

퍼지 논리와 신경망에 기반한 공정 예측 및 품질 추정을 위한 공정관리 의사지원시스템

Decision Support System for Prediction and Estimation of Qualities Based on Neural Networks and Fuzzy Logic

배현, 우영광, 김성신, 우광방*
부산대학교 전자전기통신공학부
*연세대학교 자동화기술연구소

Hyeon Bae, Young-Kwang Woo, Sungshin Kim, Kwang-Bang Woo*
School of Electrical and Computer Engineering, Pusan National University
*Automation Technology Research Institute, Yonsei University
E-mail: baehyeon@pusan.ac.kr

요 약

차세대 생산 시스템(Next Generation Manufacturing System: NGMS)의 핵심 개념은 분산 생산 시스템과 다품종 소량의 유연 생산 시스템의 지원이다. 이러한 시스템의 구성을 위하여 실시간 데이터에 기반한 예측 모델이 필수적인데, 이러한 예측 기능을 통하여 생산공정의 관리와 운영, 특히 전체 공정관리를 효율적으로 수행할 수 있다. 한편, 공정으로부터 전송된 데이터는 특정한 형태의 지식으로 표현된다. 이러한 지식들은 시스템에 대한 다양한 정보를 가지고 있으므로 정보를 이용하여 시스템 상태를 빠르고 쉽게 진단할 수 있다. 공정 진단은 현재 공정 상태에서 생산되는 제품의 품질을 추정할 수 있는 정보로 활용된다. 본 논문에서는 이러한 개념이 바탕이 되어 공정관리 시스템을 설계하였다. 제안된 시스템의 적용 대상은 반도체 제조 공정의 단위 공정인 에칭 공정이다. 에칭 공정은 공정 중에 연속적인 검사가 수행되지 않고 최종 제품에 대한 검사가 수행되므로 불량 원인을 찾는 것이 쉽지 않다. 따라서 본 논문에서는 공정관리를 위한 의사지원시스템을 통해 공정의 연속적인 간접진단을 수행하고자 하였다. 본 연구에서 사용된 의사지원시스템은 각 공정에서 얻어지는 데이터와 경험적 지식을 토대로 공정시스템의 해석과 진단이 가능한 시스템이다.

Keywords: 의사지원시스템, 퍼지 논리, 신경망, 데이터 마이닝, 공정 예측, 품질 추정

1. 서론

인공지능은 특정 분야에서 전문가의 축적된 지식과 경험을 시스템화하여 필요할 때 사용하도록 하는 알고리즘이다. 여기서 지식이라는 용어는 정보와 혼용하여 사용되기도 하며 구분되기도 하는데, 데이터 관점에서 엄밀하게 정의하면 정보란 데이터를 가공, 처리해서 얻어지는 산출물이고, 지식이란 정보의 체계적인 활용을 통해 축적된 노하우이다[1]. 정보는 그 자체로서도 경쟁력의 원천이 되지만, 지식으로 발전하게 되면 경쟁력은 더욱 강화된다. 정보화시대의 기업경영에서 정보란, 전통적인 자원들을 효과적으로 운영, 관

리하며 새로운 제품이나 서비스를 창출하는 역할을 담당하는 또 다른 자원이다. 이렇게 기업간의 경쟁이 심화되고 정보의 중요성에 대한 인식이 확산됨에 따라 대량의 데이터에서 유용한 정보를 추출하는 것이 중요하게 인식되고 있다. 여기서 등장하는 것이 데이터 마이닝(data mining)의 개념이다[2]. 데이터 마이닝이란 자동화되고 지능을 갖춘 데이터베이스 분석기법으로 대량의 데이터로부터 의미 있는 정보를 추출하여 의사결정에 활용하는 작업인 것이다. 이러한 정보기반 시스템은 보다 효율적인 생산계획과 공정장비관리를 가능하게 하고, 생산 또는 운전 조건의 변동에도 유연하게 대처할 수 있다. 최종적으로 생산공정

과 운영상태를 신속하게 분석하여 현재 상황에 대한 의사결정을 용이하게 해주는 것이 목표이다.

