

# ISO/IEC 9126-2 메트릭을 활용한 소프트웨어 정량적 평가 지표 도출

조성호<sup>1†</sup> · 장중순<sup>2</sup>

<sup>1</sup>한국산업기술시험원, <sup>2</sup>아주대학교 산업공학과

## Quantitative Evaluation Index Derivation of the Software Based on ISO/IEC 9126-2 Metrics

Sungho Cho<sup>1†</sup> · Joongsoon Jang<sup>2</sup>

<sup>1</sup>Korea Testing Laboratory, <sup>2</sup>Graduate School of Ajou University

**Purpose:** Many domestic companies have to make out quantitative evaluation table in their proposal when they conduct the software R&D project. However, most of companies have a difficulty to select the evaluation items and criteria, also to derive a quantitative results. Therefore, we propose a method to derive the quantitative evaluation index by utilizing the ISO/IEC 9126-2.

**Methods:** Analyzing ISO/IEC 9126-2, and we classify the quality metrics as high-classification and sub-classification for Web/App software, Embedded software and Installation software. Next, Conduct the metrics selection survey depending on importance and necessity. Then, carry out the case study. Verify the correspondence between evaluation items and criteria from original suggestion of company and from outcome by utilizing the ISO/IEC 9126-2 quality metrics.

**Results:** It is possible to classify into two metrics, one for common software or one another for only special software. Furthermore, there is quality metrics that is more important and more necessary depending upon characteristics of the software.

**Conclusion:** ISO/IEC 9126-2 quality metrics can be used to make an evaluation items and criteria for quantitative evaluation table of software product.

**Keywords:** ISO/IEC 9126-2, Web/App Software, Embedded Software, Installation Software, Software Quality Metrics, Quantitative Evaluation Metrics

### 1. 서론

IT 분야의 지속적 경쟁력 확보를 위한 소프트웨어 R&D 사업 및 과제가 매년 추진되고 진행된다. 소프트웨어 개발을 계획하고 준비하는 많은 기업들은 국내 정부기관 및 공공기관에서 공고한 R&D 사업 및

과제를 참여한다. 참여하는 기업은 R&D 사업 및 과제의 요구사항을 파악하여 소프트웨어 제품에 반영할 수 있도록 해야 하며 발주기관 및 수요업체는 참여 기업의 소프트웨어가 사용 목적과 용도에 적합한지 확인해야 한다[1].

최근의 소프트웨어를 개발하는 과정은 매우 복잡

† 교신저자 shcho@ktl.re.kr

2016년 5월 30일 접수, 2016년 6월 10일 수정본 접수, 2016년 6월 13일 게재 확정.

하여 개발 과정 동안 요구사항에 부합되지 않은 기능이 추가되기도 하고, 예측하지 못한 결과가 나타나기도 한다. 따라서 소프트웨어 품질을 측정하고 평가하는 일이 중요한 문제로 대두되어지고 소프트웨어 검증에 대한 부분이 필수적이 되었다. 요구 사항이 소프트웨어에 반영되어 있는지 여부 및 소프트웨어가 사용 목적 및 용도에 적합한 품질을 보증하는지 여부 등을 객관적으로 확인하기 위하여 과제 참여 시 기업들은 제안서에 성과목표 달성도를 정량적으로 측정하는 잣대인 성과지표를 기입하도록 요구받고 있다[2].

현재 소프트웨어 R&D 분야 사업 및 과제의 제안서에 작성하도록 요구하는 항목 가운데 정량적 평가 지표 표가 있다. 이 표에 기입해야 하는 항목으로는 주요 성능지표, 단위, 최종 개발목표, 세계최고수준, 가중치, 객관적 측정방법이 있다. 많은 기업이 주요 성능지표 항목을 잘못 선정하여 소프트웨어의 정량적 결과 도출에 어려움을 겪는다. 또한 항목에 대한 상세한 기준 및 절차를 고려하지 않고 성능지표 항목에 대한 객관적 측정방법(시험 규격) 작성이 어려움을 겪고 있다.

