

# 분산분석에 의한 SPICE 심사의 신뢰성 검증 모델 설계의 사례연구

송기원<sup>0</sup>, 박정환, 이경환  
(Kiwon<sup>0</sup>, Jhpark Kwlee)@object.cau.ac.kr

## Case study for confidence verification model design of the SPICE assessment

Ki-Won Song<sup>0</sup>, Jeong-Whan Park, Kyung-Whan Lee  
Dept. of Computer Science and Engineering, Chung-Ang University

### 요 약

수준 높은 소프트웨어의 품질과 개발, 유지보수 비용의 최소화, 제품 출시시간의 단축을 위하여 소프트웨어 프로세스에 대한 예측, 통제 가능성을 증가시키기 위함이다. 기업이 최고도의 높은 수준에 도달하기 위해서는 정량적인 모델에 의한 프로젝트 관리가 필요하다. 따라서 기업들은 SPICE / CMM와 같은 표준을 사용하여 조직의 프로세스 능력 수준을 평가하고 수준향상을 꾀한다. 조직의 프로세스의 능력을 평가하고 수준향상을 위해서는 신뢰성 있는 SPICE 심사의 심사결과에 대한 객관적인 신뢰성의 보장과 좀더 적은 비용으로 프로세스의 수준향상을 할 수 있는 방법이 필요하다. 본 논문에서는 SPICE 심사의 신뢰성을 얻기 위해 CMM/KPA 설문서를 통해 심사하고 SPICE심사의 결과를 비교 분석하여 SPICE심사의 신뢰성을 검증한다. 또한 이를 기반으로 CMM/KPA 설문서의 정량적인 모델을 제안함으로써 좀더 적은 비용과 시간으로 SPICE 심사의 결과와 같은 효과를 얻을 수 있게 한다.

### 1. 서 론

기업의 경쟁력을 향상시키기 위해서 제품을 개발하고 유지보수 하는데 소요되는 비용을 최소화하고, 제품의 출시시간을 단축시키며 제품을 만드는 프로세스에 대한 예측과 통제 가능성을 목표로 두고 있다[1][4].

이러한 목표를 성취하기 위해 조직에서는 프로세스 개선의 요구가 증대되고 조직들은 많은 노력을 들이고 있다.

SPICE 심사를 통해 나온 조직의 수준(Level)은 심사원들이 프로세스안의 작업을 통해 나온 각각의 산출물에서 증거(evidence)의 도출을 통해 등급(rating)을 정하게 된다. [2]

수준3 까지는 소프트웨어 프로세스 심사 (SPA : Software Process Assessment)로 나온 개선사항만 가지고 기존의 프로세스 개선 절차(SPI : Software Process Improvement)를 이용하여 수준향상에 큰 어려움이 없으나 수준4(Predictable), 수준5(Optimizing)을 달성하기 위해서는 기존의 SPICE의 SPI 만으로는 어려움이 있다[1].

이를 보완하기 위해서는 정량적인 모델을 통해 어느 프로세스가 가장 큰 개선사항인가 하는 우선순위가 결정되어야 한다.

여기서 조직은 적지 않은 시간과 자금을 투자해 나온 심사 결과가 얼마나 신뢰할 수 있는지와 도출된 개선 사항 중 어느 개선사항이 더 중요한 요인(factor)인가를 알고자 한다.

본 논문에서는 기존의 SPICE 심사와 CMM/KPA 설문서의 결과를 비교 분석하여 SPICE 심사의 신뢰성을 검증하고 중요한 개선사항을 도출할 수 있는 검증모델을 제안하고자 한다.

이를 위해 관련 연구로서 SPICE의 심사절차를 기술하고 기반 연구로서 기존연구에서 신뢰성이 입증된 CMM/KPA 설문서를 통해 SPICE심사의 신뢰성 검증을 위한 SPICE 심사 검증모델을 제안하고 사례연구를 통한 신뢰성을 평가하고 결론을 맺는다.

### 2. 관련 연구

#### 2.1 SPICE 심사의 rating

SPICE는 ISO 12207에 정의된 소프트웨어 생명주기 프로세스 전체 또는 각각에 대하여 독립적인 심사와 개선 활동을 위한 지침을 제공하고 있다.

