http://dx.doi.org/10.14386/SIME.2018.26.3.37

데이터 마이닝 기법을 이용한 차량용 반도체의 불량률 예측 연구

윤경식* · 정희운** · 박승범***

<목 차>

Ⅰ. 서론
Ⅱ. 이론적 배경
Ⅲ. 연구 방법
Ⅳ. 데이터 분석 및 결과
Ⅴ. 결론

국문초록 : 본 논문에서는 차량용 반도체가 제품 출하 후 사용 환경에 따라 발생되는 불량률을 데이터 마이닝 기법을 이용하여 분석하였다. 20세기 이후 가장 보편적인 이동 수단인 자동차는 전자 컨트롤 장치와 자동차용 반도체의 사용량이 급격히 증가하면서 매우 빠른 속도로 진화하고 있다. 자동차용 반도체는 차량용 전자 컨트롤 장치 중 핵심 부품으로 소비자들에게 안정성, 연료 사용의 효율성, 운전의 안정감을 제공하기 위해 사용되고 있다.

자동차용 반도체는 가솔린엔진, 디젤 엔진, 전기 모터를 컨트롤하는 기술, 헤드업 디스플레이, 차선 유지 시스템 등 많은 부분에 적용되고 있다. 이와 같이 반도체는 자동차를 구성하는 거의 모든 전자 컨트롤 장치에 적용되고 있으며 기계적인 장치를 단순히 조합

* TEXAS Instruments, 부장 (limit7828@gmail.com)
** 호서대학교 기술경영대학원 교수 (hwcheong@hoseo,edu)
*** 호서대학교 기술경영대학원 교수, 교신저자 (parksb@hoseo.edu)

한 이상의 효과를 만들어 내고 있다.

자동차용 반도체는 10년 이상의 자동차 사용 기간을 고려하여 높은 신뢰성, 내구성, 장기공급 등의 특성을 요구하고 있다. 자동차용 반도체의 신뢰성은 자동차의 안전성과 직접적으로 연결되기 때문이다. 반도체업계에서는 JEDEC과 AEC 등의 산업 표준 규격을 이용하여 자동차용 반도체의 신뢰성을 평가하고 있다. 또한 자동차 산업에서 표준으로 제시한 신뢰성 실험 방법과 그 결과를 이용하여 개발 초기 단계 및 제품 양산 초기 단계에서 제품의 수명을 예측 하고 있다. 하지만 고객의 다양한 사용 조건 및 사용 시간 등 여러 변수들에 의해 발생되는 불량률을 예측하는 데는 한계가 있다.

이러한 한계점을 극복하기 위하여 학계와 산업계에서 많은 연구가 있어왔다. 그 중 데이터 마이닝 기법을 이용한 연구가 다수의 반도체 분야에서 진행되고 있지만, 아직 자동차용 반도체에 대한 적용 및 연구는 미비한 상태이다.

이러한 관점에서 본 연구는 데이터 마이닝 기법을 이용하여 반도체 조립(Assembly)과 패키지 테스트(Package test) 공정 중 발생 된 데이터들간의 연관성을 규명하고, 고객 불량 데이터를 이용하여 잠재 불량률 예측에 적합한 데이터 마이닝 기법을 검증하였다.

주제어 : Automotive, 차량용 반도체, 데이터 마이닝, 불량률 예측

Prediction of field failure rate using data mining in the Automotive semiconductor

Gyungsik Yun · Hee-Won Jung · Seungbum Park

Abstract : Since the 20th century, automobiles, which are the most common means of transportation, have been evolving as the use of electronic control devices and automotive semiconductors increases dramatically. Automotive semiconductors are a key component in automotive electronic control devices and are used to provide stability, efficiency of fuel use, and stability of operation to consumers .

For example, automotive semiconductors include engines control, technologies for managing electric motors, transmission control units, hybrid vehicle control, start/stop systems, electronic motor control, automotive radar and LIDAR, smart head lamps, head-up displays, lane keeping systems. As such, semiconductors are being applied to almost all electronic control devices that make up an automobile, and they are creating more effects than simply combining mechanical devices.

Since automotive semiconductors have a high data rate basically, a microprocessor unit is being used instead of a micro control unit. For example, semiconductors based on ARM processors are being used in telematics, audio/video multi-medias and navigation. Automotive semiconductors require characteristics such as high reliability, durability and long-term supply, considering the period of use of the automobile for more than 10 years. The reliability of automotive semiconductors is directly linked to the safety of automobiles. The semiconductor industry uses JEDEC and AEC standards to evaluate the reliability of automotive semiconductors. In addition, the life expectancy of the product is estimated at the early stage of development and at the early stage of mass production by using the reliability test method and results that are presented as standard in the automobile industry. However, there are limitations in predicting the failure rate caused by various

parameters such as customer's various conditions of use and usage time. To overcome these limitations, much research has been done in academia and industry. Among them, researches using data mining techniques have been carried out in many semiconductor fields, but application and research on automotive semiconductors have not yet been studied.

In this regard, this study investigates the relationship between data generated during semiconductor assembly and package test process by using data mining technique, and uses data mining technique suitable for predicting potential failure rate using customer bad data.

Key Words : Automotive, Semiconductior, Data mining, Prediction of field failure rate

Ⅰ. 서론

20세기 이후 가장 보편적인 이동수단인 자동차는 전자 컨트롤 장치와 자동차용 반도체의 사용량이 급격히 증가하면서 진화하고 있다. 자동차용 반도체는 차량용 전자 컨트롤 장치 중 핵심 부품으로 소비자들에게 안정성, 연료 사용의 효율성, 운전의 안정감을 제공하기 위해 사용되고 있다(전황수, 2013).

예를 들어, 자동차용 반도체는 가솔린엔진, 디젤 엔진, 전기 모터를 컨트롤하는 기술, 트랜스미션 컨트롤 유닛, 하이브리드 자동차용 트레인 컨트롤, 스타트(Start)/스탑(Stop) 시스템, 전자 펌프 모터 컨트롤, 자동차용 레이다 및 라이다, 나이트 비전, 스마트 헤드램프, 헤드 업 디스플레이, 차선 유지 시스템 등에 적용되고 있다. 이와 같이 반도체는 자동차를 구성하는 거의 모든 전자 컨트롤 장치에 적용되고 있으며 기계적인 장치를 단순히 조합한 이상의 효과를 만들어 내고 있다.

자동차용 반도체는 높은 데이터 레이트를 기본적으로 가지고 있기 때문에 최근에는 마이크로 컨트롤 유닛이 아닌 마이크로 프로세서 유닛이 사용되고 있다. 예를 들어, ARM 프로세서를 기반으로 디자인/제작된 반도체가 텔레매틱스, 오디오/비디오 멀티미티어, 네비게이션 등에 사용되고 있다(D.S Ahn, 2012). 자동차용 반도체는 10년 이상의 자동차 사용 기간을 고려하여 높은 신뢰성, 내구성, 장기공급 등의 특성을 요구하고 있다. 자동차용 반도체의 신뢰성은 자동차의 안전성과 직접적으로 연결되기 때문이다. 반도체업계에서는 JEDEC[1]과 AEC[2] 규격을 이용하여 자동차용 반도체의 신뢰성을 평가하고 있다. 또한 자동차 산업에서 표준으로 제시한 신뢰성 실험 방법과 그 결과를 이용하여 개발 초기 단계 및 제품 양산 초기 단계에서 제품의 수명을 예상 하고 있다. 하지만 고객의 다양한 사용 조건 및 사용 시간 등 여러 변수들에 의해 발생되는 불량률을 예측하는 데는 한계가 있다. 이러한 한계점을 극복하기 위하여 학계와 산업계에서 많은 연구가 있어왔다. 그 중 데이터 마이닝 기법을 이용한 연구가 다수의 반도체 분야에서 진행되고 있지만, 아직 자동차용 반도체에 대한 적용 및 연구는 미비한 상태이다.

---

1)JEDEC : Joint Electron Device Engineering Council, 세계반도체표준협회 또는 국제반도체표준협의기구, 미국 전자 공업 협회(EIA)의 하부 조직으로, 제조업체와 사용자 단체가 합동으로 집적회로(IC) 등 전자 장치의 통일 규격을 심의, 책정하는 기구

2) AEC : Automotive Electronics Council, 전자장치의 승인 요구 사항들을 정의하는 위원회, The Automotive Electronics Council (AEC) was originally established by Chrysler, Ford, and GM for the purpose of establishing common part-qualification and quality-system standards.

이러한 관점에서 본 연구는 데이터 마이닝 기법을 이용하여 반도체 조립(Assembly)과 패키지 테스트(Package test) 공정 중 발생 된 데이터들간의 연관성을 규명하고, 고객 불량 데이터를 이용하여 잠재 불량률 예측에 적합한 데이터 마이닝 기법을 검증할 것이다. 본 연구는 다음과 같이 구성 되어 있다.

제 Ⅱ장 이론적 배경에서는 반도체와 자동차 산업, 차량용 반도체의 종류, 차량용 반도체의 신뢰도 요구 사항 및 신뢰도 예측 방법에 대해 기술한다.

제 Ⅲ장에서는 본 연구에서 제안하는 자동차용 반도체 제조 공정 중 조립(Assembly) 공정과 패키지 테스트(Package test) 공정에서 생성된 데이터를 이용하여 시장 불량률을 예측하는 방법에 대해 서술하였고, 제 Ⅳ장에서는 반도체 제조 공정에서 추출한 측정 데이터에 제안 기법을 적용하여 효과를 입증하였다.

제 Ⅴ장에서는 본 연구의 결론 및 경제적인 기대 효과와 함께 한계점에 대하여 논의하고 향후 연구 방향을 모색하였다.

