

준지도학습 기반 반도체 공정 이상 상태 감지 및 분류

이용호^{*}·최정은^{*}·홍상진^{*†}

^{*†}명지대학교 공과대학 전자공학과

Semi-Supervised Learning for Fault Detection and Classification of Plasma Etch Equipment

Yong Ho Lee^{*}, Jeong Eun Choi^{*} and Sang Jeen Hong^{*†}

^{*†}Department of Electronics Engineering, Myongji University

ABSTRACT

With miniaturization of semiconductor, the manufacturing process become more complex, and undetected small changes in the state of the equipment have unexpectedly changed the process results. Fault detection classification (FDC) system that conducts more active data analysis is feasible to achieve more precise manufacturing process control with advanced machine learning method. However, applying machine learning, especially in supervised learning criteria, requires an arduous data labeling process for the construction of machine learning data. In this paper, we propose a semi-supervised learning to minimize the data labeling work for the data preprocessing. We employed equipment status variable identification (SVID) data and optical emission spectroscopy data (OES) in silicon etch with $SF_6/O_2/Ar$ gas mixture, and the result shows as high as 95.2% of labeling accuracy with the suggested semi-supervised learning algorithm.

Key Words : Machine Learning, Semi-supervised Learning , Labelling, Fault Detection and Classification, Optical Emission Spectroscopy

1. 서 론

고집적 회로로 구성된 고성능 디바이스를 생산하기 위해서 반도체 소자의 미세화가 요구된다. 하지만 미세화 공정이 도입되면서 반도체 제조 공정 단계가 늘어나 제조과정은 더욱 복잡해지고 기술적 문제의 수도 증가하였다 [1]. 이러한 공정 환경의 변화에 대응하기 위해 반도체 제조공정에서는 빠르고 정확하게 장비의 이상상태를 감지하여 장비 정지 시간을 효율적으로 단축시킬 수 있는 방법이 요구된다. 이를 위한 방법으로 이상 감지 기술(Fault detection classification, FDC)은 반도체 제조업계에서 장비의 가동시간을 늘리고 테스트 웨이퍼 사용을 효과적으

로 줄이는 방법으로 인식되었다 [2].

과거 반도체 제조업계에서는 결함을 감지해내기 위한 시스템으로 통계적 공정 제어(Statistical Process Control, SPC)를 사용한다. 하지만 SPC의 단일 변수 분석은 공정이 복잡해지면서 이전보다 더 다양한 변수가 공정에 영향을 주는 최근 반도체 제조공정에서는 적합하지 않다. 주성분 분석(Principal Component Analysis, PCA)과 같은 방법으로 여러 변수에 대응할 수 있지만 원래 변수 정보의 대부분이 손실될 수 있다는 문제점이 있다. 이에 대한 해결책으로 다양한 변수에 보다 능동적으로 대처할 수 있는 기계학습을 적용한다면 복잡도가 증가하는 반도체 제조공정에서 효율적인 FDC 시스템을 구성할 수 있다 [3].

반도체 분야에서 기계학습은 제조공정의 효율성 증대를 위해 다양한 측면에서 활용되어왔다. 1) 계측 시간과

[†]E-mail: samhong@mju.ac.kr

비용을 감소시키기 위해 랜덤 포레스트와 같은 트리 양상을 방법을 기반으로 한 가상 계측(Virtual Metrology, VM) 모델을 사용한 연구가 진행되었다 [4]. 2) 또한 광 스펙트럼 분광분석기(Optical Emission Spectroscopy, OES)와 같은 플라즈마 측정 센서를 활용하여, 센서 데이터와 공정 결과의 복잡한 관계를 연결시킬 수 있는 인공 신경망 모델링이 제시되었다 [5]. 3) 수율 향상 및 검사 장비의 결합 검출 능력 향상을 위해 웨이퍼 맵의 이미지 분석이 가능한 합성곱 신경망(Convolutional neural network, CNN) 기반 모델 연구도 진행되었다[6-7]. 4) 모듈형 신경망 모델링(MNN)을 통하여 장비의 툴 데이터를 활용한 식각 장비의 이상 탐지 방법이 제시되기도 하였다 [8]. 이외에도 CNN를 적용하여 센서 신호에서 추출한 이상 변수에서 공정 이상을 탐지하는 방법이 제안되었다 [9].