SPICE 심사의 rating은 심사대상 프로세스마다 각각의 심사원의 미니팀으로 구성하여 문서심사와 인터뷰심사를 통해 심사대상프로세스의 능력을 결정한다.[2]

Rating은 F(Fully),L(Largely),P(Partially),N(None)의 4단계로 평가를 하게 되며 이때 각각 프로세스의 성취도 비율에 따라 <표 1>과 같이 평가한다[2][3].

<표 1> SPICE심사의 일반적인 rating 척도

Rating Notation	F	L	P	N
프로세스 성취도	100~86%	85~51%	50~16%	15~0%

#### 2.2 KPA(KPA의 Maturity Questionnaire)의 신뢰성 검증

KPA(Key Process Area)는 CMM에 의한 프로세스 평가를 하기 위한 기준이다. KPA는 총 18단계의 핵심공정 영역으로 이루어져 있으며 각각의 핵심공정에는 약 6~8개의 Key practice questionnaire가 존재한다[4].

예로, 18개 KPA항목 중 13번째 KPA인 peer review는 4개의 설문서로 구성되어 있다. SEI가 제시한 항목 이외에도 설문 항목을 추가시킬 수 있지만 4개의 설문서를 일관성 있게 rating하는 것은 쉬운 일이 아니다.

기존 연구를 위한 data collection은 Pareto rule에 의해서 4개의 설문서만 선정하였다.

Pareto rule을 적용한 결과는 <그림 1>에 나타나있다[9].

항목	1	2	3	4	5	6	7	8	9	10	11
1	25	25	20	30	24	25	20	20	30	25	20
2	20	15	20	15	20	20	24	20	15	20	15
3	20	15	15	15	20	10	20	24	15	20	18
4	15	15	15	10	13	15	10	12	13	10	13
5	12	10	10	10	11	10	8	9	10	10	10
6	3	12	10	15	5	3	8	10	11	7	9
7	3	8	10	5	13	10	6	8	8	11	7
8	100	100	100	100	100	100	100	100	100	100	100

<그림 1> Pareto rule에 의한 peer review의 Maturity Questionnaire 선정표

기존 KPA설문서는 CMM KPA의 18개항목의 Goal을 전반적인항목으로 두고 Key Practices를 총 104개를 세부항목으로 두었다. <그림 1>은 세부 항목을 SPICE 심사원에게 설문 후 가장 중요도가 높게 선택받은 항목이다. SPICE 심사의 내부 심사원과 외부심사원이 이와 같이 작성된 KPA심사 설문서에 응답한다. 그리고 이를 <그림 2>와 같은 COCOMOII의 EPML계산식을 통해 rating을 하게 된다.

$$EPML = 5 \times \left( \sum_{i=1}^n \frac{KPA\%_i}{100} \right) \times \frac{1}{n}$$

<그림 2> EPML 계산식

우리는 1년간 30여과제의 SPICE심사에서 기존의 KPA 설문서를 통해 4개사에 수집한 KPA 자료를 통해 KPA EPML을 확인한 결과 <표 5>의 기존 KPA EPML과 같은 결과를 얻었다.

우리는 2003년 6월 샌프란시스코에서 개최된 SERA'03 (Software Engineering Research Association)에 SEI의

Maturity Questionnaire 보정모델의 신뢰성의 검증을 위한 가설검증과 보정모델 및 SPICE 수준 변환모델에 대한 논문을 발표했다.[6] 이 연구결과로 KPA 보정모델이 99.8%의 신뢰성을 가지고 <표 5>에서와 같이 수치적으로 정량화한 SPICE 변환 EPML이 실제 직관적으로 보이는 SPICE rating Level과 일치함을 증명하였다[6][7].

<표 5> 보정 KPA EPML과 SPICE EPML의 비교

	A사	B사	C사	D사
기존 KPA EPML	3.38	3.77	3.19	3.67
보정후 KPA EPML	3.4	3.23	3.61	3.59
SPICE 변환 EPML	3.43	3.17	3.71	3.51

2.3 Software Reading 기법

S/W Reading 기법은 Basili 교수가 S/W의 검증을 위한 방법으로 개발한 기술로서 작업산출물에 대해서 수행할 수 있는 개별적인 분석기법이다.[5]

속성 분류법에 따라서 문제점을 특성화시켜서 High level goal, Specific goal, Document, notation, form, family, technique 등으로 나누어서 분석한다.[5] 속성 분류법은 <표 6>과 같다.