Ⅱ. 이론적 배경

1. 반도체와 자동차 산업

반도체 산업은 1947년 벨 연구소에서 트렌지스터를 발명한 이후 전자, 정보, 통신 사업 부문과 함께 두드러진 성장 과 발전을 기록하고 있다. 특히 산업 발전과 고도 정보화 사회로 진입하면서 산업의 핵심 요소로 반도체의 수요가 급속히 확대되고 있으며 종류도 다양화 되고 있다. 개인용 및 산업용 컴퓨터, 핸드폰, 모바일 기기, 자동차까지 첨단 산업의 발전과 수요처의 고도화 및 다변화에 따라 반도체 시장 규모는 매우 빠르게 성장하고 있다(Pieter Pete B, 2000).

특히 자동차 산업에서는 자동차의 다양한 요구 기능을 해결 하기 위하여 기존의 기계식 제어 방식에서 전자식 제어로 대체하면서 자동차의 전자화가 시작되었으며(김경호, 2011), 이는 자동차용 반도체 시장 규모를 급격히 성장 시켰다. 자동차에 사용되는 전기 전자 제어 장치의 비율이 지속적으로 증가하여, 2010년 32%에서, 2015년에는 40%까지 원가 비율이 증가하였다. 이에 전기 전자제어 장치의 핵심 구성 요소인 차량용 반도체의 시장 규모도 2015년 41조로 증가하였다[4]. 자동차의 전자화가 점차 빠르게 진행 되면서 핵심 부품인 자동차용 반도체의 역할이 더욱 중요해지고 있다(김경호, 2011). 특히 영상

및 음성 처리와 더불어 MP3와 같은 외부 기기와 연결성이 증가되는 인포테인먼트 영역에서는 다양한 기능을 포함하는 반도체의 수요가 증가 하고 있다(김경호, 2011).

2. 자동차용 반도체의 종류

자동차에는 메모리/비메모리, 전원 반도체, 아날로그 IC, 디스크리트 파워, 센서, MCU (Micro Controller Unit)등 200여개 이상의 반도체가 사용되고 있으며, 고급 차종은 100개 이상, 일반 차종에는 평균 40개 정도의 반도체를 내장하고 있다(Chris Webber, 2013). 자동차용 반도체는 변속기와 엔진 상태를 제어하며 자동차 속 컴퓨터 역할을 수행하는 ECU(전자 제어 장치, Electronic Control unit)나 신호 처리, 센서 신호의 증폭, ASIC (주문형 반도체, Application Specific Integrated Circuit), ASSP (특정 용도 표준 제품, Application Specific Standard Product) 및 Analog IC 등의 비메모리 반도체, 자동차 내/외부를 측정하는 각종 센서, 각종 장치들을 직접 구동하는 구동장치 (Actuator)에 사용되는 파워 반도체 등을 총칭하는 용어로 통용되고 있다(이승규, 하상현 & 변영재, 2012).

3. 자동차용 반도체의 신뢰도 요구 사항

자동차용 반도체는 일반 반도체 보다 온도, 습도, 허용 불량률 등에 대한 요구 수준이 높다. 차량 고장 및 결함 등의 문제는 인명 사고와 직결되므로 차량용 반도체 단위부터 고 신뢰도를 확보하고 있다. <표 1>은 10년 동안 10만 mile(16만 km) 운행 기준으로 차량용 반도체 컴포넌트(component) 및 소자(element)별 결함 허용 목표치를 보여 준다 (D.S Ahn, 2012)

차량용 반도체는 고품질 확보를 위해서 다음의 조건들을 만족해야 한다.

첫째. 차량용 반도체는 자동차 부품 업체가 기본적으로 만족해야 하는 품질 경영 체계를 충족해야 한다.

둘째. 차량용 반도체 개발 프로세스는 제품 사양, 설계, 검증 등의 엔지니어링 활동과 프로세스 관리, 프로젝트 관리, 지원 활동을 포함해야한다.

셋째. 차량용 반도체는 차량의 신뢰도 확보를 위한 안전 요구 사항 및 이를 반영한 안전 설계 메커니즘을 만족해야 한다(도성룡 & 한혁수, 2014).

<표 1> 컴포넌트 및 소자의 결함 허용 목표치 (D.S Ahn, 2012)

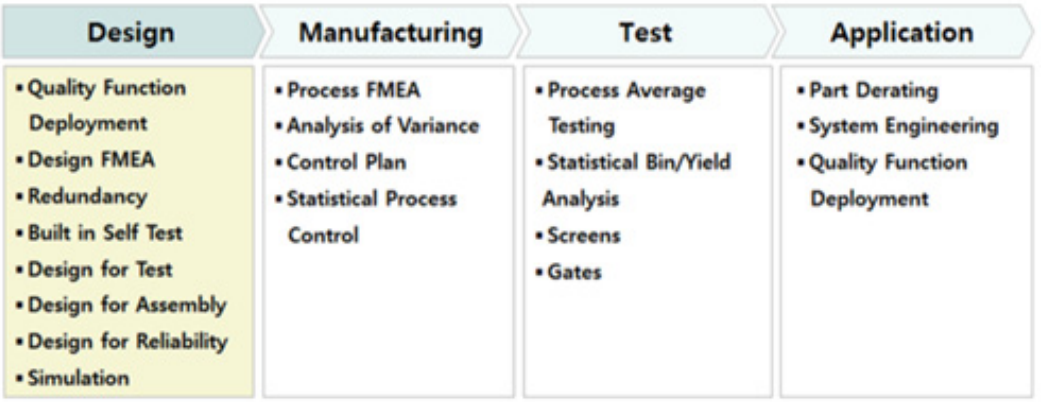
Division	Type of Failure Rete	Fault Tolerance Target
Component	ECU failure rate(field)	< 50 ppm
	ECU failure rate (0miles)	<15 ppm
	Modules and sensors failure rate	< 10 ppm
Element (Semiconductor)	ASIC failure rate	< 3 ppm
	IC failure rate	< 1 ppm
	Discrete components failure rate	< 0.5 ppm

고품질 차량용 반도체의 필수 조건은 신뢰도(Reliability)와 안정성(Safety)이다. 신뢰도(Reliability)는 제품이 5년, 10년 등 특정 기간 동안 고장(Failure) 없이 작동하는 정도를 의미한다. 안정성(safety)은 제품의 고장으로 인하여 사람에게 미치는 피해의 최소화 정도를 의미한다(S.Y Chae & W.J Kim, 2012).

차량용 반도체의 신뢰도 요구 사양은 <표 2>와 같이 타 산업에 사용되는 반도체에 비하여 매우 높다(B.C Kim, 2012).

<표 2> 산업별 반도체 신뢰성 사양 비교(B.C Kim, 2012)

Division	Home	Industrial	Automotive
Lifetime	1 - 3 years	5 -10 years	> 15 years
Temperature	0'C - 40'C	-10'C - 70'C	-40'C - 150'C
Humidity	Low	High	0 - 100%
Failure rate	<10%	< 1%	0%
Supply period	> 2 years	> 5 years	> 30 years



<그림 1> AEC-Q004 의 Zero Defect Flow

<표 3> AEC Q100[3]에 열거된 시험 조건과 기간의 적합성 여부

Loading	Mission Profile Input	Stress Test	Stress Condition	Acceleration Model	Calculated Test Duration	Q100 Test Duration
Operation	tu =12,000hr (Average operating use time over 15 yr) Tu = 87℃ (Average junction Temperature in use Environment)	HTOL	Tj = 125℃	Arrhenius	1393 hr	1000 hr
Thermo-Mechanical	nu = 54,750 cycle (number of engines on/off cycles over 15 yr of use) ΔTu = 76℃ (Average thermal cycle temperature change in use environment)	TC	ΔTu = 205℃ (-55℃~150℃)	Coffin Manson	Nt = 1034 cyc	1000 cyc
Humidity	tu = 131,400 hr (Average operating use time over 15 yr) RHu = 74% (Average relative humidity in use environment) Tu = 32℃ (Average temperature in use environment)	THB	RH = 85% Tt = 85℃	Hallberg-Peck	Tt = 960 hr	1000 hr

3) AEC-Q100: Failure mechanism Based Stress Test Qualification for integrated circuits (based document)

이에 따라 차량용 반도체는 자동차용 전자부품에 대한 신뢰도 평가 절차를 규정한 AEC(Automotive Electronic Council) 품질 규격을 적용한다. 특히 <그림 1>의 [4] AEC-Q004(Zero Defects Guideline) (AEC-Q004, 2006)와 같이 Zero Defect Flow를 이용한 전자 부품의 개발 단계 별 신뢰성 확보를 위한 기법을 적용하고 있다.

이러한 요구 사항을 만족하는 반도체의 제조 공정은 수 백 개의 정밀 공정으로 이루어진다.

4. 반도체 제조 공정 및 패키징 기술

반도체 제조 공정을 4가지로 분류하여 구분할 수 있다. <그림 2> 는 반도체 제조 공정에 대한 설명이다.

