP모형의 구성에 대해 설명하면 입력층은 가지고 있는 데이터 (독립변수)에 따라 노드의 수가 정해지고, 출력층은 구하고자 하는 목적에 따라 노드의 수가 정해진다. 예를 들어, 선형함수로 수치를 예측하고자 하면 1개의 y 를 구성하게 되며, 이진 분류나 다중 클래스 분류하고자 하면 2개 또는 3개 이상의 y 를 구성하게 된다.

각각의 층은 데이터 입력을 받을 때 가중치들의 합을 구하게 되고 활성화 함수를 통해 가중치의 값의 크기에 따라 출력하게 된다. 함수에 따라 모양과 쓰임새가 다르기 때문에 각각 상황에 맞게 적절한 활성화 함수를 사용 할 필요가 있다. 입력뉴런과 가중치로 계산된 결과 값이 그대로 출력으로 나오는 기본 함수인 linear, 은닉층에 많이 쓰이는 rectifier 함수, 이진 분류에 사용되는 sigmoid 함수, 다중 클래스 분류 문제에서 사용되는 softmax 함수 등이 있다. 또한 함수에 입력하는 범위가 크면 오차가 발생할 확률이 크기 때문에 데이터를 표준화 또는 정규화해서 입력 할 필요가 있다.

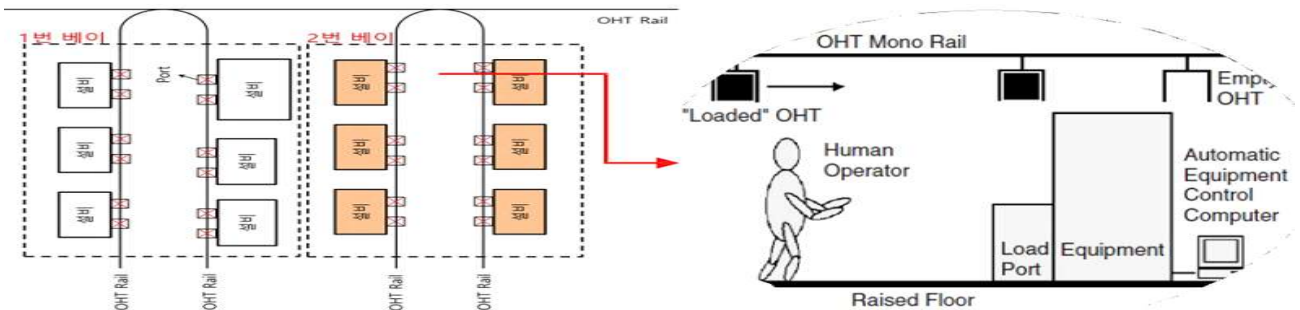


Fig. 2 FAB 내부 및 설비 구성 모습

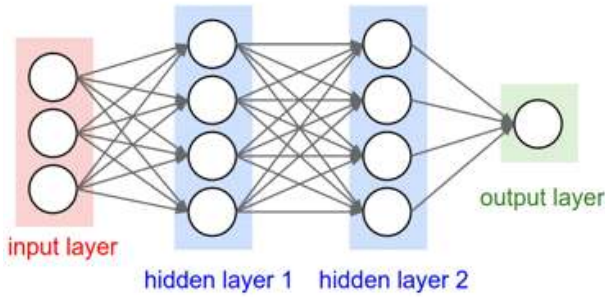


Fig. 3 다층 퍼셉트론의 구조

2.3 Automod 시뮬레이터

AMHS를 분석하기 위해서는 시뮬레이션을 주로 활용한다. 반도체 물류 시스템을 분석하는 시뮬레이션 도구 중 AutoMod를 많이 이용한다. 본 연구에서는 Automod 이용해 Fab을 구축(Fig. 4), 정체 현상 파악(Fig. 5)하여 예측모델에 필요한 데이터를 수집한다^[5].