<표 6> Families of Reading technique

가법	High level goal	Specific goal	Document	Notation/Form	Family	Technique
Reading	Construction	제사용	Test Plan	...	...	...
			Code library	...	...	...
			Code	White box	Scope based	System wide Task oriented
			...	...	...	...
			...	...	...	...
	Analysis	Fault detection	Design	...	...	...
			Req.	SCR	Defect based	inconsistent Incorrect omission ambiguity
			...	...	...	...
			English	Perspective based	Tester User Developer	
			Traceability	...	...	...
Performance	...	...	...			

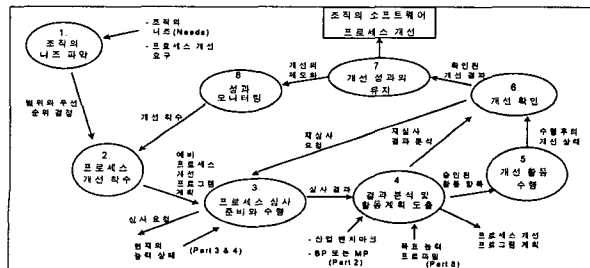
3. SPICE 심사의 신뢰성 검증을 위한 검증모델

우리는 약 1년 6개월간 수집한 CMM/KPA 데이터를 가지고 KPA 설문서의 신뢰성을 99.8%까지 보장할 수 있는 보정모델을 기존 연구에서 제안하였다.

하지만, 이는 CMM/KPA 설문서 자체의 신뢰성만 입증할 뿐 SPICE 심사의 신뢰성은 보장하지 않는다. 따라서 본 논문에서는 SPICE 심사 신뢰성 검증을 위한 CMM/KPA rating 정량적 모델을 제안한다.

3.1 SPICE 심사 결과 분석

SPICE의 심사는 <그림 5>와 같은 프로세스로 구성된다.



<그림 5> SPICE의 기본 프로세스 개선절차

우리는 <그림 5>의 4.결과분석 및 활동계획 도출의 전 단계의 SPICE 심사 최종 결과보고서를 분석하였다. 2003년 심사를 실시한 E사의 SPICE 심사 최종결과보고서의 분석결과를 <표 7>과 같다.

<표 7> E사 SPICE 심사 개선사항 분석

프로세스	개선사항
일반적 개선사항	- 과제 유형 분석을 통한 개발 방법론 정립이 필요함 - 소규모 단기간 프로젝트의 Process 관리를 위한 Pattern 개발이 필요함. - Process 관리 지표 개선을 위한 지속적인 활동이 필요함 - 프로젝트 특성에 준해서 체계적, 조직적 Agile 방법의 도입 계획의 수립이 필요함

<표 6>의 개선사항은 심사원들간의 consolidation을 통한 개선사항을 분석한 것이다. 그러나 이러한 심사결과가 얼마나 신뢰할 수 있는지는 알 수가 없었다. 또한 어느 개선사항이 중요한 요인(factor)인지도 알 수가 없었다. 따라서 우리는 기존연구에서 99.8%의 신뢰성이 보장된 CMM/KPA의 rating 결과와 비교 분석 함으로써 SPICE 심사의 신뢰성을 검증하고자 한다.

3.2 CMM/KPA rating 분석

CMM/KPA rating은 SPICE 심사시 동일한 심사원들에게 CMM/KPA 설문서를 통해 얻은 데이터를 우리가 개발한 CMM/KPA Tool을 이용하여 <그림 6>과 같은 결과를 얻었다.

<그림 6> CMM/KPA rating 분석결과

<그림 6>을 이용한 CMM/KPA rating의 분산을 통해서 본 수준 평가 결과는 <표 8>과 같은 결과가 나왔다.