<그림 2> 반도체 제조 공정

반도체 제조 공정은 크게 원재료가 투입된 후 Fabrication (이후 FAB 공정), 프로브(Probe) 검사, 조립(Assembly) 그리고 패키지 테스트(Final test) 공정으로 나눌 수 있다 (R Uzsoy, CY Lee and LA Martin, 1992). FAB 공정은 실리콘을 이용하여 단결정 잉곳(Ingot)을 만든 후 가공하여 웨이퍼(Wafer)를 만든다. 이후 웨이퍼 표면에 패턴 및 층(Layer) 생성 공정 등의 세부 공정 반복을 통하여 수천, 수만 개의 집적 회로를 형성하는 단계이다. 프로브(Probe) 검사 공정은 FAB 공정이 완료된 웨이퍼 표면의 각 칩을 전기적 테스트를 통하여 양품과 불량품으로 판별하는 단계이다. 조립(Assembly) 공정은 웨이퍼의 표면의 각 칩(Chip)을 분리하여 PCB(Printed circuit board) 또는 리드프레임(Lead frame)에 실장하여, 칩의 물리적, 전기적 특성을 향상 시키고 외부 충격 및 데미지(Damage)로부터 칩을 보호하기 위한 단계이다. 마지막으로 패키지 테스트(Package

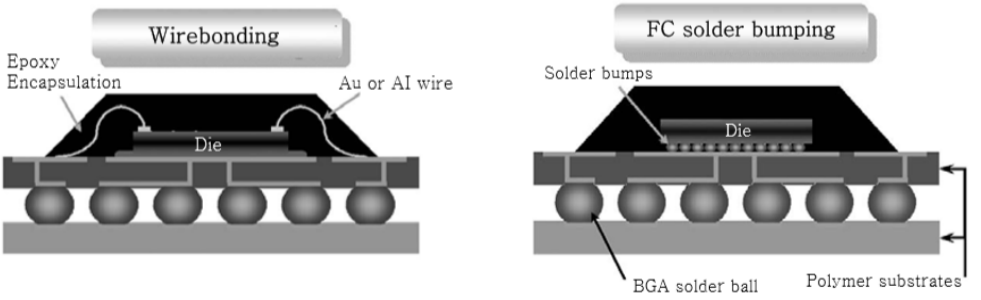
4) AEC-Q004: Zero defect guideline(describes a set of tools and processes which suppliers and users of integrated circuits can use to approach or achieve the goal of zero defect during a product's lifetime)

Test) 공정은 조립된 칩의 전기적 특성 및 기능, 신뢰성 등을 검사하여 양품과 불량품으로 분류하는 단계이다(DH Baek & CH Han, 2003).

멀티미디어 시스템 전자 제품들은 빠른 신호처리가 가능한 반도체 칩(Chip)의 개발 및 칩과 칩 또는 칩과 다른 주변 장치들 간의 상호 신호 전달을 위한 전자 패키징(electronic packing) 기술이 요구된다. 패키징 기술은 칩의 물리적인 보호와 보드와 칩 간의 전기적 연결이 외에 칩 성능의 최적화, 발열처리 문제 등을 해결하고 있다(JW Yoon, JW Kim & et all, 2007). 칩(chip)의 미세화, 고집적화가 진행되고 있으며, 이와 함께 패키지 타입도 경박단소화(經薄短小化) 되고 있다(J.H, Lau, 2001). 패키지 타입은 리드프레임을 이용한 QFP (Quad Flat Package), PCB를 이용한 BGA (Ball Grid Array)로부터 CSP (Chip Scale Package) 및 플립칩 (Flip Chip) 패키지의 형태로 발전해 오고 있다(J.H, Lau, 2001).

패키지 타입은 칩을 기판에 실장하는 기술에 따라 와이어 본딩(Wire bonding) 패키지와 플립칩 (Flip chip) 패키지로 구분된다. <그림 3> 은 와이어 본딩과 플립칩 본딩 패키지 형태이다.

와이어 본딩(Wire bonding) 패키지는 매우 작은 직경(12um 이상)으로 가공된 금, 은, 알루미늄 및 구리 등의 금속선을 사용하여 칩과 패키지를 연결한다. 패키지 특성상 입력(input) 부분을 칩의 외곽에만 배치하는 주변정렬(Peripheral array) 방식의 칩을 사용하기 때문에 입력(Input)/출력(Output) 수에 제한이 있다.

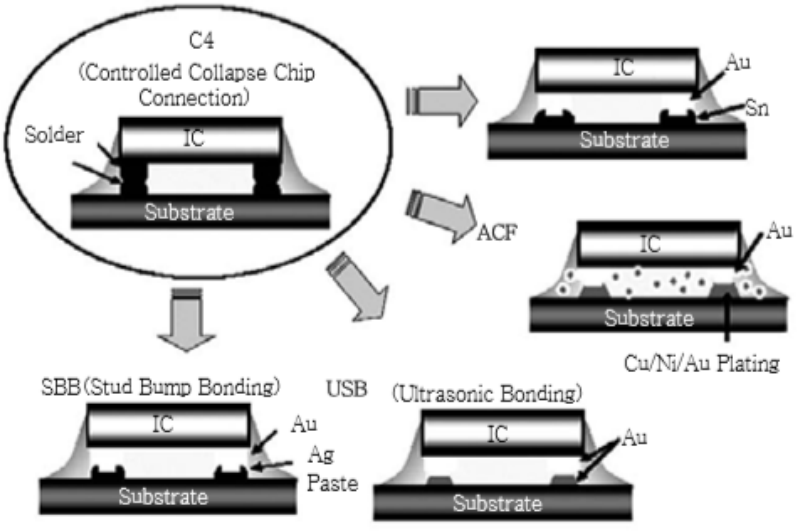


<그림 3> 와이어 본딩과 플립칩 본딩 패키지 비교 (JW Yoon, JW Kim & et all, 2007)

플립칩(Flip chip) 패키지는 다양한 재료(Solder bump, Cu pillar, Au stud bump, conductive polymer and paste) 및 방법(Deposition, plating, Ball screen printing, Thermal compression bonding, Mass flow)을 이용한 접속(interconnection)을 통하여

칩의 표면이 기판을 향하도록 하여 칩과 패키지를 연결한다. 와이어 본딩 패키지와 달리 칩(Chip) 표면 전체에 입력(input) 부분을 배치하는 면 정렬(Area array) 방식을 사용하여 입력(Input)/출력(Output) 수를 증가 시킨다(J.H, Lau, 2001). <그림 4> 다양한 플립칩 본딩 방법이다. 본 연구에서는 솔더 범프(solder bump)와 리플로우(Reflow)를 이용한 플립칩 패키지 데이터를 이용하였다.

반도체 제조는 FAB공정부터 패키지 테스트까지 수 백 개의 제조 공정을 수개월에 걸쳐 진행하여 각 공정에서 가공되는 제품 및 품질 관리가 복잡하고 어려우며, 제조 공정이 나노 미터급 공정으로 소형화, 미세화 됨에 따라, 각 공정 중 발생된 가공 단계 영향이 최종 제품의 성능에 크게 영향을 미친다. 특히 각 공정에서 가공된 제품의 작은 영향이 지속적으로 누적되어 최종 제품 불량으로 나타나므로 관리가 더욱 어렵다 (안대웅, 고효헌 & 백준걸, 2009). 이러한 제품 불량은 반도체 제조 공정에서 검출되기도 하지만 고객에 판매 한 이후 불량이 발생하기도 한다. 이러한 고객 판매 후 불량을 예측하는 것을 신뢰도 예측이라고 하며 이에 대한 연구가 진행되어 왔다.



<그림 4> 다양한 플립칩 본딩 방법 (JW Yoon, JW Kim & et all, 2007)

5. 차량용 반도체의 신뢰도 예측

신뢰도 예측이란 “ 부품 또는 시스템의 운영 및 사용 환경과 조건을 고려하여 고장률이나 평균수명 등과 같은 신뢰성 척도를 예측하는 것”을 말한다. 시장 불량률 예측 연구는 신뢰성 공학을 기반으로 한 고장률 예측 방법론과 신뢰성 공학 이외 분야에서 연구된 방법론으로 나눌 수 있다.

신뢰성 공학 분야의 예측 방법론은 수명 분포함수를 이용하는 방법으로서 가속 수명시험에 의한 예측과 시장 반입 데이터를 이용한 예측으로 구분된다. 이외에 신뢰성 부품 규격을 이용한 방법이 있다(Shin K.Y and Han J.H, 2011).

가. 신뢰성 공학 분야의 예측 방법론

첫 번째 가속수명시험 (ALT: Accelerated Life Test) 에 의한 수명분포 함수 추정 방법은 전자제품, 반도체 분야의 발전으로 제품의 수명 주기가 매우 짧아지고 신제품 개발이 지속적으로 이루어짐으로써 제품의 품질, 신뢰성을 짧은 개발 기간에 검증하고 보증하기 위하여 개발된 방법이다. 2013년에 [5]IEC 62506(Methods for Product Accelerated Testing)가 제정되었으며 가속수명시험(ALT)을 기반으로 한 가속열화시험(ADT: Accelerated Degradation Test)과 초가속수명시험(HALT; Highly Accelerated Life Test) 이 있다(Yum B-J, Seo S-K and et all, 2014). 가속수명시험에 의한 수명 분포 함수 추정 방법은 제품을 가혹한 조건에서 시험하는 도중 고장이 발생한 시간 정보를 기초로, 일반 사용 조건에서의 고장 시간으로 환산하여 제품의 수명을 예측하는 방법이다. 이때 사용되는 가속 모형은 제품의 수명과 스트레스의 관계로서 수명 분포의 모수를 스트레스 변수들의 함수로 표현하므로, 가속수명시험 계획의 수립과 분석이 매우 중요하다(Lee S-H and Chung H-J, 2009). 여러 부품을 조립한 시스템의 고장률을 예측하기 위해서는 부품 단위의 고장률을 기초로 각 부품들의 특성을 찾아서 조립 제품의 신뢰도를 추정하는데, 부품마다 고장 형태를 가속시키는 스트레스의 종류가 다르기 때문에 적당한 가속인자를 선정하기 어렵다(Shin K.Y and Han J.H, 2011). 또한 고가의 제품을 시험하는 경우 비용 및 시간의 제약으로 소량의 샘플을 평가한다. 이러한 경우 불량수가