이후 연구는 다음과 같이 진행된다. 먼저 수집한 데이터를 MS Excel을 이용해 전처리를 진행한 후 데이터의 경향과 탐색을 진행한다. 이 후 예측모델을 설계하기 위해 Keras 라이브러리를 이용한 다층 퍼셉트론구조를 활용한다.

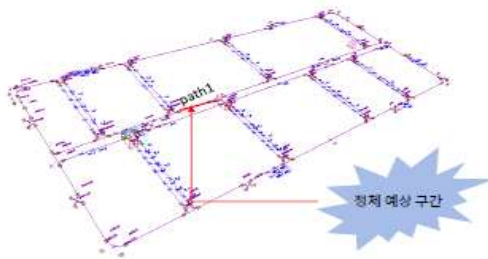


Fig. 4 오토모드를 활용한 Fab구조

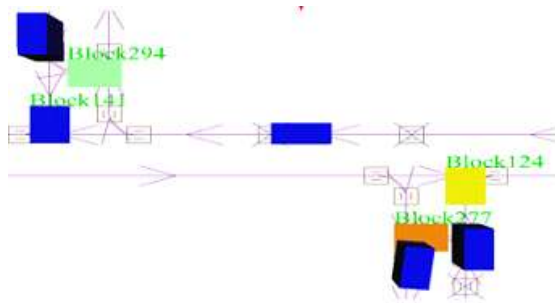


Fig. 5 오토모드를 활용한 Fab의 정체현상

3. 연구 절차

3.1 시뮬레이션 Fab 설계

반도체 물류 시스템을 구축하기 위해 Automod 시뮬레이션을 이용했다. 구성요인으로는 rail과 port이다. Rail은 직선부 rail과 곡선부 rail로 각각 948개, 260개, port는 782개로 구성되어있다. 시간대 마다의 통계량을 구하기 위해 5분, 10분, 15분 별로 설정을 하였고, 로트 (lot)의 발생 간격은 정규분포를 따르는 평균 10초, 분산 5초로 구성했다. port에 로트를 적재하고 하역하는 시간은 12초로, 정체현상을 유도하기 위해서 100대라는 대수를 임의로 설정했다. OHT의 규격은 2018년도 한국 대학생 시뮬레이션 경진대회 예선문제를 참조하였다.

Table 1 오토모드 시뮬레이션 실험 조건

구성요인	값
시뮬레이션 시간(min)	5,10,15
로트(lot)발생 간격(sec)	n(10,0.5)
UnLoading/Loading시간(sec)	12
OHT 대수	100
크기(가로*세로*높이)(mm)	700*450*500
직선부 속도(m/s)	4
곡선부 속도(m/s)	0.8
가속도(m/s)	1.6
감속도(m/s)	2.4

3.2 시뮬레이션 및 데이터 수집

데이터를 수집하기 위한 변수설계를 진행했다. 연구 배경에서 사전 fab의 정보만을 가지고 예측모델을 설계한다고 하였다. 이를 위해 정체현상을 발생시키는 주요 요인들을 설계적 요인과 운영적 요인을 구성하였다. 설계적 요인으로는 rail의 모양, 거리, port의 위치 등이 있고, 운영적 요인으로는 흐름량 정보로 하여 변수를 도출하였다. 본 논문에서 rail은 path, port는 cp로 부르겠다. Sector는 fab이 구분되는 지점으로부터 모양을 구성하였다. Path shape는 path의 곡선과 직선으로 구분지어 0 또는 1로 이진변환을 하였다. path distance는 좌표간의 직선거리 식을 이용하였다. cp distance는 port가 있는 지점부터해서 다음 지점까지의 거리를 나타낸다. Before after shape는 path의 전 후 모양이 직선인지, 곡선인지, 합류점인지, 분기점인지를 구분하였다. 운영적 요인은 작업횟수가 많을수록 그 부근의 정체현상이 많이 발생한다는 가정하여 한 출발port의 작업횟수, 도착port의 작업횟수와 cp whole은 출발설비와 목적설비의 작업횟수를 모두 합한 값을 변수로 지정하였다.