그 중에서도 장비의 이상 상태를 감지하기 위한 FDC에 기계학습 알고리즘을 적용하기 위해서는 충분한 양의 실험 데이터가 필요하다. 하지만 집적회로 제조 기술의 빠른 미세화로 인해 제조 과정이 매우 복잡해지고 실험 비용이 크게 증가하면서 충분한 양의 실험 데이터를 얻기 위해서는 큰 비용이 요구된다. 뿐만 아니라 높은 성능의 기계학습 알고리즘을 구성하는 데에는 각각의 데이터에 대한 정의를 지칭하는 레이블링(labeling)이 필요하다. 반도체 공정 특성 상 실험데이터에 레이블을 지정하기 위해서는 공정 결과에 대한 계측이 필요하다. 미세한 반도체 공정 결과물에 대한 레이블을 얻기 위해 추가적인 계측 과정이 필요한데, 이는 작업자가 진행하면서 시간적으로도 오래 걸리고 운영자의 일관성을 보장하지 못한다는 문제가 있다 [10]. 따라서 본 연구에서는 실험데이터 중 일부에 대해서만 레이블을 지정하고 레이블이 지정되지 않은 데이터(unlabeled data)를 같이 사용하여 주어진 지도학습 알고리즘의 분류 정확도를 향상시키는 준지도학습(Semi-supervised learning) 알고리즘을 사용하여 보다 적은 비용으로 고효율의 이상탐지를 실현할 방법을 제시하고자 한다 [11].

레이블 전파(Label Propagation) 알고리즘은 labeled data와 unlabeled data를 모두 사용하여 알고리즘을 학습시키는 준지도학습 방법이다 [12]. Labeled data만 사용하여 학습을 진행하는 지도학습(Supervised learning)보다는 일반적으로 정확도가 떨어지지만, labeling 과정이 적기 때문에 한정된 자원에서 새로운 데이터를 취득하는 경우에 데이터셋 구성 비용을 절감할 수 있다는 강점을 가지고 있다.

본 연구에서는 반도체 식각 공정 중 MFC 이상상태에 대한 모사 데이터셋을 레이블 전파 알고리즘에 적용하여 일반적인 지도학습 기반 이상분류모델에 근사한 분류 정확도를 기대할 수 있는 저비용 학습 방법론을 제안한다.

2. 데이터 및 전처리 방법

2.1 MFC 의 성능 저하 모사 데이터

반도체 공정 과정에서는 공정 조건을 정확하게 조절하기 위해 다양한 부분품을 사용한다. 공정 과정에서 단일 부분품의 역할은 작지만, 미세화된 반도체 공정에서는 부분품의 작은 성능 변화도 공정결과에 영향을 주는 변수가 된다. 따라서 부분품에서 발생하는 변수를 정확하고 빠르게 감지할 수 있는 FDC 시스템이 필요하다.

반도체 부분품에서 발생하는 오류는 대부분 빠른 조치가 가능하지만, 부분품 자체의 노후화로 인해 성능이 감소된 경우에는 변수가 다양하기 때문에 단일 부분품 문제로 정의하기 어려우며, 특히 센서나 제어기가 노후화되어 잘못된 측정수치를 나타낼 경우에는 장비 운용자가 이를 확인하는데 한계가 있다. 따라서 이러한 문제는 부분품이 아닌 공정 과정에서 원인을 찾아야 한다. 반도체 식각 공정 과정에서 사용되는 플라즈마는 다양한 원인으로 공정 중 실시간으로 변화된다. 특히 chamber 내부에 공급되는 불소 계열의 가스의 유량이 변화될 경우 플라즈마의 특성이 변하는 모습을 OES로 확인할 수 있다.

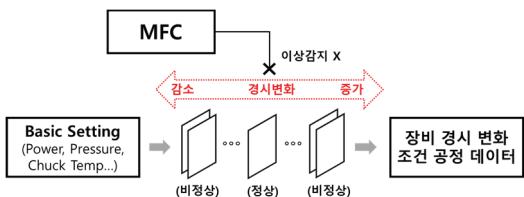


Fig. 1. Structure diagram of MFC performance degradation simulation experiment.