---

5) IEC-International Electronical Committee, 전기 기술에 관한 표준의 국제적 통일과 조정을 목적으로 1906년에 설립된 ‘국제전기표준회의’를 이른다. 1947년에 ISO(국제표준화기구)가 발족됨에 따라 ISO의 전기 부문으로서 가입했다. 80개의 기술 위원회가 있으며, 현재까지 약 2,000개의 IEC 국제 규격(IEC publication)을 제정

극히 적어 가속수명시험 결과로 시장 불량률을 예측하는데 한계가 있다(Lee D.K, 2005). 두 번째로 시장반입 데이터에 의한 수명분포함수 추정 방법은 일정 기간 동안의 고객 불량 수와 불량이 발생하지 않은 제품 수를 바탕으로 수명분포함수를 추정하여 제품의 수명을 평가한다. 여러 수명분포함수 중 모수에 따라서 증가 불량률, 감소 불량률, 상수 불량률을 나타낼 수 있는 와이블 분포 (Weibull distribution)를 일반적으로 사용한다. 와이블 분포는 신뢰성 공학에서 실패 분석, 부품의 수명 추정 분석, 산업 현장에서 어떤 제품의 조제에 걸리는 시간을 추정 할 때에도 널리 사용된다(CF Chien, WC Wang and JC Cheng, 2007). 이 방법은 분포를 추정하기 위한 시장반입 데이터의 기간에 따라서 예측 시점이 달라질수 있으며, 시장 반입 데이터에는 실제 제품의 사용 시간 뿐 아니라, 운송 시간, 고객이 사용하지 않는 시간을 포함하고 있기 때문에 실제 환경에서는 수명분포의 적용이 어렵다는 한계점이 있다(Wilson, S, Joyce, T & Lisay, E, 2009).

마지막으로 신뢰성 부품 규격을 이용한 예측 방법으로 MIL- HDBK-217F에서 Bellcore Method, Telcordia SR332, RiAC HDBK- 217PLUS(통계적 접근) 등 라이브러리를 이용한다. 이러한 방법은 가속수명시험보다 신속히 신뢰도를 평가 할 수 있는 장점이 있지만 부품 규격에 맞춰 전체 시스템의 신뢰도를 예측하는 방식으로서 예측 값과 실제 값과 차이가 크고, 규격에 제약이 많은 단점이 있다(Lee D.K, 2005).

나. 신뢰성 공학 이외 분야의 예측 방법론

2000년대로 오면 지금까지 문제점들이 노출된 고장률 예측을 위한 각종 신뢰성 데이터베이스(예를 들어, MIL-HDBK 217[6]) 등을 대체할 수 있는 고장물리 분야가 더욱 다양하게 연구되었으며, 시스템 신뢰성 분야에서는 대형시스템분석을 위한 상용 소프트웨어의 개발, 시뮬레이션 방법론의 응용, 기능안전 (functional safety) 등 안전성 및 위험분석과의 통합화에 관한 연구가 수행되었다(Yum B-J, Seo S-K and et all, 2014). 특히 최근에는 재료, 부품, 장비의 고장 예측 진단 기술의 개발과 활용 (Prognostics and Health Management) 이 요구되고 있으며 이를 위한 확률과정 이론, 베이지안 방법론에 관한 연구가 활발히 이루어지고 있다(Yum B-J, Seo S-K and et all, 2014).

신뢰성 공학 외 기타 분야에서 데이터 마이닝을 이용한 고장률 예측 연구 사례로는 서포트 벡터머신 (Support vector machine) 또는 순위합과 나이브베이즈 (Naive Bayes)를 결합한 알고리즘기반 하드드라이브 고장률 예측, 다중 회귀분석 기반 소프트웨어 고

6) MIL- HDBK-217F(군용 핸드북- 전자장비 신뢰도 예측, Military Handbook. Reliability prediction of electronic equipment)

장률 예측 그리고 의사결정 나무기반 주상변압기 고장 여부 예측 연구를 들 수 있다.

그러나 이러한 데이터마이닝 기반의 연구들은 대부분 반도체 외 다른 산업 분야에서 연구된 것이며, 반도체 산업에 적용된 사례는 FAB 공정에서 발생된 공정 변수를 이용한 변수선택기법, 인공 신경망 기법과 메모리 기반 추론(Memory Based Reasoning) 기법을 사용한 랏 단위 프로브 수율 예측 모델 구축 (CK Shin, SC Park, 2000), 저수율의 원인이 되는 FAB 공정의 공정 변수를 이용한 유전자 프로그래밍(Genetic Programming) 기법 기반 랏 단위 프로브 수율 예측 및 분류하는 모델 구축 (TS Li, CL Huang& ZY Wu, 2006), FAB 공정에서 측정된 데이터를 변수로 이용하여 K-평균 군집 분석, Kruskal-Wallis 검정, 의사결정나무기법을 사용하여 랏 단위로 프로브 수율을 예측하는 모델 구축 (CF Chien, WC Wang and JC Cheng, 2007), 공정 제어 모니터링 데이터와 프로브 검사 공정에서 출력된 데이터를 변수로 이용하여 SSVM(Stepwise Support vector machine) 기법을 사용하여 각 랏 단위의 최종 수율을 고수율과 저수율로 구분하는 분류 예측 모델 구축(D An, HH Ko & T Gulamber, 2009) 등이 있다. 이러한 연구들은 FAB공정에서 발생된 데이터를 이용하여 수율을 예측하는 연구이며, 차량용 반도체의 시장 불량률에 관한 연구 문헌을 찾아보기 힘들다.

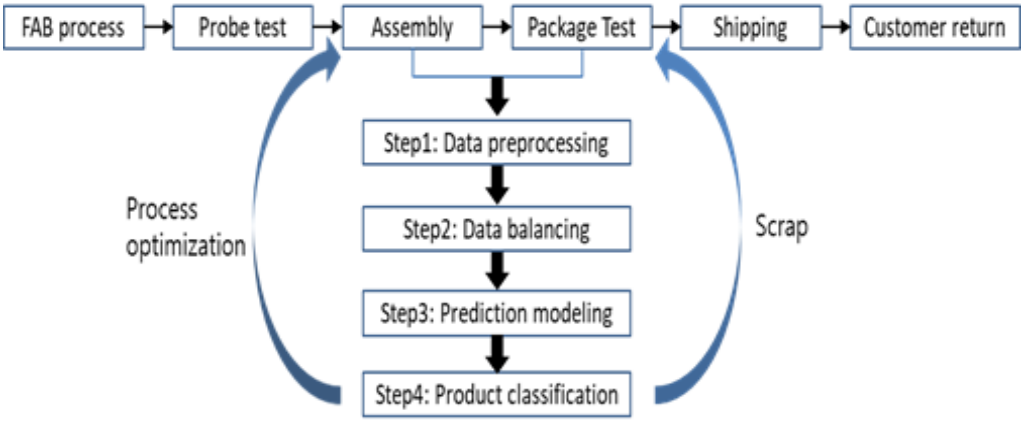
본 연구에서는 데이터 마이닝 기법을 이용한 차량용 반도체 제조 공정 중 조립(Assembly) 공정과 패키지 테스트(Package test) 공정에서 생성된 데이터를 공정변수로 사용하여 시장 불량률을 예측하는 모델을 제안하고자 한다. 이를 통하여 조립(Assembly) 공정과 패키지 테스트(Package test) 공정에서 잠재 불량 제품을 선별 폐기하는 방안을 마련하도록 한다. 이에 따라 잠재적인 고객 불량률 감소와 불량 제품에 대한 원인 및 개선 방안 도출 등에 필요한 품질 비용 절감 효과 그리고 Zero Defect[7] 실현을 통한 고객 신뢰성 향상을 기대 할 수 있다.

---

7) Zero Defect: 무결점. 최근 글로벌 자동차 회사들의 잇따른 대량 리콜 사태는 제조 기업들에게 품질에 대한 경각심이 높아지고 있으나, 기업 간의 경쟁이 치열한 상황에서 원가를 고려하지 않고 품질 경쟁력 향상에만 매진 할 수도 없다. 제조 기업들은 시스템적 접근을 통해 불량률 제로(Zero Defect)을 실현하고 있다.

Ⅲ. 연구방법: 자동차용 반도체 시장 불량률 예측 모델링 방법

본 연구에서 제안하는 시장 불량률 예측 모델은 반도체 제조 공정 중 조립(Assembly) 공정과 패키지 테스트(Package test) 공정에서 생성된 공정 데이터를 기반으로 잠재 불량 제품과 관련된 주요 공정 변수를 구분하는 분류 모델을 구축하는 것이다. 본 연구에 자세한 설명은 다음과 같다. 서 제안하는 시장 불량률 예측 모델링 프로세스는 <그림 5>와 같이 4단계로 이루어지며



<그림 5> 시장 불량률 예측 모델 프로세스

1. 데이터 전처리 (Step1)

첫 번째 단계는 데이터 전처리 작업이다. 자동차용 반도체 제조 공정 중 조립(Assembly) 공정과 패키지 테스트(Package Test) 공정 지식을 바탕으로 제품의 품질을 대표하는 특징을 갖는 변수를 선택한다. 데이터 마이닝 알고리즘을 기반으로 하는 여러 가지 공정 변수 선택 기법들이 있으나, 엔지니어의 경험을 통해 선택하는 방법이 강건하므로(Nam W.S&Kim S.B, 2015), 본 연구에서는 최종적으로 공정 지식을 바탕으로 시장 불량률에 영향을 미치는 주요 변수를 선정 하였다.

