

XGBoost를 사용한 반도체 노광 공정 계측 결과 예측

신정일*, 박지수**, 손진곤*

*한국방송통신대학교 대학원 정보과학과

**전주대학교 컴퓨터공학과

cutekk@knou.ac.kr, jisupark@jj.ac.kr, jgshon@knou.ac.kr

Prediction of Semiconductor Exposure Process Measurement Results using XGBoost

Jeong Il Shin*, Ji Su Park**, Jin Gon Shon*

*Dept. of Computer Science Graduate School, Korea National Open University

**Dept. of Computer Science and Engineering, Jeonju University

요약

반도체 회로의 미세화로 단위 공정이 증가하면 TAT(turn-around time) 증가에 따른 제조 비용이 늘어난다. 반도체 공정 중 포토 공정은 마스크의 회로를 웨이퍼에 전사하는 공정으로 전사를 담당하는 노광장비의 성능에 의해 회로의 정확성이 결정된다. 이런 정확성을 검증하는 계측공정은 회로의 미세화가 진행될수록 필요성은 증가하나 TAT 증가의 주된 요인으로 최근 기계학습을 사용한 다양한 예측 모형들의 개발로 계측 결과를 예측하는 실험들이 진행되고 있다. 본 논문은 노광장비 센서들의 이상값을 감지하여 분류 후 계측공정을 진행하는 LFDC(Lithography Fault Detection and Classification) 시스템의 문제인 분류 성능이 떨어지는 것을 해결하기 위해 XGBoost를 사용하여 계측공정을 진행하지 않고 노광장비 센서의 이상값을 학습된 학습기를 통해 분류하여 포토 공정을 재진행하거나 다음 공정을 진행하는 방법을 실험하였다. 실험에서 사용된 계측 결과 예측 모형은 89%의 정확도를 확보하였고 반도체 데이터 특성인 심각한 불균형의 데이터에 대해서도 같은 정확도를 얻었다. 이런 결과는 노광장비 센서들의 이상값에 대해 89%는 정상으로 판단하였고 정상으로 판단한 웨이퍼를 실제 계측 시 예측과 같은 결과를 얻었다. 계측 결과 예측 모형을 사용하면 실제 계측을 진행하지 않고 노광장비 센서들의 이상값에 대한 판정을 할 수 있어 TAT 단축으로 제조 비용감소, 계측 장비 부하 감소 및 효율 향상을 할 수 있다. 하지만 본 논문에서는 90%의 성능을 보이는 계측 결과 예측 모형으로 여전히 10%에 대해서는 실제 계측이 필요한 문제에 대해 추후 더 연구가 필요하다.

1. 서론

최근 반도체는 미세화가 진행되면서 많은 단위 공정들이 추가되고 단위 공정 진행을 위한 천문학적인 비용이 소요된다. 이러한 단위 공정에는 전 공정과 후공정으로 나누어지고 전 공정은 산화, 포토, 식각, 증착, 배선 공정 다섯 가지로 크게 나누고 이 다섯 가지 공정이 수백 단계로 진행하면 하나의 웨이퍼에 수천 개의 칩이 만들어진다. 반도체 제조업은 장치 산업으로 불리고 제품의 개발부터 양산까지 걸리는 시간(TAT, turn-around time)이 제조원가에 가장 큰 비중을 차지하고 있어 이러한 단위 공정을 줄이는 연구개발이 활발하게 이루어지고 있다. 특히 웨이퍼에 회로를 인쇄하는 포토 공정에 사용되는 노광장비는 빛을 사용하여 마스크의 회로를 웨이퍼에 전사하는 노광 공정에 사용이 되고 노광장비의 성능이 반도체 칩의 성능을 결정한다. 노광 공정 진행 시

노광장비의 미세한 변화에 의한 선폭의 변화를 확인하기 위해 노광 공정 후에 계측공정으로 웨이퍼에 전사된 패턴의 정확도를 검증하고 계측 장비의 계측 결과에 따라 웨이퍼를 재진행하거나 다음 공정으로 넘긴다. 이에 따른 재진행 비용과 계측공정 진행으로 인한 TAT 증가가 반도체 제조 비용을 증가시키는 요인으로 작용한다. 현재 노광 공정 진행 시 노광장비의 이상 유무를 감지하기 위한 시스템으로 LFDC(Lithography Fault Detection and Classification)로 노광장비의 센서값들을 모니터링 및 이상값을 감지하여 분류 후 계측공정을 진행하고 있다. 분류된 모든 웨이퍼는 계측공정을 진행 후 계측 결과에 따라 포토 공정을 재진행하거나 다음 공정으로 넘어간다. 하지만 LFDC의 분류 성능은 15% 이하로 모든 이상값의 웨이퍼 중 15%만 실제 불량으로 재진행이 이루어지고 나머지 85%는 정상 웨이퍼로 분류되어 다음 공정으로 넘어가기 때문에 심각한 계측 장비의