데이터를 구성하고 난 뒤 아래와 같이 필요한 데이터를 코드를 이용해 시뮬레이션을 구동하고 발생된 데이터를 MS Excel에서 전처리를 실시하였다.

Table 2 변수설계표

변수	요인 구분	변수 명	변수설명
독립 변수	설계적 요인	Sector	Divide zones from 0 to 10
		Path shape	Shape of Path (1-Line, 2-Curve)
		Path distance	Extract the x-y coordinate of the path to obtain the distance (straight line only)
		Cp distance	Distance between the port and the next
		Before after shape	Indicates the path (line, curve, junction, and breakpoint) by coding from 0 to 10
	운영적 요인	Cp start	Number of operations for starting equipment
		Cp goal	Number of operations of the purpose facility
Cp whole		Total number of operations	
종속 변수	Max	Maximum traffic of OHT	

3.3 모델의 구성

Sector, path shape, before_after 라는 변수를 범주화, 범주형은 category로 형변환을 한다. 그리고 정규화와 경사 하강법을 사용하였다. 다음으로 관측된 데이터 범위(99.7%) 밖이면 이상치로 판단하여 제거하였다. Train 데이터 수는 900개, Test데이터 수는 298개로, 7:3 비율로 훈련시킬 데이터와 Test할 데이터를 나누었다.

우리는 머신러닝 keras패키지의 다층 퍼셉트론(Multi-Layer Perceptron)로 모델링 하였고 Row 는 1198, Columns는 8로 구성하였다. 활성화 함수는 relu, 손실 함수는 mse, 알고리즘은 msprop, batch 사이즈는 20개로 모델을 구성하였다. 오버 피팅을 방지하기 위해 10 번 이상 손실 값에 변화가 없을 시 조기 종료하도록 모델링 하였다 [4].

3.4 예측모델 검증

다층퍼셉트론 구조(layer 수=6)를 이용하여 모델을 구축하였다. 아래의 그림은 모델의 검증결과과정을 그래프로 나타내었다. 이 과정은 예측치에 대한 최저 손실값을 찾는 과정이라 볼 수 있다. Loss값과 acc값으로 구성되어있으며 training loss/acc값과 test loss/acc값을 비교했을 때 일정 변동폭을 볼 수 있지만 그 폭이 높지 않다는 것을 볼 수 있다(Fig. 6).

본 연구는 FAB 기획단계에서 사전예방차원의 AMHS 평가방법을 개발하고 생산 시 위험구간을 파악하는 것을 목적으로 두고 있다. 이를 위해 머신러닝 기법 중 다층퍼셉트론 구조를 이용하여 위험구간을 예측하는 모델을 설계하였다. 모델검증과정에서 모델에 정확도는 87%가 나왔고 손실값이 0.3901이 나와 검증된 모델로써 효력을 발휘할 수 있다. 즉, 사전 FAB의 도면 정보와 구간 사이의 정보들을 통해 최대통행량을 예측할 수 있다고 결론 내릴 수 있다.

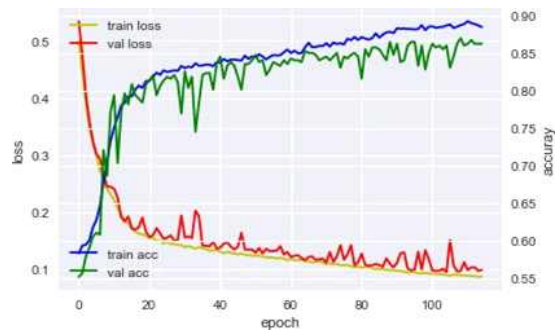


Fig. 6 예측모델의 결과

4. 결론

반도체 제조 시스템 생산도면 및 흐름량 정보만을 가지고는 통행량을 정확히 예측하기 어렵다. 하지만 전·후 path의 정보를 이용하면 관심 있는 path에 최대 통행량을 예측하는데 정확도를 높일 수 있었다. 여기서 전·후 path의 정보란, 관심 구간의 전 path와 후 path의 정보를 이야기하는데 path의 모양이나, 흐름량 등이 있다.