현재 국내 유일의 소프트웨어 품질 인증 제도인 GS 인증은 소프트웨어 제품 품질 특성과 품질 평가의 메트릭을 정의하는 ISO/IEC 9126-2 국제 표준을 기반으로 한다. 이 표준은 제품이 갖추어야 할 품질 특성으로 기능성, 신뢰성, 사용성, 효율성, 유지보수성, 이식성을 정의한다[3]. 그리고 품질 특성 별로 품질 메트릭을 나열하고 메트릭의 목적과 적용 방법을 기술하고 계산 방법에 대하여 수식으로 표현하며 해석 방법

에 대해 상세히 서술되어 있다. 따라서 모든 품질 메트릭 항목이 모두 정량적 결과 도출이 가능하게 작성되어 있다[4].

그러나 ISO/IEC 9126-2 표준은 모든 소프트웨어를 대상으로 적용되는 메트릭을 나열하고 있다. 소프트웨어 제품의 특성이나 제품의 분류에 따른 메트릭을 구분하지 않고 있어 다양한 소프트웨어 제품별로 모든 품질 메트릭을 모두 적용하기는 현실적으로 불가능하다. 따라서 소프트웨어 제품 특성에 따라 메트릭을 선택적으로 적용할 필요가 있다.

본 연구에서는 소프트웨어의 정량적 성능지표 선정 및 기준 및 측정 방법 등을 결정하는데 있어 ISO/IEC 9126-2 품질 메트릭을 활용하는 것을 제안한다. ISO/IEC 9126-2의 품질 메트릭을 활용하는데 있어 3가지 소프트웨어 제품(웹/앱 소프트웨어, 임베디드 소프트웨어, 설치형 소프트웨어) 특성별로 메트릭을 상위 분류와 하위분류로 구분하여 제시한다. 이를 통하여 현실성을 반영한 소프트웨어의 정량적 성능 지표 항목 선정과 객관적인 시험 기준 및 측정 방법 작성이 가능하다.

## 2. ISO/IEC 9126과 소프트웨어 국내 평가 기준

### 2.1 ISO/IEC 9126

소프트웨어 품질보증 분야에 대한 국제표준화 작업은 ISO/IEC JTC1과 ITU-T에서 주로 이루어진다. 이 중 ISO/IEC JTC1에서 추진 중인 소프트웨어 품질

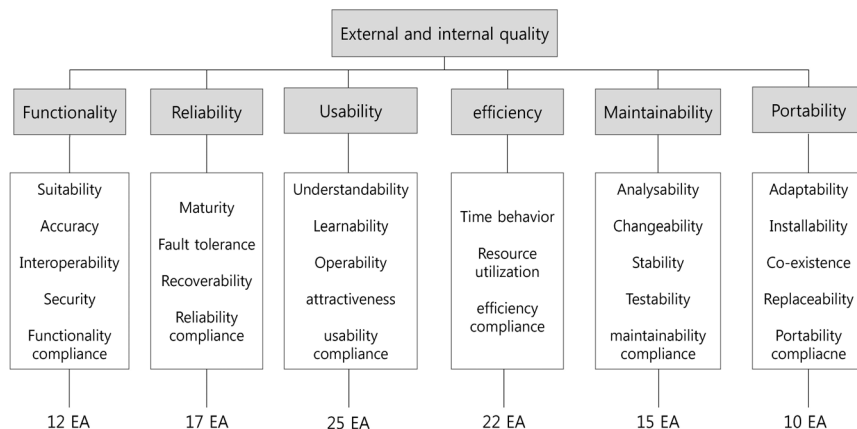


Fig. 1 ISO/IEC 9126 quality model

보증 분야에 대한 국제표준은 ISO/IEC 9126, ISO/IEC 14598, ISO/IEC 25051 등이 있다[5].

이 중, ISO/IEC 9126은 소프트웨어 전체 품질의 정량적 측정을 위해 제정된 국제 표준이며 소프트웨어 제품의 품질 특성을 정의하고 품질 평가의 메트릭을 정의한다. 해당 표준은 총 4가지의 시리즈로 구성되어 있다. ISO/IEC 9126-1은 소프트웨어 품질 모델을 정의한다. ISO/IEC 9126-2, 3은 개발 수명주기 동안에 소프트웨어가 실행될 수 없을 때와 시스템 환경에서 소프트웨어가 실행할 수 있을 때 제품에 대한 품질을 측정할 수 있는 외부 메트릭과 내부 메트릭을 설명한다. ISO/IEC 9126-4는 생산성, 안전성, 만족도 등과 같이 명시된 사용자의 요구를 만족하는지 여부를 측정하는 사용 품질 메트릭을 설명한다. ISO/IEC 9126에서는 소프트웨어 품질 모델을 기능성, 신뢰성, 사용성, 효율성, 유지보수성, 이식성으로 제시한다. 그리고 각각의 품질 특성은 아래 그림과 같이 부특성과 품질평가 지표로 구성된다[6].