본 연구에서는 반도체 식각 공정 과정에서 유량제어기(MFC)가 노후화되어 SF₆ 가스가 설정한 수치와 다르게 공급되는 상황에서 식각 공정 중 발생하는 플라즈마의 상태를 대변할 광 스펙트럼 데이터 및 실시간 장비 조건 변화량 데이터를 취득하였다. Figure 1은 MFC 노후화 데이터를 얻기 위한 모사 실험 진행 구조도이다.

광 스펙트럼 데이터는 OES 센서를 사용하여 식각 공정 중 발생하는 실시간 플라즈마 화학종 데이터를 취득하였다. OES 센서는 플라즈마 내에 존재하는 이온과 라디칼로부터 발생하는 광원을 흡수 스펙트럼으로 변화시켜, 특정한 파장에서 발생하는 스펙트럼의 강도(intensity)를 모니터링 하여 해당 화학종의 변화를 관찰한다. 데이터는 공정에 사용한 가스 변화량을 나타낼 수 있는 파장대의 스펙트럼 강도 변화량으로 이루어져 있다.

본 연구에 사용된 데이터셋은 MFC 이상상태 모사 과정에서 발생한 파장강도의 실시간 변화량으로 이루어져

있다. 파장강도는 공정에 사용된 각 가스의 변화량을 반영해 줄 수 있는 파장대를 선별하였고, 최종적으로 총 8개의 파장대에 대한 강도 변화량을 취득하여 데이터셋을 구성하였다.

2.2 레이블 전파 학습(Label Propagation)

Label Propagation은 데이터 내에 레이블이 없는 노드에 대하여 레이블을 할당해주는 과정을 반복하는 학습 방법이다. 이 방법은 2002년 Xiaojin Zhu와 Zoubin Ghahramani가 처음 제안하였다 [13]. Label Propagation은 하나의 node가 속한 영역은 주변 노드들이 속한 영역일 확률이 높다는 직관적이지만 강력한 개념이 담겨있다. 이 알고리즘은 단순히 주변 노드들 간의 연관성으로 학습하기 때문에 빠른 속도의 검출이 가능하다.

Label Propagation은 아래의 순서에 따라 Fig. 2 의 형태로 학습을 진행한다.

1. 일부 학습데이터 라벨 제거
2. Unlabeled node 주변 labeled node들을 확인하여 가장 빈도가 높은 label을 unlabeled node로 전파
3. 모든 unlabeled node에 대해 2번 과정 진행

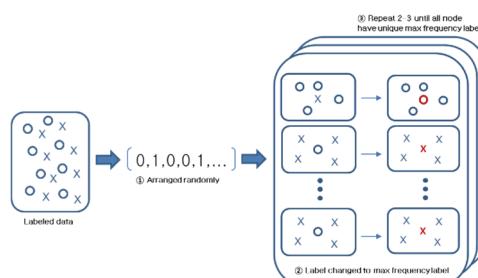


Fig. 2. Structure of Label Propagation algorithm.

알고리즘 조건인 3번은 안정적인 영역설정을 위한 방법이다. 하나의 노드가 매번 다른 레이블로 지정된다면 여러 영역의 중간에 포함되어 있는 형태일 가능성이 높다. 따라서 유일하게 빈도수가 가장 높은 라벨을 각 노드가 가질 수 있도록 학습을 반복한다.

3. 실험결과 및 분석

3.1 SF_6 가스 유량 변화 식각 실험

실험을 위해 13.56MHz RF 전력 유도 결합 플라즈마 반응성 이온 식각 (ICP-RIE) 과 함께 300mm 플라즈마 식각 시스템을 사용하였다. SF_6 용 MFC가 공정 시간에 따라서 노후화되는 상황을 모사하기 위해 기본 식각 레시피에서 SF_6 가스 유량이 2 sccm씩 변화하였을 때 플라즈마에서 나

타나는 파장 강도 변화를 OES 센서로 취득하였다. 본 공정에 사용된 이상 레시피(Best-known Method, BKM)은 Table 1 과 같다.