정확도를 더욱 올리기 위해서는 추가적인 변수가 필요하다. 왜냐하면 복잡한 반도체 제조 시스템에서 설계정보와 흐름량 정보만을 고려하였기 때문이다. 좀 더 여러 가지의 설계정보라든지 흐름에 영향을 미칠 수 있는 다른 변수가 필요할 것으로 생각된다. 또한 정확도를 올리기 위해서 전·후 path의 정보에서 전전·후후 rail까지 확대해 보다 더 많은 데이터를 추가하는 방법도 고려해 봐야 할

것이다. 다른 고려 사항으로는 다층 퍼셉트론모델만을 이용하여 반도체 제조 시스템의 최대 통행량을 예측해 보았는데, 향후에 예를 들어 RNN(Recurrent Neural Network)같은 다른 모델로 추가 분석모델을 이용해 정확도를 비교분석해 볼 필요가 있다.

본 연구에서는 기존 방법보다 광범위한 제조설비상황의 요인들을 고려하였다. fab의 각 구간에 대한 최대 통행량을 예측함에 따라 FAB설계 및 공장 증축 단계에서 정체 위험 구간을 사전에 파악해 기획 단계 신속한 의사결정에 도움을 줄 수 있어 반도체 수율 상승에 이바지할 수 있다. 이를 통해 복잡해지는 물류환경에서 막대한 비용으로 인한 제조 환경 개선을 지능화를 통해 저비용 고효율의 효과를 얻을 수 있다.

본 연구의 활용 방안은 다음과 같다. 반도체 제조물류시스템을 연구대상으로 하였기에 반도체 물류 환경이 구축된 기업에서 본 연구의 결과를 활용할 수 있고 기업에서 사용하고 있는 제조물류시스템 분석 방법에 본 연구결과를 기업 환경에 맞게 설정하여 평가 기능으로 추가해 기존 분석보다 더 정확하고 신속한 문제 대응하는데 활용할 수 있다. 또한 반도체 공정은 매우 복잡하기 때문에 연구에서 고려된 변수 외에 다른 변수들도 고려할 필요가 있기에 본 연구를 바탕으로 향후 다른 요인을 고려한 연구를 진행하는데 참고가 될 수 있다.

연구 결과 FAB 구간의 전/후 구간의 정보를 이용하면 정체위험 구간을 예측할 수 있다는 것을 알 수 있었다. 이에 더 앞서서 관심구간에 대한 전전/후후 등의 더 많은 데이터를 확보해 연구를 진행하는데 도움을 준다.

본 연구의 이용된 머신러닝 기법 중 다층 퍼셉트론 구조 기법을 활용하였는데 연구방법 범위를 넓혀 다른 분석 기법들을 본 연구에 활용하여 각 분석기법간의 정확도 및 손실값을 비교분석하는데 활용될 수 있다.

References

- [1] Kenji Kumagai, 2017, Maximum Transportation Throughput of Automated Guided Vehicle System by use of Models of Traffic Capacity and Traffic Capacity Consumption, Okayama, Japan.
- [2] Kelly K.Bartlett, 2014, Congestion-aware dynamic routing in automated material handling systems, Atlanta, Usa.
- [3] 강연국, 조성준, 2018, 딥러닝을 이용한 제조 물류시스템 Vehicle 구간 정체 예측, 한국경영과학회 학술대회논문집.

[4] 김태영, 2017, 블록과 함께하는 파이썬 딥러닝 케라스 이야기, 디지털북스.

[5] 김태복, 2014, 물류시뮬레이션 : AutoMod 활용방법을 중심으로, 박영사.