앞에서 설명한 ISO/IEC 9126 표준에서 ISO/IEC 9126-2는 소프트웨어 외부 메트릭에 대하여 정의하고 있다. <Fig. 1>에서 보여준 각 품질 주특성과 부특성에 대한 외부 메트릭(품질평가지표)의 적절한 직접 및 간접 측정 방법을 제안한다. 외부 메트릭이란 소프트웨어 제품을 한 부분으로 사용하는 시스템의 행위를

측정함으로써 소프트웨어 제품의 품질을 측정하는데 사용하는 것을 말한다. ISO/IEC 9126-2는 ISO/IEC 9126-1에서 정의한 품질 모델 중 기능성 12개, 신뢰성 17개, 사용성 25개, 효율성 22개, 유지보수성 15개, 이식성 10개를 나열한다. 각각의 메트릭은 항목별 적용방법, 계산방법, 해석방법을 상세히 기술하고 있다[4, 7].

### 2.2 소프트웨어 국내 평가 기준

국내에서 제시하는 소프트웨어를 정량적으로 평가하기 위한 기준 중 3가지를 소개하면 다음과 같다. 공공부문 소프트웨어 사업의 상용 소프트웨어 및 정보시스템에 적용되는 기준, 그리고 소프트웨어 공인 시험 기관인 한국정보통신기술협회에서 발간한 웹 소프트웨어에 대한 평가 기준, 국내 무기체계 임베디드 소프트웨어 개발 기관인 LIG넥스원에서 제시한 대한 평가 기준을 나열한다.

첫째, 소프트웨어 기술성 평가 기준 적용 가이드이다. 여기서는 상용 소프트웨어 기술성 평가항목별 세부 내용을 명시하고 있다. 이 중 평가 부문은 크게 기능성, 사용성, 이식성, 효율성, 유지보수성, 신뢰성, 공급업체 지원으로 구분하고 있다. 평가 부문에 따른 평가 항목은 <Table 1>과 같다[8].

총 7개 부문 28개 평가 항목으로 공급업체 지원 평

Table 1 Cots software test sector and test item

| Test sector           | Test item   | Test sector     | Test item                               |
|-----------------------|---|-----------------|---|
| Functionality         | Functional implementation completeness                | Efficiency      | Response time                           |
|                       | Functional adequacy                                   |                 | Resources utilization                   |
|                       | Data exchangeability                                  |                 | Throughput                              |
|                       | Security  | Maintainability | Support error diagnosis/resolution      |
|                       | Standard compliance                                   |                 | Possibility to change the configuration |
| Usability             | Function understandability                            | Reliability     | Ease of update                          |
|                       | Understandable input and output                       |                 | Ease of backup/recovery                 |
|                       | Possibility of adjustment of the user interface       |                 | Stability of operation                  |
|                       | Consistency of the user interface                     |                 | Ease of failure recovery                |
|                       | Ease to understand the progress                       |                 | Persistence of service                  |
| Portability           | Possibility of the adjustment of operating procedures |                 | Data recovery                           |
|                       | Operating environment compatibility                   | Vendor support  | Support the maintenance                 |
|                       | Ease of installation                                  |                 | Support the education and training      |
| Continued use of data | Product reliability                                   |                 |   |

가 부문을 제외한 6개 부문이 ISO/IEC 9126에서 제시한 주특성과 동일하다. 그리고 공급업체 지원 부문을 제외한 25개의 평가항목은 ISO/IEC 9126-2에서 설명하는 품질 메트릭의 일부를 선택하여 사용하고 있다. 하지만 이 평가 기준 역시 모든 소프트웨어를 대상으로 하고 제품 특성에 따른 분류를 하고 있지 않다.

다음으로 한국정보통신기술협회에서 발간한 소프트웨어 품질평가 프레임워크 개발에 관한 연구보고서에 기재된 웹 기반 소프트웨어 품질평가 지침이다 웹 기반 소프트웨어 품질 평가 지침에서는 웹 기반 소프트웨어 품질평가 모델을 아래의 그림으로 제시한다[9].