<표 8> CMM/KPA rating의 분산을 Reading을 통해서 본 수준평가

KPA 일반항목	분산이 큰 항목 : 심사원들의 평가의견이 차이가 많음	
	소프트웨어 프로젝트의 추적과 감독	3번
KPA 일반항목	형상관리	6번
	진척도 변경관리	18번
	메이스라인에 준한 표준보고서가 존재하는가?	6-4번
KPA 세부항목	Defect의 양적관리(수출, 제거, 비용추정)을 하고있는가?	16-1번
	분산 항목이 큰 세부항목을 가진 KPA 일반항목 (세부항목이 3번이상 나타난 KPA 일반항목)	
KPA 일반항목	형상관리 (6번 항목)	3회
	Defect 예방 (16번 항목)	3회
	품질관리 (15번)	4회

CMM/KPA Tool을 이용한 EPML 수준은 실제 SPICE 심사결과와 유사한 결과를 보였다. 하지만 이러한 수치가 실제 SPICE 심사 결과의 개선사항과 일치하는가 하는 문제가 남아있다.

따라서 우리는 CMM/KPA rating의 표준편차의 분산분석을 통해 KPA Defect(편차가 심하게 나는 것은 심사원들간의 의견이 다른 것이다. 이는 프로세스의 정립이 올바르지 않다는 것이다. 따라서 우리는 이를 KPA Defect라 부른다.)를 찾아내기 위한 정량적인 모델을 제안한다.

4. CMM / KPA rating의 분산분석의 정량모델

지금까지 CMM/KPA의 심사원의 편차를 Reading 기법을 통해 보통 분석해 왔다. Reading 기법보다 더 신뢰성 있는 방법을 제안하고자 한다. Minitab의 NP관리도를 이용하여 CMM/KPA의 심사원간 편차를 정량적으로 관리하는 모델을 제안하였다. NP관리도란 불량 갯수(KPA Defect) 관리도로 공정(Process)을 불량 갯수 PN에 의거하여 관리할 경우에 사용한다. 여기서 입력되어야 할 값은 결합데이터(심사원들의 편차, 그룹의 사이즈(최소 편차값과 최대편차값의 합), P값(기대 신뢰구간 값)을 입력하게 된다. 입력데이터는 다음과 같다.

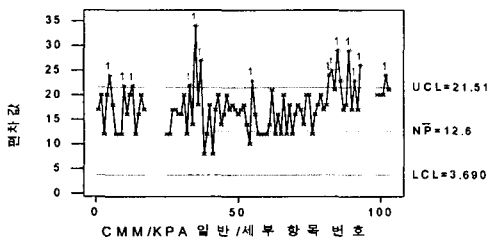
불량 데이터 : 총 103개 항목 심사원들의 편차  
 그룹사이즈 : 최소 편차(8) + 최대편차(34) = 42  
 신뢰구간 P : 0.3(유의수준 ±0.15 : 85%의 신뢰구간)

8사의 경우 CMM/KPA 심사원의 편차를 입력한 결과는 <그림 7>과 같다.

NP Chart: A사 CMM/KPA 심사원 편차

TEST 1. One point more than 3.00 sigmas from center line.  
 Test Failed at points: 5 10 13 33 35 37 55 82 83 85 86 89 91 93 102

NP Chart for A사 CMM/KPA 심사원 편차



<그림 7> NP관리도를 통한 분산분석

위의 결과를 분석한 결과 관리 지표는 최대 수용할 수 있는 편차의 값(UCL)은 21.51개 평균 편차의 값은 NP는 12.6개, 최대 하한 편차의 값(LCL)은 3.69개로 이것이 의미하는 것은 심사원의 편차의 평균이 12.6개 최대 ±1.5개의 값 안에 있으면 85%의 신뢰성을 인정할 수 있다. 실제 심사원의 평균의 편차는 11.74개로 85% 신뢰구간 안으로 들어왔다.

또한 심사원의 편차가 3.69개가 되지 않는 항목은 심사원들의 설문에 응한 성실도가 떨어지는 항목이라 볼 수 있으며 21.51개가 넘는 항목은 KPA Defect라고 볼 수 있다.