2. 에칭공정을 위한 의사지원시스템

2.1 시스템 개요

본 연구에서는 최종 결과물인 공정관리를 위한 의사지원시스템 개발을 위하여 데이터 마이닝 기법이 사용된다. 데이터 마이닝에 의해 데이터로부터 지식이 도출되어 각 공정에서 얻어지는 데이터나 경험적 지식을 토대로 공정시스템의 해석과 진단이 가능한 시스템이다. 오늘날의 생산 환경에서는 공정설계, 조립, 제어, 스케줄링, 유지·보수 등과 같은 연관된 모든 영역으로부터 수많은 양의 데이터를 획득하여 데이터 베이스에 저장한다. 본 연구에서는 운전자의 경험을 접목하여 적절한 지식 추출이 가능한 통합형 데이터 마이닝 기법이 적용된다. 그림 1은 의사지원시스템의 공정 적용에 대한 개념도이다.

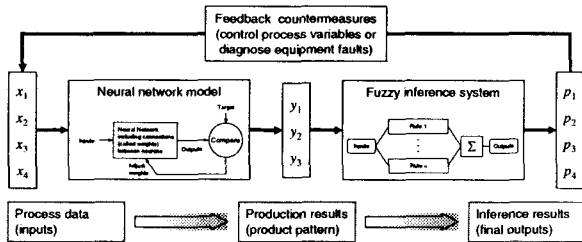


그림 1. 의사지원시스템의 개략적인 개념.

2.2 에칭공정

반도체 공정 중 에칭 공정은 중요한 단위 공정이다. 에칭 공정은 슬라이스의 두께를 줄이거나, 표면 결함 혹은 표면 응력을 제거하기 위하여 슬라이스를 화학용액에 담가 화학용액과 반응케 하는 공정을 말한다. 에칭은 크게 건식 에칭과 습식 에칭으로 나뉜다. 본 연구에서는 건식 에칭을 대상으로 하였다. 그 이유는 건식 에칭이 submicron 단위의 초미세 회로선폭 구현에 있어서 습식 에칭(wet-etching)으로는 도달할 수 없는 높은 분해능과 평탄화 기술의 필수 불가결한 에칭단면의 이방적 특성을 제공하므로 보다 널리 사용된다.

3. 시뮬레이션 결과 및 평가

본 연구에서 적용한 의사지원시스템은 그림 2에서 보는 바와 같이 5단계로 이루어진다. 입력 단계에서 추출한 제품의 측정 데이터를 이용하여 공정과 제품의 관계를 예측할 수 있는 모델을 구성하였다. 일반적으로 반도체 공정에서 생산되는

제품의 경우 전수검사가 아닌 샘플링 검사에 의존하는 경우가 많은데, 측정장비가 고가이고 전수검사를 위한 시간이 많이 소비되기 때문이다. 따라서 이력데이터를 통해 현재 공정 데이터에 기반한 품질추정이 가능한 모델은 현재 공정 조건에서의 향후 제품의 품질을 예측한다. 예측된 값이나 실제 측정 값을 바탕으로 추론시스템에서는 최종적인 제품의 상태를 추정한다.