효율 저하와 LFDC가 분류한 이상값의 웨이퍼 모두를 계측하여 계측 장비 부하로 이어지고 계측 장비에 웨이퍼들이 대기하기 때문에 TAT가 증가하는 문제가 있다. 본 논문에서는 실제 계측을 진행하지 않고 기계학습을 사용하여 LFDC의 이상값을 학습 후 계측 결과를 예측하는 방법을 제시한다. 본 논문의 실험에 사용된 데이터는 LFDC에서 추출한 센서값과 계측 장비에서 실제 계측한 결과로 XGBoost 알고리즘을 사용하여 학습하고 테스트 데이터로 성능을 검증하는 방법과 교차검증 알고리즘을 사용하는 방법으로 학습기의 성능을 검증한다.

2. 관련 연구

2.1 반도체 포토 공정

포토 공정은 Resist를 웨이퍼 표면에 입히는 도포 공정, 디자인된 회로 패턴을 마스크에 형성 후 빛을 이용하여 해당 마스크의 패턴을 프로젝션 렌즈를 통해서 웨이퍼에 전사하는 노광 공정, 노광 된 웨이퍼를 현상 용액을 사용하여 현상하는 현상 공정, 현상된 패턴이 이전 진행한 단위 공정과 정확하게 중첩이 되었는지 확인하는 오버레이 계측공정, 회로 패턴이 정확한 크기로 노광이 되었는지 확인하는 CD(Critical Dimension) 계측공정, 웨이퍼 표면에 이물질이나 다른 요인으로 패턴의 불량이 발생했는지에 대해 계측하는 정밀검사 공정의 단위 공정으로 이루어진다. 반도체는 여러 층으로 이루어지고 노광과 식각 공정을 지나면서 순차적으로 다양한 층들을 쌓아서 설계된 패턴의 회로가 웨이퍼에 그려진다.

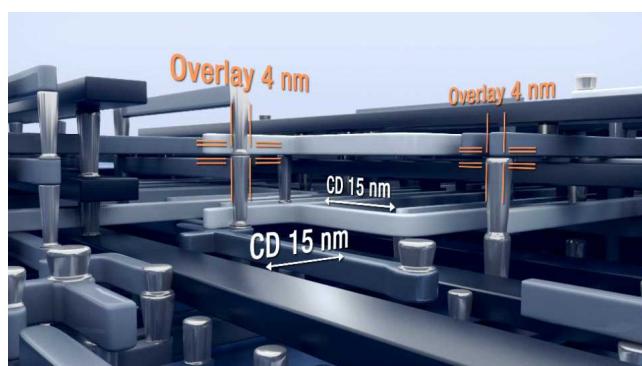


그림 1 반도체 구조 (www.asml.com)

<그림 1>에서 보여주는 Overlay 4nm의 의미는 층과 층 사이의 정렬이 4nm의 차이가 있다는 의미로 오버레이 계측공정은 층과 층 사이의 정렬을 계측하여 수치로 나타내고 하부층과 상부층이 정확한 위치에

정렬되었는지 검증하는 공정이다.

2.2 XGBoost 알고리즘

XGBoost(eXtreme Gradient Boost)는[1] 트리 기반의 앙상블 학습 알고리즘으로 분류와 회귀 영역에서 뛰어난 예측 성능을 가지고 있다. XGBoost는 분산 컴퓨팅 기법을 활용한 병렬 수행으로 그레디언트 부스트보다 빠른 속도를 보여주고 있고 자체 과적합 규제 기능이 있다. 트리 가지치기 기능으로 이득이 없는 분할을 가지치기해서 분할 수를 더 줄이고 반복 수행마다 내부적으로 교차검증을 수행하여 최적화된 반복 수행 횟수를 가진다. XGBoost는 그레디언트 부스트, 정규화, A unique regression tree 학습 알고리즘을 개선하였고, 근사 탐욕 알고리즘, 병렬 학습, Weighted Quantile Sketch, Sparsity-Aware Split Finding, Cache-Aware Access, Blocks for Out-of-Core Computation과 같은 최적화 기술을 도입하였다. XGBoost는 유사성 점수라는 인덱스로 트리를 만든다. <수식 1>은 분류를 위한 유사성 점수 인덱스 트리 구성 수식이다.