자동차용 반도체 제조 공정 중 조립(Assembly) 공정은 기존의 반도체 제조 공정과 동일하지만 통계적 공정 관리 기법을 적용하고 있으며, 고객별 특별 요구 사항에 따라

특별 관리가 이루어지고 있다. 조립(Assembly) 공정은 수백에서 수 천 개의 제품으로 구성된 랏(Lot) 단위로 관리되며, 각 랏(Lot)은 모든 공정에서 고유의 측정(Measurement) 및 모니터링(Monitoring) 결과를 가진다. 하지만 조립(Assembly) 공정의 특성상 모든 제품에 대하여 test를 진행할 수 없기 때문에 초기 셋업(Set-up) 에는 모니터링 용 더미(Dummy) 자재를 이용하여 랏 에 대한 전체 대용 특성 값을 측정한다. 동일한 제품을 연속 생산 할 경우에는 모니터링 방법을 통하여 랏의 특성 값으로 사용한다. 각 공정 별로 셋업(Set-up)과 모니터링(Monitoring)은 관리 항목으로 관리 되고 있다. 따라서 변수 선정에 따라 시장 불량률 예측 모델의 신뢰도가 달라진다.

패키지 테스트(Package test) 공정에서는 자동차용 반도체의 사용 환경을 고려하여 상온 (Room temperature) 및 고온(High Temperature) 조건에서 반복 테스트를 진행한다. 각 테스트 결과는 각 단위 테스트 별 수 백 개의 테스트 로그 기록을 생성한다. 이렇게 생성된 로그는 제품의 기능 특성을 고려하여 그룹화 하는데 이것을 빈(BIN)이라고 한다. 각 단위 테스트 별 최종 결과는 각 빈(BIN)의 특성에 따라 관리한다.

본 연구에서는 조립 공정 중 모든 랏에 대한 테스트 결과를 가지는 공정 변수를 선택하였고, 패키지 테스트 공정에서는 전체 빈(BIN)을 선정하였다. 전체 데이터에서 Missing value가 존재하는 record를 제거하였다. 테스트수량, 불량률, 수율등의 수치형변수(Numeric data)들은 사용되는 측정 단위의 정규화를 통하여 분석시 성능향상을 도모하였다.

2. 데이터 불균형 해결 (Step2)

현실에서 다루는 많은 분류 문제의 데이터는 불균형 상태에 놓여 있으며(C Chen, A Liaw and L Breiman, 2004), 이는 알고리즘의 예측 성능을 저하시키는 주요 요인 중 하나이다(P Kang& S Cho, 2006). 자동차용 반도체 제조 공정의 안정화로 공정 수율이 매우 높고, 시장 불량률이 매우 낮아, 조립(Assembly) 공정과 패키지 테스트(Package test)에서 양품과 불량 랏의 비율이 심각한 불균형 상태가 된다. 그러므로 각 랏(Lot) 를 양품과 불량으로 구분하는 분류 모델 구축을 위해서는 데이터 불균형 문제를 해결해야 한다. 본 연구에서는 데이터 불균형 문제를 해결하기 위하여 SMOTE(Synthetic Minority Over-sampling Technique) 기법을 사용하였다. SMOTE 알고리즘은 소수 계층 과다 추출, 다수 계층 과소 추출의 조합으로 데이터의 불균형을 해결하는 방법이다.

샘플링 방법으로 K이웃근접기법을 사용하였다. K이웃근접기법은 소수 계층의 데이터를 새롭게 생성하여 소수 계층의 결정 경계를 넓히는 효과를 나타내 보편적으로 사용되는 복원 샘플링 기법이 가지고 있던 데이터 과적합 문제를 해결하는 장점이 있다 (NV Chawla, KW Bowyer & Lo Hall, 2002).

본 연구에서는 조립(Assembly) 과 패키지 테스트(Package test) 공정에서 랏(Lot) 선정 시 시장 불량 랏(Lot)를 기준으로 한 K 이웃근접기법을 사용하여 샘플링을 하였다.

3. 시장 불량률 예측 모델 구축 (Step3)

이 단계에서는 전 단계에서 샘플링 된 랏(Lot)의 조립(Assembly) 및 패키지 테스트(Package test) 데이터를 독립변수로 구성하였다. 독립변수는 총 86개로 조립(Assembly) 데이터 39개와 패키지 테스트 47개로 구성하였다. 종속변수는 시장 불량과 시장 미 불량으로 구분한 범주형 변수를 활용하여 분류 모델을 구축한다. 본 연구에서는 분류 문제에 널리 사용되고 있는 예측 모델의 성능을 비교 검증하기 위해 다음 기법들을 고려하였다.

첫째, 결과에 대한 해석이 쉽고 학습 시간이 짧고 의사 결정 규칙을 도표화하여 관심 대상이나 집단을 몇 개의 소집단으로 분류하거나 예측하는 방법인 의사결정나무(Quinlan,J.R, 1986)를 고려하였다. 둘째로, 문서 분류, 영상 인식, 문자 인식 등 여러 분야에서 우수한 성능을 보여주며, 분류와 예측에 모두 응용 할 수 있는 알고리즘으로 크게 SVC(Support Vector Classification)과 예측을 위한 SVR(Support Vector Regression)으로 구분되며 최적의 초평면으로 분류 문제를 해결하는 서포트벡터머신(SVM) (Vapnik,V, 2000), 셋째로, 선형 모델로 분류 문제를 해결하는 로지스틱 회귀분석(Hosmer, D. & Lemeshow, W, 1989), 패턴인식이나 데이터 마이닝 등 다양한 분야에 응용되고 있으며 분류 문제 해결을 위한 예측 알고리즘으로도 사용되는 인공신경망(Rosenblatt, F, 1958) 기법을 고려하였다. 마지막으로 랜덤 포레스트(이하 RF) 기법을 사용하였다. RF 알고리즘은 나무 유도 과정에서 임의의 변수 또는 특정 변수를 선택하여 학습 데이터의 부트스트랩(Bootstrap) 샘플로 유도된 잘라내지 않은 분류 또는 회귀 나무의 앙상블 이다(C Chen, A Liaw and L Breiman, 2004). RF 알고리즘은 다른 분류 알고리즘과 비교하여 높은 예측력을 보여주고 있다(Breiman, L., 2001). RF 알고리즘은 모델 구축 시 사용자가 결정해야하는 매개변수의 수가 적고, 그 값에 덜 민감하기 때문

에 모델 구축이 매우 용이하다(Liaw, A. & Wiener, M., 2002).

본 연구에서는 랜덤포레스트 기법을 이용하여 불량 예측 모델을 구축하였고, 교차 검증 방법을 이용하여 불량 예측 알고리즘의 신뢰도를 평가하였다.

4. 새로운 제품의 불량률 예측 및 공정 관리 (Step4)

Step1부터 Step3까지의 단계적인 과정을 거쳐 구축된 불량 예측 모델을 이용하여, 새롭게 생산되는 제품의 Zero defect 을 실현하기 위하여 조립(Assembly) 및 패키지 테스트(Package test) 공정 관리 인자 및 관리 범위를 최종 설정하였다.

Ⅳ. 데이터 분석 및 결과

1. 분석 데이터

본 연구에서 제안하는 시장 불량률 예측 모델의 효용성을 입증하기 위해 아날로그 반도체, 모바일 AP 반도체 및 automotive 반도체등을 주력으로 생산하는 외국계 반도체 기업인 A 기업의 실제 조립(Assembly), 패키지 테스트(Package test) 및 시장 불량 데이터를 이용하였다. 실험에 사용된 데이터는 최근 3년간 발생된 제품의 데이터이며 사용변수는 A 기업이 반도체 생산 및 품질관리를 위하여 선정 및 관리하는 86개의 품질지표 <붙임1>를 이용하였으며, 시장 불량 발생 랏(Lot) 데이터 132개, 시장 불량이 발생하지 않은 랏(lot) 데이터 1175개로 구성되었다. 시장 불량 발생 랏과 불량이 발생하지 않은 랏 간 비율이 8.9대 1로 범주 간 불균형을 보여주고 있다. 데이터는 변수의 특성에 따라 연속형 독립변수와 범주형 종속변수로 이루어진 데이터를 구성하였다.

2. 분석 계획

여러 기계학습 알고리즘은 서로 저마다 적합한 가설의 데이터에 적용될 때 최고의 성능을 발휘한다. 본 연구에서 사용한 데이터의 경우 변수가 86차원 수준으로 그리 차원이 높지 않고, 각각의 변수들 간의 상호 의존성이 매우 적은 특징을 가지고 있다. 각 변수간 의존성이 적은 이유는 조립 공정에서 검사하는 항목과 패키지 테스트 공정에 검사하는 특성이 서로 중복되지 않는 각 공정의 특성 요인에 대한 검증을 하기 때문이다. 또한 변수 자체가 이미 고수준의 추상적인 정보를 담고 있기 때문에 이를 따로 인지하기 위한 Feature engineering 이나 deep learning 을 이용한 feature learning등이 크게 필요하지 않은 상황이다.

이러한 메타 정보 기반의 데이터는 Tree 계열 알고리즘을 적용 하였을 때 매우 높은 일반화 성능을 기대할 수 있다. Tree 계열 알고리즘은 모두 명시적인 input 변수가 주어진 상황을 가정하고, 각 변수의 threshold 를 학습하여 공간을 구분하는 알고리즘이기 때문이다. 다른 서포트벡터머신(SVM) 이나 Deep learning 등과 비교 하였을 때 적은 수의 데이터만으로 충분히 높은 예측 성능을 기대할 수 있으며 학습 속도 또한 매우 빠르기 때문이다.