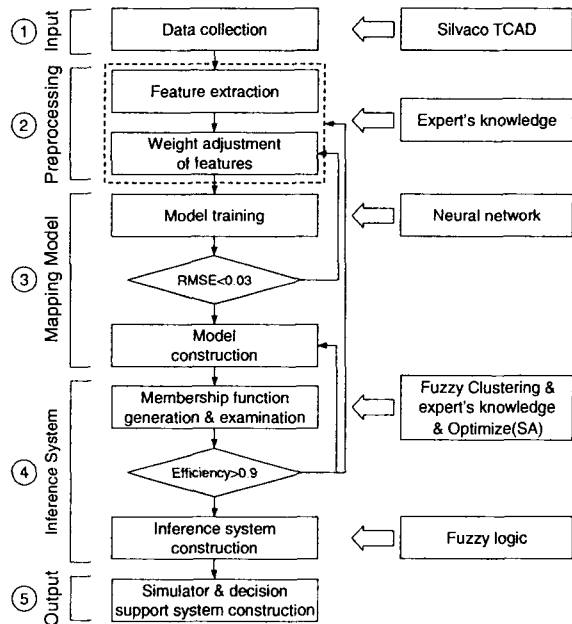


그림 2. 전체 알고리즘 흐름도.

3.1 Step 1: 입력단계(input)

본 연구의 최종적인 목표인 Digital Factory 시뮬레이터와 공정 관리 시스템의 구성을 위해 데이터 모델과 추론시스템을 적용하였다.

TCAD에서 반도체 관련 툴인 ATHENA에서 건식 에칭 중 Trench 에칭에 대한 시뮬레이터를 사용하였다. 그림 3에서 보는 바와 같이 제일 윗층(Oir32)의 모양이 에칭 공정의 결과로 생기는 결과물로 제어 대상이다. 이 공정의 입력 변수는 에칭 시간(bake time), 에칭 온도(bake temp.), 분사구 구경(circle sigma), 분사구 각도(projection na.)로 네 가지이다. 이러한 입력변수들에 따른 36개의 데이터 세트를 구성하고 Oir32의 에칭된 모양을 결과로 나타내었다.

3.2 Step 2: 전처리단계(preprocessing)

입력변수들에 따른 결과를 공정 모델과 의사결정 지원 시스템에 직접 적용하기 어려우므로 전문가의 지식을 사용한다. 각 변수의 범위를 정하기 위하여 다구찌 방법을 사용하는데[3], 이는

모든 범위에 대한 경우를 실험하는 것이 경제적으로 고가이기 때문이다. 표 1은 선택된 4가지의 입력변수와 각 범위값들을 나타내고 있다.

그림 3에서 보는 바와 같이 에칭 공정의 특성을 나타내는 9개의 특징점을 추출하여, 특징점 간의 관계로 공정의 특성을 나타내는 세 가지 공정변수를 정의하였다. 공정변수는 에칭 제품의 품질을 나타낸다. 입력변수에 대한 공정변수를 추출하고, 추출된 공정 변수에 따라 결과 데이터를 에칭공정이 얼마나 잘 되었는지를 전문가의 지식에 기반하여 4개의 패턴으로 나누었다. 목표 패턴은 3으로 두었고 그 특징에 따라 각각 1, 2, 4의 번호를 부여하였다. 이러한 입력변수, 공정변수, 그리고 최종적인 패턴으로 36개의 학습데이터를 구성하였다.

표 2. 입력변수의 영역별 대표값.

| 입력변수 | 변수명 | 범위 | 범위값 |
|-------|--------|-----|-----------------|
| x_1 | 에칭 시간 | 2단계 | 30, 60 |
| x_2 | 에칭 온도 | 2단계 | 125, 185 |
| x_3 | 분사구 구경 | 3단계 | 0.3, 0.5, 1.0 |
| x_4 | 분사구 각도 | 3단계 | 0.28, 0.4, 0.52 |

표 3. 공정의 특성을 나타내는 변수(특징점).

| 공정변수 | 변수 내용 |
|-------|----------------------|
| y_1 | 초기 식각된 부분의 깊이(④의 깊이) |
| y_2 | 중앙의 높이(④와 ⑤의 차이) |
| y_3 | 좌반부의 기울기(②와 ③의 기울기) |

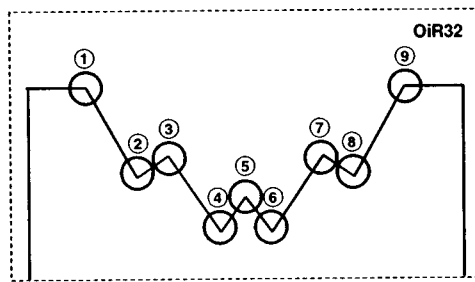


그림 3. 에칭결과로부터 특징점 추출.