$$\text{유사성 점수} = \left(\sum_{i=1}^n \text{Residual}_i \right)^2 / \left(\lambda + \sum_{i=1}^n p_i(1-p_i) \right) \text{ 수식 1}$$

*Residual*은 잔차이고, λ 는 정규화 파라미터이고 $p_i(1-p_i)$ 는 가중치를 나타낸다. 즉, 모든 가중치의 합과 정규화 매개변수의 합에서 모든 잔차의 합을 나누면 유사성 점수가 된다.

3. XGBoost 기반 계측 결과 예측 실험 방법

본 논문의 계측 결과를 예측하는 실험 방법의 전체적인 과정은 <그림 2>와 같다.



그림 2 계측 결과 예측 실험 단계

계측 결과 예측 실험은 입출력 정의, 정의된 입출력의 데이터를 추출 및 전처리를 진행하고 학습에 사용하는 알고리즘을 선택 후 학습, 검증, 테스트의 과정으로 진행한다. 입력 데이터는 단 변량 변수 분류 시스템에서 추출 후 Min-Max Scale 정규화를 진행하여 0에서 1 사이로 변환된 값을 사용한다. 학습

과 테스트 데이터의 비율은 70%와 30%이며, 레이블은 계측 장비의 계측 결과에서 추출하여, 상한 5%, 하한 5%를 임계치로 정하고 레이블을 ‘1’로 나머지 90%는 ‘0’으로 설정한다. 학습은 XGBoost 알고리즘에서 scikit-learn의 XGBClassifier 함수로 이진 분류하고, 평가는 accuracy_score 함수로 30%의 테스트 데이터를 평가하는 방법과 교차검증으로 평가하는 방법을 사용한다. 최종적으로 데이터 불균형에 따른 정확도 비교를 위해 <수식 2>의 정확도로 확인 후 최고점이 나올 때까지 하이퍼파라미터를 변경하는 방법을 사용한다.

$$Accuracy(\text{정확도}) = \frac{TP + TN}{TP + FN + FP + TN} \quad \text{수식 2}$$

<그림 3>은 본 논문에 사용한 계측 결과 예측에 사용된 학습 및 평가의 전체적인 과정을 보여준다.

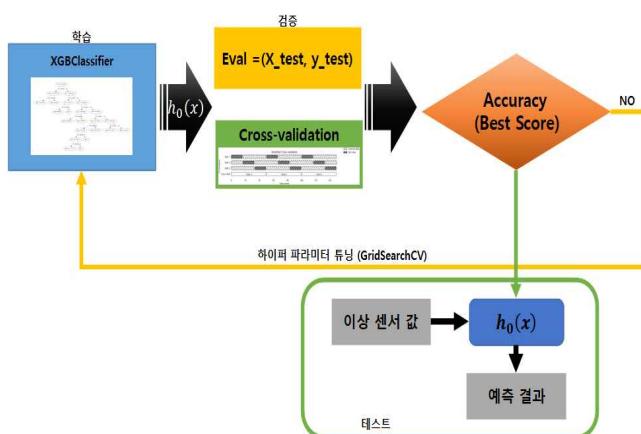


그림 3 계측 결과 예측 모형

4. 실험 결과

실험은 AMD Ryzen 7 4800H, 16G RAM 및 64 Bit Window 운영체제에서 실험이 이루어졌고 파이썬을 사용하여 진행하였다. 입력은 Reticle Alignment에서 나온 센서값 중 Zernike 상수 5에서 37까지의 offset, tilt, curvature, 3rd order distortion 총 131개를 사용하였고 출력은 오버레이 계측 결과에서 average x/y, 3sigma x/y, residual x/y를 상한 5%, 하한 5%를 불량으로 레이블 설정하였고 90%의 데이터는 정상 레이블 설정으로 데이터의 불균형을 크게 하였다. 실험의 하이퍼파라미터 변경은 GridSearchCV 알고리즘을 사용하여 최고의 정확성을 얻는 최적의 하이퍼파라미터를 얻었다. 교차검증은 StratifiedKfold 알고리즘으로 데이터 불균형에 따른 문제를 해결하였고

데이터는 10개의 단위로 나누어서 학습과 테스트에 사용하였다. 데이터 개수 별 정확도 비교 실험에서는 <표 1>의 결과를 얻었다.