<Fig. 1>과 <Fig. 2>를 비교해보면 알 수 있듯이, 기

존의 ISO/IEC 9126-2 품질 모델에서 기능성의 부특성에 있는 적합성, 정확도를 사용성으로 변경하고 이식성 품질 모델을 삭제하는 등 ISO/IEC 9126-2 품질 모델을 변경 및 삭제하여 웹 기반 소프트웨어 평가 모델을 제시한다. 그리고 이 기준의 많은 메트릭이 제품 설명서나 사용자 문서와 제품 설명서와 같은 문서와의 비교를 제시하고 있다.

세 번째는 임베디드 소프트웨어를 대표하는 무기체계 내장형 소프트웨어에 대한 시험 성과 지표 선정 및 적용 방안에 대하여 LIG 넥스원에서 연구한 내용이다. 이 역시 ISO/IEC 9126-2의 기능성과 유지보수성, 신뢰성의 부특성의 지표를 해석하여 무기체계에

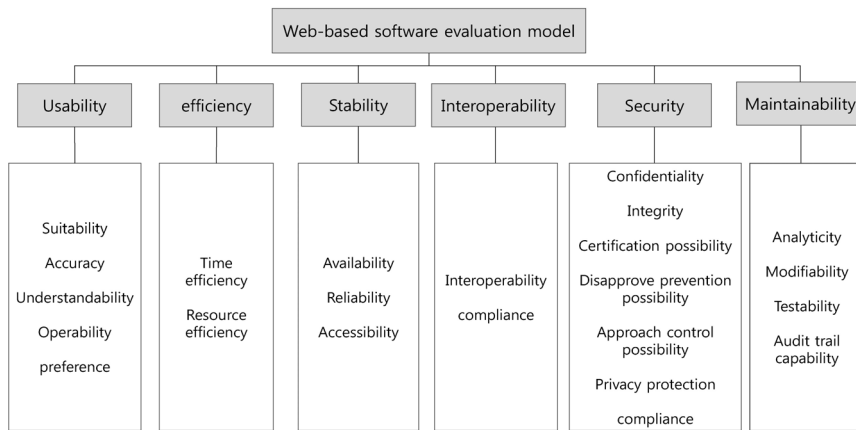


Fig. 2 Web-based software quality evaluation model

Table 2 Weapon systems embedded software test performance metrics

| Characteristic   | Indicator name                               | Measurement formula   |
|------------------|--|---|
| Functionality    | Requirements coverage                        | Number of Requirements confirmed by the test cases / Total number of requirements |
| Reliability      | Defect density                               | Number of occured defect / KLOC   |
|                  | Defect measure rate                          | Number of measured defect / Number of occurred defect                             |
|                  | Test coverage                                | Performed test cases / Designed test cases  |
|                  | Test maturity                                | Passed test cases / Designed test cases   |
|                  | Syntax Coverage                              | Execution rate the code at execute test cases                                     |
| Maintainability  | Source code comment rate                     | Comment line / Total line   |
|                  | Coding standard compliance rate              | Violation coding rule / Designed coding rule                                      |
|                  | Complexity                                   | Conditions in the functions + 1   |
| Interoperability | Interoperability requirements coverage       | Number of interoperability requirements / Total number of requirements            |
|                  | Standard compatibility requirements coverage | Number of standard compatibility requirement / Total number of requirements       |

적합한 성과 지표를 재 정의한다. 이 성과 지표에 대한 특성별 성과 지표 명과 측정 공식을 기술한 표는 <Table 2>와 같다[10].

위 3가지 기준 모두 ISO/IEC 9126-2 국제 표준을 활용하여 평가 기준을 제시하고 있다 하지만 평가 기준 항목이 선정된 이유에 대한 설명이 없고 어떤 항목이 먼저 선택 및 고려될 수 있는가 하는 등의 분류가 부족하다.