Table 1. Best-known Method (BKM) recipe

Experiment process recipe	
Source Power [W]	800
Bias Power [W]	50
Time [s]	180
Pressure [mTorr]	20
SF_6 Flow Rate [sccm]	174
O_2 Flow Rate [sccm]	124
Ar Flow Rate [sccm]	10

가스 유량을 변경하여 이상상태를 모사하는 실험을 통해 가스 유량의 변화가 플라즈마 및 공정 결과에 영향을 미치는 것을 확인하였다. Fig. 3은 SF_6 가스량이 BKM으로 진행하였을 때와 의도적으로 가스 유량을 변경하였을 때의 식각률을 보여주기 위해 공정 결과를 단면 주사 전자 현미경(Scanning Electron Microscope, SEM)으로 확인한 이미지이다. Fig. 3 (b)의 경우 의도적으로 SF_6 가스 유량을 16sccm 증가시킨 경우이며, BKM으로 진행한 (a)에 비해 식각 깊이가 감소한 모습을 보여준다. 이 현상은 주어진 압력에서 가스 분자의 수가 증가하여 플라즈마에서 충돌량이 변화하며 발생한 것으로 볼 수 있다.

또한 취득한 OES 센서 데이터의 경우 SF_6 가스 유량과 OES 파장 강도가 반비례하는 경향을 보여주었다. 플라즈마의 전자는 충돌 전자에 의해 생성되며, 에너지 적으로 여기 된 전자가 에너지를 잃었을 때 플라즈마 글로우 방전이 발생한다. 증가된 가스 흐름의 양이 플라즈마의 충돌 단면을 증가시켜 관련 파장대에서 OES 파장 강도를 증가시킬 수 있다고 추론하였다. OES 센서 데이터는 Fig. 4 와 같은 스펙트럼 형태로 취득되며, 공정에 사용된 Ar , SF_6 , O_2 가스의 영향을 가장 많이 받는 파장대를 Table 2와 같이 8가지 선정하여 각 파장대의 파장강도를 알고리즘의 입력 변수로 사용하였다.

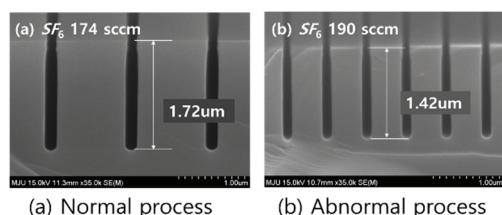


Fig. 3. Scanning Electron Microscope (SEM) image of normal and abnormal state.

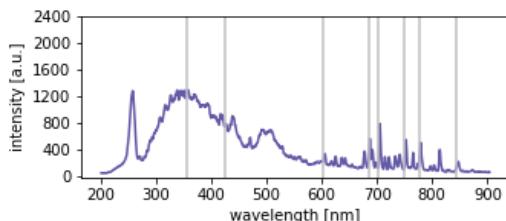


Fig. 4. OES full spectrum graph.

Table 2. Variables in the dataset.

Sensor data (Intensity)	
<i>F</i>	685 nm
	703 nm
<i>O</i>	775 nm
	844 nm
<i>Ar</i>	357 nm
	425 nm
	603 nm
	750 nm

3.2 데이터 구조 및 전처리

본 연구에서는 labeling이 이루어진 MFC 성능저하 모사 데이터셋에서 일부 실험데이터의 레이블을 제거하는 작업을 추가하여 unlabeled data가 포함된 데이터셋으로 학습을 진행하였다. 결과에 대한 신뢰도를 높이기 위해 레이블을 지우는 실험데이터는 무작위로 선정하였다. 학습된 알고리즘의 평가를 위해 실험데이터를 80%의 학습 데이터와 20%의 테스트 데이터로 분류하였다.

고차원 데이터를 시각적으로 표현하기 위해 데이터의 차원을 축소하는 t-Stochastic Neighbor Embedding (t-SNE)을 적용하였다. 실제 데이터는 OES 8개 파장대에 대한 intensity 값으로 8차원 형태로 이루어져 있다. 이를 보다 직관적으로 확인하기 위해 Fig. 5와 같이 3차원으로 차원을 축소하여 시각화 하였다.