표 1 데이터 개수 별 정확도

데이터 개수	Accuracy [%]					
	avrx	avry	3sigx	3sigy	rsdx	rsdy
300개	89.0	89.4	89.0	88.0	88.3	89.7
1,000개	89.0	89.5	89.7	89.5	90.6	89.6

300개와 1,000개를 사용한 데이터에서 1,000개를 사용하여 학습한 결과가 좀 더 높은 정확성을 보여주었다. 두 번째는 검증 방법에 따른 정확도를 실험하였고 <표 2>의 결과를 얻었다.

표 2 데이터 분류별 정확도

검증 방법	Accuracy [%]					
	avrx	avry	3sigx	3sigy	rsdx	rsdy
CV	89.0	89.5	89.7	89.5	90.6	89.6
7:3	98.1	96.6	89.6	86.7	97.0	95.9

교차검증을 사용한 실험에서는 모든 출력에서 89%의 고른 정확도를 나타내었지만 7대 3의 비율로 나누어 학습시킨 실험에서는 98%에서 86%까지 편차가 심한 결과를 얻었다. 이는 90대 10의 극심한 불균형 데이터에 따른 결과로 보이고 교차검증을 사용하면 데이터 불균형 문제는 해결되었다.

5. 결론

본 논문에서는 최근 반도체 제조 공정에 도입되고 있는 다양한 기계학습 알고리즘 중에 이진 분류에 우수성을 보이는 XGBoost 알고리즘을 사용하여 노광장비의 센서값과 계측 장비의 계측 결과를 학습하여 계측 결과를 예측 및 분류하는 실험을 하였다. 기존의 단 변량 변수 분류 시스템은 이상값을 띠운 모든 웨이퍼를 계측하고 계측 결과 정확성은 15% 이내로 낮아 계측 장비의 부하는 높아지고 효율은 떨어졌다. 본 실험에서는 기계학습을 사용하여 학습한 모델을 통해 계측 없이 계측 결과를 예측 및 분류하여 89%의 정확성을 얻었다. 이는 이상 값 중

89%의 웨이퍼가 정상으로 분류되어 계측 없이 다음 공정을 진행할 수 있어 계측 장비 부하는 줄이고 정상 웨이퍼를 측정하는 빈도도 줄어 계측 장비의 효율도 높였다. 하지만 본 논문은 90% 이상의 예측 및 분류 정확성은 얻을 수 없었고 데이터 수에 따른 예측 및 분류 정확성에 차이는 없으나 심각한 데이터 불균형에 따른 예측 및 분류의 어려움이 있어 향후 불균형 데이터 해소를 위한 데이터의 추가 및 데이터 증강 알고리즘 적용을 통한 예측 및 분류 성능을 높이는 추가 연구가 필요하다.

본 연구는 과학기술정보통신부 및 정보통신기획 평가원의 대학ICT연구센터 육성지원사업의 연구 결과로 수행되었음. (IITP-2021-2020-0-01789)

참고문헌

- [1] Tianqi Chen, Carlos Guestrin, “XGBoost: A Scalable Tree Boosting System”, KDD’16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785-794, Aug. 2016
- [2] Rusheng Hu, Xiaolong Li, and Yunxin Zhao, “Gradient Boosting Learning of Hidden Markov Models” IEEE, 1-4244-0469-X, 2006
- [3] Jerome H. Friedman, “Greedy function approximation: A gradient boosting machine”, Ann. Statist. 29 (5), 1189-1232, Oct. 2001
- [4] Yoav Freund, Robert E. Schapire, “A Short Introduction to Boosting”, Journal of Japanese Society for Artificial Intelligence, 14(5):771-780, Sep. 1999
- [5] Alexey Natekin, Alois Knoll, “Gradient boosting machines, a tutorial”, METHODS ARTICLE, Dec. 2013
- [6] Andreas C. Muller & Sarah Guido, “Introduction to Machine Learning with Python”, Sebastopol CA, O’Reilly Media, Oct. 2016