본 연구에서 사용한 트리 기반의 알고리즘은 기본적인 decision tree와 앙상블 및 부스팅 계열의 Random forest, AdaBoostClassifer, ExtraTreesClassifier등을 활용하였다. 또한 어떠한 변수가 잠재적인 고객 불량을 예측하는데 중요한 정보를 지니고 있는지를 알아내는 Feature importance analysis를 하기 위한 연구를 진행하여 중요도가 높은 Top-N 개의 변수를 확인하였다.

3. 분석 결과

본 연구에서는 가장 먼저 앙상블 기반의 알고리즘인 Random Forest 를 이용하여 데이터에서 추후 불량이 발생한 경우를 True, 그렇지 않은 경우를 False로 예측하는 2 class 분류 문제를 학습하였다. Random Forest 는 다른 알고리즘에 비해서 하이퍼 파라미터 튜닝을 적게 하여도 충분히 높은 예측 성능을 나타내며, 학습속도가 빨라 컴퓨팅 환경에 대한 제약이 적은 알고리즘이다. 이 알고리즘에서 50개의 tree를 사용하도록 지정하고 3-Fold Cross-validation(Kohavi, R., 1995)을 수행한 결과 92.8%의 예측정확도

를 얻었다. 본 분류 문제의 성능 척도를 정오분류표 <표 4>를 통하여 한 정확도(Accuracy)를 계산하여 사용한다. 하지만 정확도는 데이터의 범주간 불균형이 심한 경우 관측치가 많은 범주에 편향되기 때문에 올바른 평가 척도로 사용되기 어렵다(P Kang& S Cho, 2006).

<표 4> Contingency table for test[8)]

		Predict	
		양호	불량
Actual	양호	1082 (TP)	93 (FN)
	불량	16 (FP)	116 (TN)

정오분류표에 따르면 본 연구에서 사용된 1307개의 training 데이터에서는 100%의 정확도를 보였다. 다만 test에서는 다소 예측 성능이 감소하였는데 이는 데이터 개수의 부족과 데이터 자체가 가지고 있는 정보량의 한계로 인해 일반화 성능이 저하된 것으로 보인다.

Zero-R classification을 한다고 가정했을 때의 성능은 0.897으로 Random Forest에 비하여 약 2% 정도 낮은 test 정확도를 보인다. 이는 데이터 구성이 각 검사단계에서의 불량만을 가지고 있고, 추후 발생한 불량의 원인이나 사용 기간, 사용처 및 사용 환경 등은 전혀 포함되지 않은 데이터이기 때문에 이 정보만 가지고도 이러한 성능 증가를 보였다는 것은 어느 정도 의미 있는 학습이 이루어졌음을 보인다.

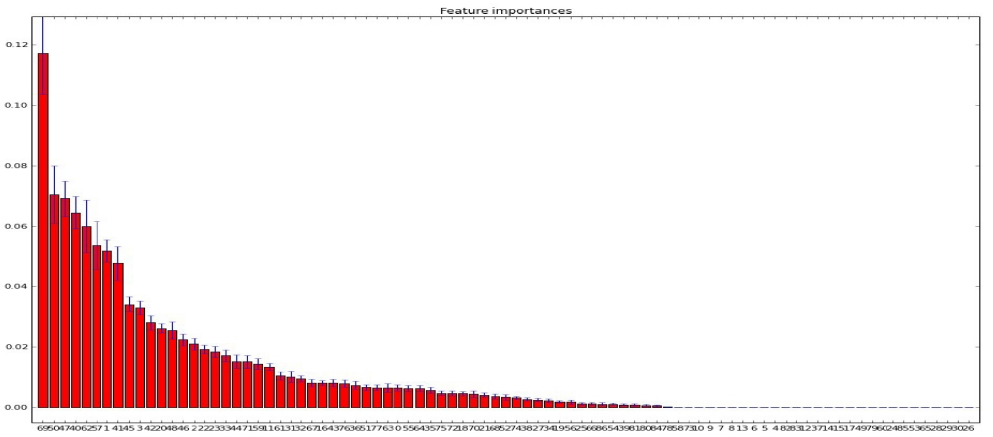
불량 여부를 예측하는 학습을 수행한 후, 학습된 모델을 이용하여 어떤 변수가 불량 여부를 예측하는데 많은 정보를 가지고 있는지를 다음과 같이 분석하였다.

위의 분석은 전체 86개의 변수 중 중요도가 높은 순서로 정렬하고, 50개의 tree에서 얻어진 각각의 변수 중요도의 평균과 분산을 표시한 것이다.

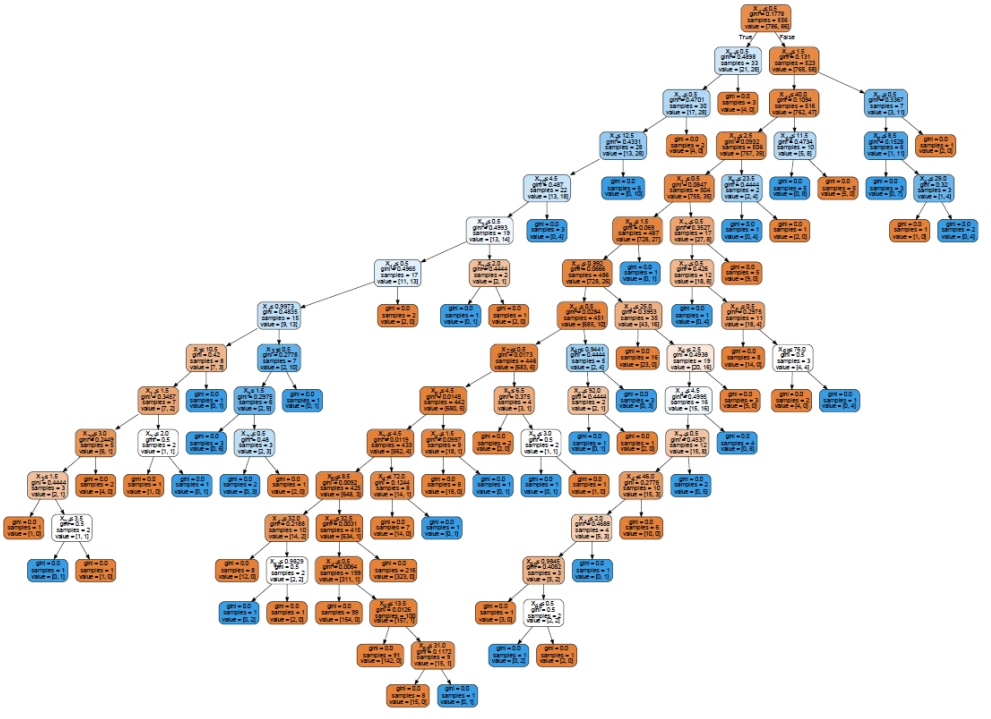
<그림 6> Feature importance의 경우, 각 tree에서 해당 변수가 얼마나 많은 데이터의 불확실성을 줄이는 데 기여했는지를 측정하여 구한다.

---

8) TP: the proportion of low yields that are correctly predicted,
FN: the proportion of low yields that are incorrectly predicted as high yields,
TN: the proportion of high yields that are correctly predicted,
FP: the proportion of high yields that are incorrectly predicted as low yields.



<그림 6> 시장 불량률 예측을 위한 feature importance



<그림 7> The structure of Decision Tree

<그림 7> The structure of Decision tree는, 50개의 앙상블된 본 tree 중 일부를 선택하여 해당 tree 의 구조를 확인하여 어떠한 과정을 거쳐 최종적인 불량 예측을 수행하였는지 명시화 하였다.

본 연구에서는 마지막으로 DecisionTree, RandomForest, ExtraTrees, AdaBoost 의 총 4가지 알고리즘을 이용해 동일한 데이터에서 앞서 분석한 feature importance를 바탕으로 Top-9 개의 feature를 선별하고, 이 중 2개의 feature만을 골라 알고리즘을 학습시켜보았다. 또한 학습된 결과를 이용해 선택된 2개의 feature에 대한 Decision Boundary를 시각화함으로써 어떠한 공간을 통해 최종적으로 분류가 이루어지는지 분석해볼 수 있다.