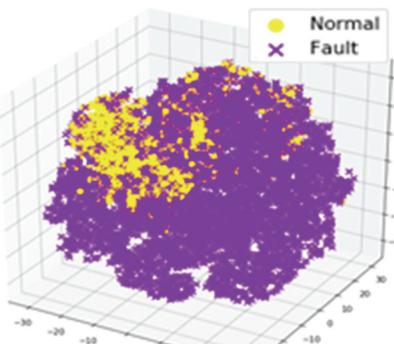


Fig. 5. Dataset visualization by t-SNE graph.

3.3 Label Propagation 학습 적용

Label Propagation학습을 진행하기 위해 모델을 구성하는 하이퍼 파라미터들을 선정하였다. 각 하이퍼 파라미터는 모델 구성에 직접적인 변화를 주며, 본 연구에서는 Table 3과 같이 설정하였다. 데이터셋이 가지는 분포적 특징에 따라 Gamma는 0.003로 설정하였다. 또한 서로 다른 kernel의 분류 정확도를 확인하기 위해 Radial Basis Function (RBF) kernel과 K-nearest neighbor (KNN) kernel을 각각 사용하여 두 개의 학습모델을 구성하였다. Kernel의 변화에 따른 알고리즘 분류 정확도의 차이는 Table 4 와 같다.

Table 3. Label Propagation hyperparameter list

Hyperparameter	
Kernel	Radial Basis Function (RBF)
	K-nearest neighbor (KNN)
Gamma	0.003
N. neighbors	15
Max-iteration	1000
Total	0.001
N. jobs	None

Table 4. Difference of model classification accuracy according to kernel

Label Propagation model accuracy					
RBF kernel			KNN kernel		
Total	Train	Test	Total	Train	Test
95.8%	96.7%	92.4%	93.2%	93.5%	91.9%

KNN kernel보다 RBF kernel을 적용한 학습모델의 분류 정확도가 더 높게 나타난 것으로 보아 현재 데이터셋에 RBF kernel이 더 효과적이었음을 확인하였다. 또한 전체 데이터로 예측한 경우와 훈련 데이터만으로 예측한 경우 모두 테스트 데이터만 적용한 경우보다 높은 분류 정확도를 보여주었다. 이는 모델이 학습데이터에 다소 과적합(Over-fitting)된 것으로 보인다. 하지만 전체적으로 90% 이상의 높은 분류 정확도를 보였다.

다음으로는 동일한 데이터셋에 대한 지도학습 알고리즘과의 분류 정확도를 비교해 보았다. 지도학습을 대표하여 사용된 알고리즘은 분류 모델 중에서도 일반적으로 많이 사용되는 Support Vector Machine (SVM)을 사용하였다. SVM에 적용한 결과 Fig. 6 같은 분류 정확도를 확인할 수 있었다. Figure 6은 모델 정확도를 평가할 때 사용하는 방법 중 하나인 ROC곡선(Receiver Operating Characteristic Curve)이다. ROC 곡선은 모든 분류 임계값에서 분류 모델의 성능을 보여준다. Figure 3에서 ROC 곡선 아래 영역(Area

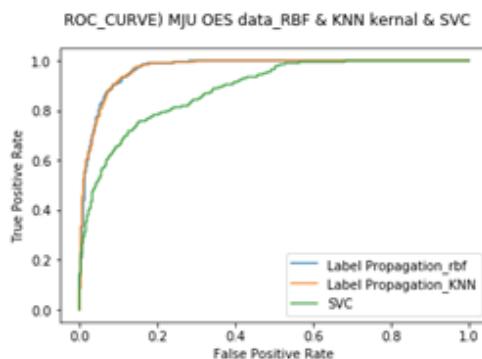


Fig. 6. ROC-curve with two kernels and SVC model.

Under the Curve, AUC)을 비교하였을 때 Label Propagation 모델이 SVM 모델보다 더 높은 AUC를 보여주었다. 이는 Label Propagation모델이 SVM모델 보다 우수한 성능을 나타낸다고 볼 수 있지만, 학습에 사용된 데이터셋의 분포도가 비선형적이기 때문에 SVM에 적합하지 않았던 것으로 추정된다. 추후 고차원의 데이터를 반영할 수 있는 kernel SVM 알고리즘과의 성능을 비교하고자 한다.