### 3. 연구 방법

#### 3.1 ISO/IEC 9126-2 품질 메트릭 분류

ISO/IEC 9126-2 표준에서는 소프트웨어의 품질특성을 기능성, 신뢰성, 사용성, 효율성, 유지보수성 이 식성 총 6개로 분류하여 정의하고 품질특성 별로 품질 메트릭을 나열하고 있다. 하지만 본 표준은 모든 상용 소프트웨어에 대하여 시험 및 운용 단계 동안 사용되는 전체 품질 평가 지표에 대하여 나열하고 있다 [7]. 따라서 소프트웨어 제품에 대한 정량적 평가 항목 및 기준을 선정하는데 ISO/IEC 9126-2에서 나열하는 전체 메트릭을 적용하는 것은 거의 불가능하고 제품에 따라 선택적으로 적용하여야 한다. 그래서 소프트웨어 공통적으로 해당되는 메트릭과 정보보호, 안전과 같은 특수한 목적으로 개발되었거나 자동 복구 기능과 같은 기능이 존재할 때만 해당되는 메트릭과 같이 분류하는 것이 필요하다.

따라서 ISO/IEC 9126-2 표준에 정의된 메트릭 테이블의 내용을 분석하여 정량적 평가 지표 및 기준을 도출하는데 품질 특성별로 공통적으로 활용 가능한 메트릭을 상위분류, 특수한 경우에만 활용 가능한 메트릭을 하위분류로 구분하였다. 각 품질 특성별로 상위분류와 하위분류로 구분한 메트릭명, 분류한 근거에 대한 사항은 <Table 3>과 같다.

먼저 기능성의 상위분류 메트릭은 모든 소프트웨어의 가장 기본이 되고 계산 기능 및 수치 표현 기능이 존재하는 소프트웨어의 평가 지표가 될 수 있다 그리고 웹/앱 소프트웨어나 외부 하드웨어와 통신하는 임베디드 소프트웨어의 중요 지표이다. 다음 하위분류 메트릭은 제품설명서, 사용자설명서와 같은 문서가 있어야만 하거나 시험 시 계산정확성 메트릭과 수행되는 내용이 동일하고 사용자의 접근통제기능이 존재하는 제품에만 평가가 가능하다. 그리고 정보보호

및 정보시스템 보호를 목적으로 하는 제품, 안전 분야 소프트웨어에만 평가가 가능하고 법/시행령/시행규칙 등 특수한 규정이 존재하는 경우에만 해당된다

신뢰성의 상위분류 메트릭은 설치형 소프트웨어의 중요 평가 지표이고 소프트웨어 개발 프로세스에서 정적 분석 및 동적 분석을 수행하는 소프트웨어의 평가 지표가 된다. 또한 오류 주입 시험이 가능한 소프트웨어의 평가 지표가 될 수 있다. 하위분류 메트릭은 자동 복구 및 자동 복원과 같은 특수한 기능이 존재하는 소프트웨어에만 평가가 가능하고 가장 적합한 신뢰성 성장 예측 모델을 선택하는 어려움이 존재하며 테스트케이스의 적합성에 따라 메트릭의 결과가 달라질 수 있는 경우가 있다. 또한 고장 발생 간격 시간의 분포, 운영 시간 간격에 따른 평균 시간의 변화 등 고려해야 할 사항이 많고 기능 준수성과 마찬가지로 법/시행령/시행규칙 등 특수한 규정이 존재하는 경우에만 메트릭 평가가 가능하다.

사용성의 상위분류 메트릭은 도움말 기능을 제공하고 사용자 인터페이스가 존재하는 소프트웨어의 평가 지표이고 매개변수 값을 선택할 수 있는 기능과 오류 메시지를 제공하는 소프트웨어의 평가 지표이다. 하위분류 메트릭은 소프트웨어 이외에 사용설명서를 평가하는 메트릭이며 데모 및 단축기능, 사용자가 기능의 내용/절차 등을 직접 변경 가능한 기능과 같이 특수한 경우에만 시험이 가능한 메트릭이다. 마지막으로 법/시행령/시행규칙 등 특수한 규정이 존재하는 경우에만 평가 항목으로 도출이 가능하다.

제품의 성능과 가장 관련 있는 효율성의 상위분류 메트릭은 Response time, Throughput, Turnaround time, Waiting time, 메모리 사용률 측정이 가능한 모든 소프트웨어의 평가지표이다. 하위 메트릭은 외부 I/O 장비의 로딩시간에 대한 정확한 측정이 어렵고 법/시행령/시행규칙 등 특수한 규정이 존재하는 경우에만 메트릭 평가가 가능하다.