<그림 8> 4종의 트리(tree) 기반 알고리즘을 이용한 학습결과

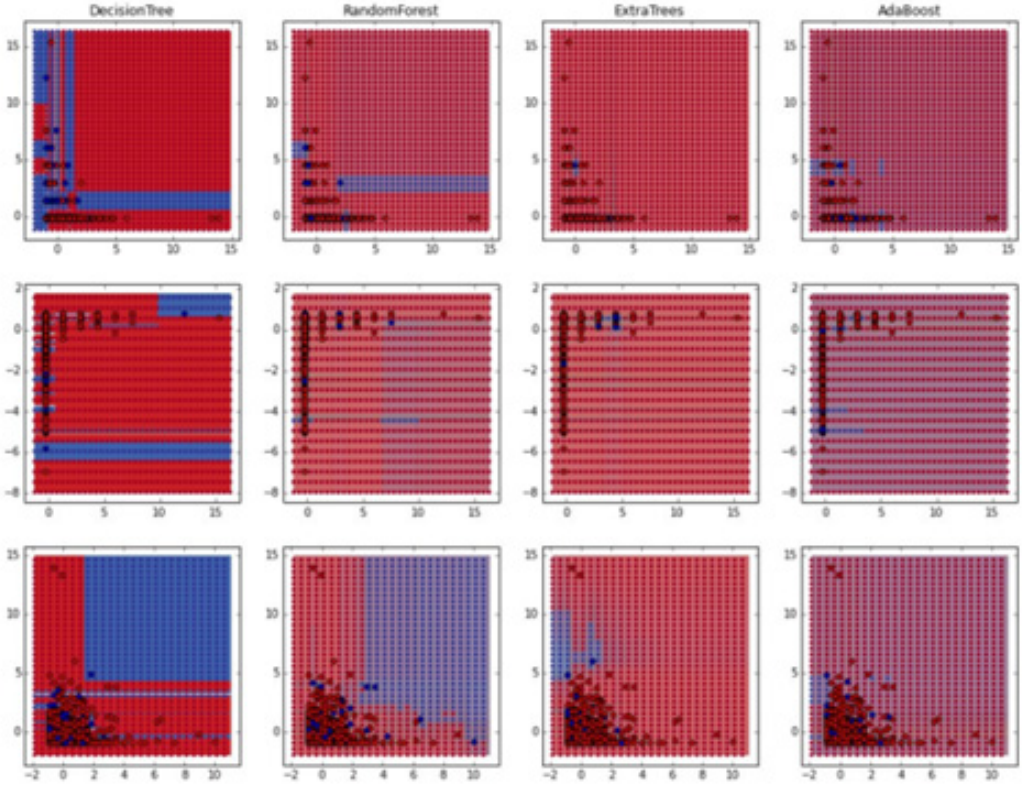
```
('DecisionTree with features', array([69, 41]), 'has a score of', 0.9357798165137615)
('RandomForest with 50 estimators with features', array([69, 41]), 'has a score of', 0.9074923547400611)
('ExtraTrees with 50 estimators with features', array([69, 41]), 'has a score of', 0.9036697247706422)
('AdaBoost with 50 estimators with features', array([69, 41]), 'has a score of', 0.9051987767584097)
('DecisionTree with features', array([41, 50]), 'has a score of', 0.9380733944954128)
('RandomForest with 50 estimators with features', array([41, 50]), 'has a score of', 0.9319571865443425)
('ExtraTrees with 50 estimators with features', array([41, 50]), 'has a score of', 0.9426605504587155)
('AdaBoost with 50 estimators with features', array([41, 50]), 'has a score of', 0.9350152905198776)
('DecisionTree with features', array([ 3, 69]), 'has a score of', 0.9258409785932721)
('RandomForest with 50 estimators with features', array([ 3, 69]), 'has a score of', 0.9273700305810397)
('ExtraTrees with 50 estimators with features', array([ 3, 69]), 'has a score of', 0.9304281345565749)
('AdaBoost with 50 estimators with features', array([ 3, 69]), 'has a score of', 0.922782874617737)
```

위의 결과는 2개의 feature pair를 선택하고 전체 데이터 중 2개만을 이용하여 알고리즘을 학습시킨 training 정확도이다. 3개의 쌍을 추출하여 실험한 결과 [41,50]의 feature pair의 성능이 가장 높았지만, 선택할 수 있는 경우의 수가 (9,2)로 매우 높기 때문에 보다 많은 반복 실험이 필요하다. 또한 알고리즘 별 성능의 경우 3개의 pair 모두에서 ExtraTrees 알고리즘이 가장 높은 성능을 보였다.

3개의 feature pair로 학습된 tree가 어떠한 Decision Boundary를 갖고 있는지 위와 같이 시각화하여 쉽고 직관적으로 알고리즘을 분석할 수 있었다.<그림 9>

분석에 사용된 86개의 변수 중 고객 불량과 큰 상관관계를 보인 상위 9개 변수를 선정하여 <표 5>에 제시하였다.

<그림 9> 3종의 feature pair로 학습된 tree별 Decision Boundary



<표 5> 고객 불량과 상관관계를 갖는 상위 9개 변수

Rank	변수	설명	공정
1	69	P2-BIN_13	패키지 테스트
2	50	P1 YLD	패키지 테스트
3	47	FT1 - BIN_13	패키지 테스트
4	62	P2 YLD	패키지 테스트
5	57	P1 - BIN_13	패키지 테스트
6	1	ASY YLD	조립 공정
7	41	FT1 - BIN_7	패키지 테스트
8	4	COMDEF	조립 공정
9	5	INSFHE	조립 공정

조립(Assembly) 공정에서는 3개의 변수가 포함되었다. 모든 랏(Lot)들은 여러 번의 테스트 과정을 거쳐 조립 공정 발생된 불량을 제거하지만, 불량률이 높은 랏은 불량률이 낮은 랏에 비하여 공정 진행 인가되는 열적, 물리적 스트레스가 상대적으로 높기 때문에 고객 불량 발생 확률이 높았다. 조립 공정 불량 모드에서는 반도체 외관적인 불량은 포함되지 않았고, 반도체의 기능과 전기적인 특성을 좌우하는 단위 유닛별 반도체와 콤포넌트의 특성이 고객 불량 발생 가능성을 높이는 인자로 나타났다.

패키지 테스트(Package test) 공정에는 6개의 변수가 포함되었다. 자동차용 반도체는 다양한 사용 환경에 노출되는 것을 고려하여 상온(Room temperature), 고온(High temperature) 그리고 장시간 고온 상태에서 유지한 후 테스트를 진행하고 있다. 각 환경 조건에서는 자동차용 반도체의 기본 성능을 나타내며, 전통적으로 품질 및 수율 관리 인자로 사용하였던 전기 통전 능력이 고객 불량 발생률에 큰 영향을 나타내고 있다. 전기 통전 테스트 불량률이 높은 랏(Lot)가 불량률이 낮은 랏에 비하여 고객 불량 가능성이 높은 것으로 나타났다. 패키지 테스트 기능 검사 항목 중 기존 불량률에 의한 관리에서는 중요하게 여겨지지 않았던 특정 기능 검사 항목이 고객 불량을 많이 유발하는 것을 확인 할 수 있었다.

Ⅴ. 결론

1. 연구 요약 및 결론

자동차에 전자 컨트럴 장치의 사용량이 증가하고 있으며, 전자장치의 핵심 부품인 자동차용 반도체는 탑승자의 Safety와 직결되므로, 자동차용 반도체의 신뢰성에 대한 중요성이 증대 되고 있다. 자동차용 반도체의 신뢰성은 AEC 규격과 전통적인 신뢰성 시험 방법을 이용하여 평가 및 관리되고 있다. 평가 결과를 바탕으로 반도체 조립(Assembly), 패키지 테스트(Package test) 공정에서 공정 관리 기준으로 적용하고 있다. 이러한 방법은 반도체 사용 시간에 따른 제품 성능 저하를 사전에 검출하는데 적합하다. 하지만 반도체 제조 방법 및 기술 종류와 제품 출하 후 발생되는 다양한 고객사용 환경 및 시간에 따른 단기 및 장기 불량률을 예측하는데 한계가 있다.

본 연구에서는 조립(Assembly) 공정과 패키지 테스트(Package Test) 공정에서 계측된 데이터를 기반으로 데이터의 전처리, 데이터의 불균형 문제를 해결한 후 데이터마이닝 분류 알고리즘 중 랜덤포레스트 기법을 이용하여 잠재적인 고객 불량 발생을 예측 할 수 있는 모델을 제안하였다. 이렇게 구축된 예측 모델을 통하여 조립(Assembly) 및 패키지 테스트(Package) 공정 단계에서 제품의 품질 상태를 구분하고 추가적인 검사 항목을 도출함으로써 잠재적인 고객 불량 발생률을 줄이는데 기여 할 수 있을 것이다.

데이터 마이닝 분류 알고리즘을 이용하여 90% 정도의 시장 불량률을 사전에 검출해 낼 경우, 제품에 대한 고객 신뢰도 향상, 30년 이상의 제품 공급 기간 내 안전 재고 물량 관리 및 감소 효과, 고객 불량 발생시 원인 분석, 개선 사항 도출 및 잠재적인 불량 랏(Lot) 스크랩(Scrap), 선적된 불량 잠재적 불량 랏을 양품으로 교환하는 등 고객 불만 해결 과정에서 발생되는 품질 비용을 절감할 수 있다. 고객 불량 해결 시까지 사용되는 품질 비용은 약 $2000로, 자동차용 반도체의 수명 보증 기간 (15년 이상)을 고려하였을 때 각 제품별로 약 12억원 정도의 품질 비용 절감 효과를 기대 할 수 있다.

2. 한계점 및 향후 연구방향

본 연구를 통해 제안된 데이터마이닝 알고리즘 (랜덤포레스트)은 불량 예측 모델의 변수로 이용되는 데이터의 정합성에 따라 예측 모델의 성능이 결정되는 한계점을 가지

고 있다. 추후 과제로는 첫째, 반도체 제조 기술인 플립칩 본딩과 와이어 본딩을 이용한 자동차용 반도체 제품에 대한 실험 및 검증, 둘째, 가정용 및 산업용 반도체 등 반도체 전체에 대한 적용, 셋째, 웨이퍼 제조 공정과 프로브(Probe) 테스트 공정 간의 연관성 규명, 넷째, 반도체 조립(Assembly) 공정에서 각 단위 공정 별 측정된 공정 능력 지수 및 측정 값 들의 범위와 패키지 테스트(Package test) 공정에서 기능별 검사 데이터를 그룹화 하지 않고 반영한 모델을 구축하여 보다 예측 성능이 높은 데이터마이닝 분류 알고리즘 모델을 구축하는 것이다. 마지막으로 연구변수간 상호의존성 등의 인지를 위한 Feature engineering 이나 deep learning 기반 feature learning등을 통하여 데이터의 특성에 부합되는 다양한 예측모델의 추천 등으로 본 연구의 범위를 확장할 필요가 있다.