4. 결 론

본 논문에서는 플라즈마 식각 공정에서 OES 데이터를 활용하여 기계학습 알고리즘을 통해 MFC의 FDC 모델을 구현하였다. 알고리즘 학습에 Label Propagation 방법을 사용할 경우 labeled data가 부족한 상황에서 unlabeled data를 함께 사용할 경우 높은 분류 정확도를 보여주었다. 이를 통해 Labeling 비용에 대한 부담이 큰 경우에서도 효과적인 기계학습 기반 FDC를 구성할 수 있다.

감사의 글

이 논문은 산업통상자원부(20006499)와 KSRC 지원 사업인 미래반도체소자 원천기술개발사업의 연구결과로 수행되었으며, 2020년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구이다. (P0008458, 2020년 산업혁신인재성장 지원사업)

참고문헌

- P. S. Peercy, "The drive to miniaturization," *Nature*., Vol. 406, No. 6799, pp. 1023-1026, 2000.
- Q. P. He and J. Wang, "Fault Detection Using the k-Nearest Neighbor Rule for Semiconductor Manufacturing Processes," *IEEE Trans. Semi. Manufac.*, Vol. 20, No. 4, pp. 345-354, 2007.
- Kim, Dongil, et al. "Machine learning-based novelty detection for faulty wafer detection in semiconductor manufacturing," *Expert Systems with Applications*., Vol. 39, No. 4, pp. 4075-4083, 2012.
- D. H. Kim, J. E. Choi, T. M. Ha and S. J. Hong, "Modeling with Thin Film Thickness using Machine Learning," *Journal of the Semiconductor & Display Technology*, Vol. 18, No. 2, pp. 48-52, 2019.
- S. J. Hong, G. S. May and D. C. Park, "Neural network modeling of reactive ion etching using optical emission spectroscopy data," *IEEE Transactions on Semiconductor Manufacturing*, Vol. 16, No. 4, pp. 598-608, 2003.
- K. Nakata, R. Orihara, Y. Mizuoka and K. Takagi, "A comprehensive big-data-based monitoring system for yield enhancement in semiconductor manufacturing," *IEEE Transactions on Semiconductor Manufacturing*, Vol. 30, No. 4, pp. 339-344, 2017.
- J. H. Han and S. S. Hong, "Semiconductor Process Inspection Using Mask R-CNN," *Journal of the Semiconductor & Display Technology*, Vol. 19, No. 3, pp. 12-18, 2020.
- S. J. Hong, W. Y. Lim, T. Cheong and G. S. May, "Fault Detection and Classification in Plasma Etch Equipment for Semiconductor Manufacturing \$e\$-Diagnostics," *IEEE Trans. Semi. Manufac.*, Vol. 25, No. 1, pp. 83-93, 2012.
- K. B. Lee, S. Cheon and C. O. Kim, "A convolutional neural network for fault classification and diagnosis in semiconductor manufacturing processes," *IEEE Transactions on Semiconductor Manufacturing*, Vol. 30, No. 2, pp. 135-142, 2017.
- Chou, Paul B., et al. "Automatic defect classification for semiconductor manufacturing," *Machine Vision and Applications*., Vol. 9, No. 4, pp. 201-214, 1997.
- P. K. Mallapragada, R. Jin, A. K. Jain and Y. Liu, "SemiBoost: Boosting for Semi-Supervised Learning," *IEEE Trans. Pattern Analysis and Machine Intelligence*., Vol. 31, No. 11, pp. 2000-2014, 2009.
- T. Sarmiento, S. J. Hong and G. S. May, "Fault detection in reactive ion etching systems using one-class support vector machines," *IEEE/SEMI Conference and Workshop on Advanced Semi. Manufac.*, pp. 139-142, 2005.
- Xiaojin, Zhu, and Ghahramani Zoubin, "Learning from labeled and unlabeled data with label propagation," *Synthesis lectures on artificial intelligence and machine learning*., Vol. 3, No. 1, pp. 1-130, 2009.