<표 9> 분산분석 통해 도출된 KPA Defect

NP num	KPA 설문항목
5	발달된 요구사항을 관리하는 활동의 상태를 점검하기 위해 사용될 수 있는 측정치가 존재하는가?
10	소프트웨어 프로젝트 계획(Chtr)을 위해 적절한 자원이 배정되어 있는가?
13	실행결과(스케줄, 사이즈, 비용)와 성능이 소프트웨어의 계획과 비교해서 추적되는가?
33	소프트웨어의 산출물이 항상관리 활동에 의해 정의되고, 조정되며, 변경관리기 이루어지고 있는가?
35	소프트웨어 케이스라인에 대한 표준 보고서가 존재하는가?
37	소프트웨어 케이스라인을 검증하기 위한 주기적인 감사활동이 존재하는가?
55	훈련과정이 조직적으로 운영되는가?
82	프로젝트에서 정의된 프로세스의 성능을 측정하고 양적으로 관리한다.
83	조직의 표준 프로세스의 능력을 분석할 수 있는 기록된 전략이 갖추어져 있다.
85	프로젝트의 양적인 프로세스 관리를 프로젝트의 프로세스 관리 계획에 의해서 수행한다.
86	프로세스에 관련된 데이터를 수집, 조정(Calibration) 및 관리한다.
89	Defect의 양적인 관리(추출, 제거, 비용추정)를 하고있는가?
91	Defect의 공통된 원인들의 우선순위가 정해지고 체계적으로 제거되는가?
93	Defect 예방 작업에서 얻은 결과를 조직의 표준 프로세스의 Revision에 적용하는가?
102	SP의 제안서가 문서화된 절차에 의해서 처리되는가?

<표 9>의 결과는 심사원들간의 보여지는 프로세스가 명확하지 않아 발생하는 것이다.

<표 10>은 SPICE 심사 및 CMM/KPA rating 결과의 mapping 분석 결과이다.

이는 SPICE 심사원들이 수행한 rating과 CMM/KPA 설문서 응답 내용을 텍스트로 하여 Reading기법에 의해서 분석한 표이다.

SPICE 심사 보고서는 일반적인 개선사항과 심사한 프로세스별 개선사항을 포함한다.

본 논문에서는 보고서중에서 프로세스별 개선 사항을 나열하고 (본문에서는 생략함) CMM/KPA의 15개의 일반항목과 세부 항목을 대응시켜 놓은 표이다.

<표 9>의 KPA Defect를 개선항목으로 선정한 결과 <표 10>과 같이 15개의 KPA Defect가 실제 SPICE 심사의 개선사항과 일치하였다.

<표 10> SPICE 심사 및 CMM/KPA rating 결과의 mapping 분석 결과

일반	항목내용	세부	항목내용	NP num
15	소프트웨어 품질관리	2	조직의 표준 프로세스의 능력을 분석할 수 있는 기록된 전략이 갖추어져 있다.	83
9	훈련 프로그램	4	훈련과정이 조직적으로 운영되는가?	55
16	Defect 예방	5	Defect 예방 작업에서 얻은 결과를 조직의 표준 프로세스의 Revision에 적용하는가?	93
15	소프트웨어 품질관리	5	프로세스에 관련된 데이터를 수집, 조정(Calibration) 및 관리한다.	86
16	Defect 예방	13	1: Defect의 양적인 관리(추출, 제거, 비용추정)를 하고있는가? or 3: Defect의 공통된 원인들의 우선순위가 정해지고 체계적으로 제거되는가?	91
15	소프트웨어 품질관리	5	프로세스에 관련된 데이터를 수집, 조정(Calibration) 및 관리한다.	86
15	소프트웨어 품질관리	2	조직의 표준 프로세스의 능력을 분석할 수 있는 기록된 전략이 갖추어져 있다.	83
6	소프트웨어 형성관리	2	소프트웨어의 산출물이 항상관리 활동에 의해 정의되고, 조정되며, 변경관리기 이루어지고 있는가?	33
6	소프트웨어 형성관리	4	소프트웨어 케이스라인에 대한 표준 보고서가 존재하는가?	35
18	진척도 변경관리	3	SP의 제안서가 문서화된 절차에 의해서 처리되는가?	102
1	요구사항관리	4	발달된 요구사항을 관리하는 활동의 상태를 점검하기 위해 사용될 수 있는 측정치가 존재하는가?	5
6	소프트웨어 형성관리	6	소프트웨어 케이스라인을 검증하기 위한 주기적인 감사활동이 존재하는가?	37
15	소프트웨어 품질관리	2	조직의 표준 프로세스의 능력을 분석할 수 있는 기록된 전략이 갖추어져 있다.	83
2	프로젝트 관리	3	소프트웨어 프로젝트 계획(Chtr)을 위해 적절한 자원이 배정되어 있는가?	10
3	소프트웨어 품질관리	1	실행결과(스케줄, 사이즈, 비용)와 성능이 소프트웨어의 계획과 비교해서 추적되는가?	13
1	요구사항 관리	4	발달된 요구사항을 관리하는 활동의 상태를 점검하기 위해 사용될 수 있는 측정치가 존재하는가?	5
15	소프트웨어 품질관리	4	프로젝트에서 정의된 프로세스의 성능을 측정하고 양적으로 관리한다.	82
15	소프트웨어 품질관리	5	프로세스에 관련된 데이터를 수집, 조정(Calibration) 및 관리한다.	86
15	소프트웨어 품질관리	1	프로젝트에서 정의된 프로세스의 성능을 측정하고 양적으로 관리한다.	82
8	조직의 프로세스 정의	2	조직의 표준 프로세스의 능력을 분석할 수 있는 기록된 전략이 갖추어져 있다.	83
16	Defect 예방	3	Defect의 공통된 원인들의 우선순위가 정해지고 체계적으로 제거되는가?	91
15	소프트웨어 품질관리	5	프로세스에 관련된 데이터를 수집, 조정(Calibration) 및 관리한다.	86
15	소프트웨어 품질관리	2	조직의 표준 프로세스의 능력을 분석할 수 있는 기록된 전략이 갖추어져 있다.	83
16	Defect 예방	3	Defect의 공통된 원인들의 우선순위가 정해지고 체계적으로 제거되는가?	91