3.3 Step 3: 모델 생성(model construction)

본 연구에서는 공정에 대한 모델링 방법으로 신경망을 사용하였다. 공정 변수를 이용하여 36개 입력 데이터를 구성하였다. 학습시 은닉층은 4개, 학습률은 0.01, epoch은 50,000번으로 설정하였다. 구성한 모델을 테스트하기 위해 입력변수로 공정변수를 추출한 방법과 마찬가지로 각 입력변수 사이의 값으로 테스트 데이터 세트를 얻는다. 표 3은 학습 후의 제품특성 예측 결과를 보여주는데 적절한 성능이 보장되고 있다.

표 4. 모델을 이용한 품질 예측 결과.

| 실험 | 실제값 | | | 예측 결과 | | |
|------|-------|--------|--------|--------|--------|---------|
| | y_1 | y_2 | y_3 | y_1 | y_2 | y_3 |
| 1 | 0.359 | 0.343 | -0.069 | 0.330 | 0.330 | -0.065 |
| 2 | 0.005 | 0.210 | 0.0223 | 0.024 | 0.209 | 0.045 |
| 3 | 0.597 | 0.2852 | 0.015 | 0.598 | 0.274 | -0.0009 |
| 4 | 0.330 | 0.292 | 0.065 | 0.343 | 0.297 | 0.046 |
| 5 | 0.363 | 0.343 | -0.055 | 0.351 | 0.321 | -0.084 |
| 6 | 0.013 | 0.179 | 0.038 | 0.018 | 0.175 | 0.013 |
| 7 | 0.603 | 0.273 | 0.007 | 0.599 | 0.267 | 0.007 |
| 8 | 0.344 | 0.257 | 0.061 | 0.334 | 0.272 | 0.036 |
| 9 | 0.398 | 0.281 | -0.076 | 0.398 | 0.283 | -0.053 |
| 10 | 0.088 | 0.141 | -0.005 | 0.091 | 0.139 | 0.052 |
| 11 | 0.625 | 0.238 | 0.004 | 0.614 | 0.235 | -0.026 |
| 12 | 0.380 | 0.202 | 0.037 | 0.463 | 0.225 | 0.04 |
| 13 | 0.406 | 0.260 | -0.093 | 0.429 | 0.284 | -0.109 |
| 14 | 0.121 | 0.131 | 0.014 | 0.134 | 0.130 | -0.022 |
| 15 | 0.63 | 0.229 | -0.019 | 0.659 | 0.221 | -0.036 |
| 16 | 0.388 | 0.188 | 0.038 | 0.386 | 0.172 | 0.042 |
| RMSE | | | | 0.0252 | 0.0125 | 0.0250 |

3.4 Step 4: 추론시스템(inference system)

본 단계에서는 표 3의 결과들을 기반으로 퍼지 추론 시스템을 설계한다. 먼저 앞서 구한 공정변수들이 원하는 결과물인지 아닌지를 판단할 수 있는 전문가의 지식이 필요하다. 앞서 이미 세 공정변수의 값에 따라 공정의 상태와 에칭의 정도를 총 4가지의 결과 패턴으로 분류하였다. 36개의 학습데이터를 이용하여 퍼지 추론 시스템을 구성한 후, 16개의 테스트 데이터로 퍼지 추론 시스템의 성능을 평가하였다.