본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 대학ICT연구센터지원사업의 연구결과로 수행되었음 (2018-08-01417)

참고문헌

(1) 국내문헌

전황수(2010),"자동차-IT 융합 기술 및 국내외 개발 동향". 한국멀티미디어학회지 14(1), 1-10 (10 pages)

김경호(2011),"자동차용 반도체에 대한 완성차업체의 대응", 오토저널, 제 33권, 제 8호, 8월, PP 93~97

이승규, 허상현, 변영재(2012), "자동차용 반도체 시장 동향 및 기술 현황", 전자공학지 D.S Ahn(2012), "Consideration of Developing Automotive Semiconductor and Reliability Test". Bizocean, 2012 in Korea

도성룡, 한혁수(2014), "차량용 반도체의 품질 확보를 위한 사양 및 설계 개발 프로세스 수립", 정보과학회논문지 41(9), 625~632page

S. Y. Chae, W. J. Kim(2012), "Plan for reaction of Automotive ECU software and Soc for supporting ISO26262", Korea information processing society review, Vol. 19. No.09, P73-81

B.C. Kim(2012), "Korean Agency for Technology and Standards", ISO26262 Requirements of Designing Automotive Semiconductor

DH Baek, CH Han(2003), "Application of Data mining for improving and predicting yield in wafer fabrication system", Journal of Intelligence and Information System

안대웅, 고효헌, 백준걸, 김성식(2009), "단계적 SVM(Support Vector Machine)을 이용한 반도체 제조 공정에서의 최종 검사 공정 수율 예측 방법"

Shin K-Y., Ji J-G., Han J-H., Lee D-G., Son Y-J., and Lee H-S(2011), "Life Analysis and Reliability Prediction of Relays based on Life Prediction Method", Journal of the Korean society for railway, Page 1327-1335.

Yum, B-J., Seo S-K., Yun W-Y., and Byun J-H(2014), "Trends and Future Directions of Quality Control and Reliability Engineering", Journal of Korean institute of industrial engineers, 40(6), 526-554.

Lee S-H. and Chung H-J(2009), "Study on Lifetime Estimation of TFT-LCD Modules by Temperature Stress", Journal of Korean Institute of information, 7(5), 1-7.

Lee D. K(2005), "Reliability Prediction using Telcordia SR-332 in Electric Home Appliance", Journal of applied reliability, 5(4), 2005, 427-438.

Nam W.S, Kim S.B(2015), "A prediction of Wafer Yield Using Product Fabrication Virtual Metrology Process Parameters in Semiconductor Manufacturing", Journal of the Korean

Institute of Industrial Engineers, pp 1-7

JW Yoon, JW Kim, JM Koo and Sang-Su Ha(2007), "Flip-chip Bonding Technology and Reliability of Electronic Packaging", Journal of KWJS, Vol. 25, No. 2, April, 2007

(2) 국외문헌

Pieter(2000), "Pete B., 2000 begins with a revised industry" Roadmap", solid state technology, PP.31~44

Chris Webber(2013), "Automotive Semiconductor Demand Forecast 2011-2020", Strategy Analytics

AEC-Q004(2006), "Automotive Electronics Council Component Technical Committee", Zero Defects Guideline

R Uzsoy, CY Lee, LA Martin(1992), "A review of production planning and scheduling models in the semiconductor industry part I: system characteristics, performance evaluation and production planning", Vega - IIE transactions, Taylor & Francis

CF Chien, WC Wang, JC Cheng(2007), "Data mining for yield enhancement in semiconductor manufacturing and an empirical study", Expert Systems with Applications, Elsevier

Wilson, S., Joyce, T., and Lisay, E.(2009), "Reliability estimation from field return data", Life time data analysis, 15(3),, 397-410.

CK Shin, SC Park(2000), "A machine learning approach to yield management in semiconductor manufacturing", International Journal of Production Research, Taylor & Francis

TS Li, CL Huang, ZY Wu(2006), "Data mining using genetic programming for construction of a semiconductor manufacturing yield rate prediction system", Journal of Intelligent Manufacturing, Springer

D An, HH Ko, T Gulambar, J Kim(2009), "A semiconductor yields prediction using stepwise support vector machine", ISAM 2009. ieeexplore.ieee.org

C Chen, A Liaw, L Breiman(2004), "Using random forest to learn imbalanced data", University of California, Berkeley

P Kang, S Cho(2006), "Ensemble of under-sampled SVMs for data imbalance problems", Neural Information Processing, Springer

NV Chawla, KW Bowyer, LO Hall(2003), "SMOTE: synthetic minority over-sampling technique", Journal of artificial,

Quinlan, J. R(1986), "Induction of decision trees, Machine learning", 1(1), 81-106.

Vapnik, V.(2000), "The Nature of Statistical Learning Theory", Springer

Hosmer, D. and Lemeshow, W.(1989), "Applied Logistic Regression", Ed. John Wolfley and Sons, 8-20.

Rosenblatt, F.(1958), "The perceptron: a probabilistic model for information storage and organization in the brain, Psychological review", 65(6), 386.

Breiman, L.(2001), "Random forests, Machine learning", 45(1), 5-32.

Liaw, A. and Wiener, M.(2002), "Classification and Regression by random Forest", R news, 2(3), 18-22

P Kang, S Cho(2006), "EUS SVMs: Ensemble of under-sampled SVMs for data imbalance problems", Neural Information Processing,

Kubat, M., Holte, R. C., and Matwin, S.(1998), "Machine learning for the detection of oil spills in satellite radar images", Machine learning, 30(2/3), 195-215.

Kohavi, R.(1995), "A study of cross-validation and bootstrap for accuracy estimation and model selection", In IJCAI, 14(2), 1137-1145

J.H. Lau(2001), "Low Cost Flip Chip Technologies", McGRAW HILL BOOK Co., 1-17, 27-90

<붙임 1>. 변수 설명

번호	변수명	설명	번호	변수명	설명
1	ASY YLD	조립수율	44	10	패키지테스트1-불량모드4
2	OS QTY	조립공정중 전기 테스트 수량	45	11	패키지테스트1-불량모드5
3	OS RATE	조립공정중 전기 테스트 불량률	46	12	패키지테스트1-불량모드6
4	COMDEF	공정1-불량 모드1	47	13	패키지테스트1-불량모드7
5	INSFHE	공정2-불량모드1	48	26	패키지테스트1-불량모드8
6	EXCFIL	공정2-불량모드2	49	32	패키지테스트1-불량모드9
7	OFFDIS	공정2-불량모드3	50	P1 YLD	패키지테스트2-수율
8	INCCOV	공정2-불량모드4	51	7	패키지테스트2-불량모드1
9	BUBBLE	공정2-불량모드5	52	8	패키지테스트2-불량모드2
10	DIECRK	공정2-불량모드6	53	9	패키지테스트2-불량모드3
11	CHPOUT	공정2-불량모드7	54	10	패키지테스트2-불량모드4
12	NODIE	공정2-불량모드8	55	11	패키지테스트2-불량모드5
13	MALDIE	공정2-불량모드9	56	12	패키지테스트2-불량모드6
14	DENPIT	공정2-불량모드10	57	13	패키지테스트2-불량모드7
15	NONWET	공정2-불량모드11	58	14	패키지테스트2-불량모드8
16	H/SSC	공정3-불량모드1	59	15	패키지테스트2-불량모드9
17	H/SOR	공정3-불량모드2	60	25	패키지테스트2-불량모드10
18	H/SCU	공정3-불량모드3	61	26	패키지테스트2-불량모드11
19	H/SAL	공정3-불량모드4	62	P2 YLD	패키지테스트3-수율
20	MISBAL	공정4-불량모드1	63	7	패키지테스트3-불량모드1
21	BLADAM	공정4-불량모드2	64	8	패키지테스트3-불량모드2

22	BALBRG	공정4-불량모드3	65	9	패키지테스트3- 불량모드3
23	ALLBAL	공정4-불량모드4	66	10	패키지테스트3- 불량모드4
24	BALNEC	공정4-불량모드5	67	11	패키지테스트3- 불량모드5
25	F/M	공정5-불량모드1	68	12	패키지테스트3- 불량모드6
26	BURRS	공정5-불량모드2	69	13	패키지테스트3- 불량모드7
27	PCBCHP	공정5-불량모드3	70	26	패키지테스트3- 불량모드8
28	VOID	공정5-불량모드4	71	P3 YLD	패키지테스트4-수율
29	SCRTCH	공정5-불량모드5	72	8	패키지테스트4- 불량모드1
30	SQUOUT	공정5-불량모드6	73	9	패키지테스트4- 불량모드2
31	PCBDEF	공정5-불량모드7	74	10	패키지테스트4- 불량모드3
32	ILLMRK	공정5-불량모드8	75	11	패키지테스트4- 불량모드4
33	O/SDEF	공정5-불량모드9	76	12	패키지테스트4- 불량모드5
34	INLREJ	공정5-불량모드10	77	13	패키지테스트4- 불량모드6
35	PCBCHI	공정5-불량모드11	78	14	패키지테스트4- 불량모드7
36	PCBSCR	공정5-불량모드12	79	26	패키지테스트4- 불량모드8
37	COP	공정5-불량모드13	80	P4 YLD	패키지테스트5-수율
38	F/F	공정5-불량모드14	81	8	패키지테스트5- 불량모드1
39	DIECHI	공정5-불량모드15	82	9	패키지테스트5- 불량모드2
40	F1 YLD	패키지테스트1- 수율	83	10	패키지테스트5- 불량모드3
41	7	패키지테스트1- 불량모드1	84	11	패키지테스트5- 불량모드4
42	8	패키지테스트1- 불량모드2	85	12	패키지테스트5- 불량모드5
43	9	패키지테스트1- 불량모드3	86	13	패키지테스트5- 불량모드6