5. 결론 및 향후 연구방향

본 논문에서는 SPICE 심사의 신뢰성을 검증하기 위해 기존 연구에서 신뢰성이 검증된 CMM/KPA를 통한 분산분석 모델을 제안하였다.

이는 심사원의 CMM/KPA 항목별 rating 결과의 Reading에 의한 결과와 정량적 분석 모델에 의한 신뢰도와 지표도출을 통해 Reading의 결과와 정량모델에 의한 신뢰성 검증이 일치함으로써 보여줌으로서 실제 SPICE 심사의 신뢰성을 85%까지 신뢰할 수 있음을 보여준다. 또한 지표통한 KPA Defect를 도출하여 SPICE 심사와 동일한 개선사항을 도출할 수 있으며 편차가 크고 자주 나타나는 항목일수록 프로세스의 정립이 되지 않음을 알 수 있다.

이를 통해 CMM/KPA의 편차를 통해 개선사항의 우선순위를 검출할 수 있으며 CMM/KPA의 설문문을 통해 좀더 적은 시간과 비용의 투자로 효율적인 프로세스의 개선을 할 수 있다.

향후 연구과제로서 좀더 많은 영역의 데이터를 사례연구를 통해 검증하고 이를 통해 Opportunity Tree를 설계하는 초석이 되는 CMM/KPA의 C-E 다이어그램(Cause Effect)을 연구하고자 한다.

참고문헌

[1] Kyung-Whan Lee, "Research for HDC Modeling", Proceedings of the 5th Korean Conference on Software Engineering, 2003.02.20 22  
 [2] ISO/IEC JTC1/SC7 15504: Information Technology- Software Process Assessment, ISO TR, ver.3.3, 1998  
 [3] KSPICE (Korea Association of Software process Assessors), SPICE Assessment Report <http://kasper.org>, 2002~2003  
 [4] Geir Amsjo, Capability Maturity Model(CMM) for Software Development, <http://www.ifi.uio.no/in331/foiler/SW-CMM.pdf>  
 [5] Basili V. R., Shull F., Rus I., and Laitenberger O., "Improving Software Inspections by Using Reading Techniques", In Proceedings of The International Conference on Software Engineering (ICSE), Tutorial, June 2000  
 [6] Ki-Won Song, "Research about confidence verification of KPA question item through SEI Maturity Questionnaire's calibration and SPICE Level metathesis modeling", SERA03, San Francisco, 2003.06  
 [7] Woo-Song Kim, Reliability Test of Maturity Questionnaire Selection Model Through KPA Rating Data Calibration, Conference proceedings, KISS, Jeju, Korea, , Apr, 25-2