표 5. 의사결정지원시스템의 입력과 출력변수.

| 공정변수 | 출력변수 |
|-------|--|
| y_1 | 1(p_1), 2(p_2), 3(p_3), 4(p_4) (식각 상태에 따른 패턴의 분류) |
| y_2 | |
| y_3 | |

3.4.1 퍼지 규칙과 소속함수

퍼지 규칙을 정하기 위해 퍼지 클러스터링을 사용하였다. 구해진 각 클러스터로 소속함수의 범위를 정의하였다. 클러스터링에서 세 번째 공정변수가 6개의 클러스터로 나뉘었지만, 실제 실험 결과에서는 왼쪽의 5개 클러스터들이 같은 영역의 값으로 묶이므로 최종적으로 2개의 클러스터로 나누었다. 이는 왼쪽에서 네 번째 까지의 클러스터들에 포함된 데이터의 수가 적고, 다섯 번째의 클러스터가 왼쪽 네 개의 클러스터의 특징까지도 포함하기 때문이다. 결국 각 공정변수들이 2개씩의 클러스터들로 나뉘므로 전체 규칙으로 조합으로 8개의 규칙으로 구성된다. 그리고 8개의 규칙은 효율성을 위해 4개로 줄일 수 있는데, 활성화되는 규칙만 사용하였기 때문이다.

IF y1 is S and y2 is S and y3 is S Then p=S
 IF y1 is S and y2 is S and y3 is L Then p=M1
 IF y1 is S and y2 is L and y3 is S Then p=S
 IF y1 is S and y2 is L and y3 is L Then p=M2
 IF y1 is L and y2 is S and y3 is S Then p=L
 IF y1 is L and y2 is S and y3 is L Then p=M1
 IF y1 is L and y2 is L and y3 is S Then p=L
 IF y1 is L and y2 is L and y3 is L Then p=M2

IF y1 is S and y3 is S Then p=S (p1)
 IF y2 is S and y3 is L Then p=M1 (p2)
 IF y2 is L and y3 is L Then p=M2 (p3)
 IF y1 is L and y3 is L Then p=L (p4)

3.4.2 최종 추론 결과

구성한 모델과 퍼지 추론시스템으로 미리 정한 결과 패턴 네가지에 대한 분류를 수행하였다. 먼저 모델링에 사용되었던 36개의 데이터 세트로서 성능을 평가해서 90%의 수율이 나올 경우를 선택해서 SA로 최적화를 수행하고, 다시 마찬가지로 모델을 테스트한 16개 데이터 세트로서 퍼지 추론 시스템을 평가하였다.

표 8. 16개의 데이터에 대한 최종적인 추론 결과.

| 실험 | y1 | y2 | y3 | 목표 패턴 | 결과 | 임계값 |
|----------------------|-------|-------|--------|-------|-------|---------------|
| 1 | 0.359 | 0.343 | -0.069 | 4 | 3.526 | 3.5 ≤ P < 4.5 |
| 2 | 0.005 | 0.210 | 0.022 | 3 | 2.527 | 2.5 ≤ P < 3.5 |
| 3 | 0.597 | 0.285 | 0.015 | 3 | 2.960 | 2.5 ≤ P < 3.5 |
| 4 | 0.330 | 0.292 | 0.065 | 3 | 2.527 | 2.5 ≤ P < 3.5 |
| 5 | 0.363 | 0.343 | -0.055 | 4 | 3.719 | 3.5 ≤ P < 4.5 |
| 6 | 0.013 | 0.179 | 0.038 | 2 | 2.475 | 1.5 ≤ P < 2.5 |
| 7 | 0.603 | 0.273 | 0.038 | 3 | 2.742 | 2.5 ≤ P < 3.5 |
| 8 | 0.344 | 0.257 | 0.061 | 3 | 2.527 | 2.5 ≤ P < 3.5 |
| 9 | 0.398 | 0.281 | -0.076 | 4 | 3.512 | 3.5 ≤ P < 4.5 |
| 10 | 0.088 | 0.141 | -0.005 | 2 | 2.527 | 오분류 |
| 11 | 0.625 | 0.238 | 0.004 | 3 | 3.413 | 2.5 ≤ P < 3.5 |
| 12 | 0.380 | 0.202 | 0.037 | 3 | 2.527 | 2.5 ≤ P < 3.5 |
| 13 | 0.406 | 0.260 | -0.093 | 4 | 4 | 3.5 ≤ P < 4.5 |
| 14 | 0.121 | 0.131 | 0.014 | 2 | 1.949 | 1.5 ≤ P < 2.5 |
| 15 | 0.630 | 0.229 | -0.019 | 4 | 3.530 | 3.5 ≤ P < 4.5 |
| 16 | 0.388 | 0.188 | 0.038 | 3 | 2.527 | 2.5 ≤ P < 3.5 |
| 분류율(입력에 대한 정확한 분류비율) | | | | | | 93.75% |

3.5 Step 5: 출력단계(output)

본 절의 목표는 먼저 Digital Factory 시뮬레이터와 e-Diagnostic에 의한 공정 기기 관리 최적화시스템을 구현하고, 공정 상태의 확인 여부와 이상상태에 대한 조절 능력 확인하는 것이다. 이와 같은 목표는 모두 전문가의 지식을 이용하여 수행된다. 앞절에서 결과물에 대한 패턴을 정의했다면 이번 절에서는 목표 패턴 이외의 패턴에 대한 수정 지침이 제시되는 것이다. 이런 지침이 야말로 진정한 공정 관리 시스템의 목표라 할 수

있을 것이다. 표 6은 이러한 지침들을 정리한 것으로 제품의 상태에 따른 운전 규칙의 제어를 위한 것이다. 이러한 지침들은 신경망을 이용한 예측과 퍼지 추론기에 의한 추정에 기반한 의사지원시스템의 최종 출력이다. 이러한 전체 단위 시스템은 본 연구의 최종적인 목표인 Digital Factory 시뮬레이터와 e-Diagnostic에 의한 공정 기기 관리 최적화를 위한 어플리케이션으로 구현된다.

표 9. 정상과 이상 상태에 대한 패턴 및 지침.

| 결과 패턴 | 조절 변수 및 방향 |
|-------|--------------------------------------|
| 1(p1) | x1(에칭시간) 작게 하시오 |
| 2(p2) | x3(분사구 구경) 크게 하시오 |
| 3(p3) | 정상상태 |
| 4(p4) | x1(에칭시간) 작게 하고, x4(분사구 각도) 크게 하시오 |

4. 결론

제조 공정은 다양한 단위 공정이 결합된 복잡하고 동적인 시스템이다. 따라서 제품의 품질이나 수율에 영향을 미칠 수 있는 결함이나 고장을 일으킬 수 있는 상황들이 발생할 수 있다. 이러한 고장 인자들은 갑자기 발생하기보다는 특징적인 징후를 보인 후 나타나는 경우가 많다. 이러한 문제점을 진단하기 위하여 다양한 방법들이 연구되어 왔다. 기존의 연구에서는 주로 모델을 이용한 진단, 사례에 기반한 진단, 그리고 전문가 지식에 기반한 진단 기법들이 주로 사용되었다. 각 방법들은 장·단점을 가지고 있기 때문에 시스템의 특성에 따라 적절한 방법을 선택하고 필요시에는 조합적으로 사용하는 것이 효율적일 수 있다. 본 연구에서는 신경회로망을 이용한 예측 모델을 이용하여 현재 입력 운전 조건에서 생산될 제품의 특성을 예측하고자 하였다. 이러한 예측 후 제품의 결과에 대한 추론 기능이 포함된 퍼지 추론 시스템을 사용하여 최종적인 의사결정이 가능한 시스템을 구성하였다.

5. 참고문헌

[1] Jay Liebowitz, Knowledge Management Handbook, pp. 1-3, CRC Press, 1999.
 [2] Michael J. A. Berry, Data Mining Techniques, pp. 5-7, John Wiley & Sons, Inc. 1997.
 [3] 임봉진, 이병운, 고선우, 김재환, "다구찌 실험 계획의 구성원리에 관한 연구." 응용통계연구 제 4권 제 1호, 1991.