유지보수성의 상위분류 메트릭은 로그를 남기는 모든 소프트웨어, 진단 기능 및 모니터링 기능, 매개변수 변경 기능을 제공하는 모든 소프트웨어의 평가 지표이다. 하위분류 메트릭은 고장의 원인 분석, 제품의 버전업, 시험 방법과 관련된 측정기 어렵거나 재시험을 해야만 메트릭 측정이 가능하다. 또한 유지보수성 관련 준수 요구사항이 현재 없는 이유로 하위분류로 구분된다.

**Table 3** Iso/iec 9126–2 metrics classification table

| No | Characteristic                       | Metric Name  | No                          | Characteristic  | Metric Name   |
|----|--------------------------------------|--|-----------------------------|-----------------|---|
|    |                                      | High-Classification                                |                             |                 | Sub-Classification  |
| 1  | Functionality                        | Functional adequacy                                | 1                           | Functionality   | Functional specification stability                                |
| 2  |                                      | Functional implementation completeness             | 2                           |                 | Accuracy to exception   |
| 3  |                                      | Functional implementation coverage                 | 3                           |                 | Access auditability   |
| 4  |                                      | Computational Accuracy                             | 4                           |                 | Access controllability  |
| 5  |                                      | Precision  | 5                           |                 | Data corruption prevention  |
| 6  |                                      | Data exchangeability                               | 6                           |                 | Functional compliance   |
| 1  | Reliability                          | Failure resolution                                 | 7                           | Reliability     | Interface standard compliance                                     |
| 2  |                                      | Fault density                                      | 1                           |                 | Mean recovery time  |
| 3  |                                      | Fault removal                                      | 2                           |                 | Mean down time  |
| 4  |                                      | Test coverage                                      | 3                           |                 | Restartability  |
| 5  |                                      | Breakdown avoidance                                | 4                           |                 | Restorability   |
| 6  |                                      | Failure avoidance                                  | 5                           |                 | Restore effectiveness   |
| 7  |                                      | Incorrect operation avoidance                      | 6                           |                 | Estimated latent fault density                                    |
| 1  | Usability                            | Help accessibility                                 | 7                           | Usability       | Failure density against test cases                                |
| 2  |                                      | Operational consistency in use                     | 8                           |                 | Mean time between failures (MTBF)                                 |
| 3  |                                      | Default value availability in use                  | 9                           |                 | Test maturity   |
| 4  |                                      | Message understandability in use                   | 10                          |                 | Availability  |
| 5  |                                      | Self-explanatory error messages                    | 11                          |                 | Reliability compliance  |
| 6  |                                      | Physical accessibility                             | 1                           |                 | Completeness of description                                       |
| 7  | Interface appearance customisability | 2  | Demonstration accessibility |                 |   |
| 1  | Efficiency                           | Response time                                      | 3                           | Usability       | Demonstration accessibility in use                                |
| 2  |                                      | Response time (Mean time to response)              | 4                           |                 | Demonstration effectiveness                                       |
| 3  |                                      | Response time (Worst case response time ratio)     | 5                           |                 | Evident functions   |
| 4  |                                      | Throughput   | 6                           |                 | Function understandability  |
| 5  |                                      | Throughput (Mean amount of throughput)             | 7                           |                 | Understandable input and output                                   |
| 6  |                                      | Throughput (Worst case throughput ratio)           | 8                           |                 | Ease function learning  |
| 7  |                                      | Turnaround time                                    | 9                           |                 | Ease of learning to perform a task in use                         |
| 8  |                                      | Turnaround time (Mean time for turnaround)         | 10                          |                 | Help frequency  |
| 9  |                                      | Turnaround time (Worst case turnaround time ratio) | 11                          |                 | Operational error recoverability in use                           |
| 10 |                                      | Waiting time                                       | 12                          |                 | Time between human error operations in use                        |
| 11 |                                      | I/O related errors                                 | 13                          |                 | Effectiveness of the user documentation and/or help system        |
| 12 |                                      | Mean I/O fulfilment ratio                          | 14                          |                 | Effectiveness of the user documentation and/or help system in use |
| 13 |                                      | User waiting time of I/O devices utilisation       | 15                          |                 | Error correction  |
| 14 |                                      | Maximum memory utilisation                         | 16                          |                 | Error correction in use   |
| 15 |                                      | Mean occurrence of memory error                    | 17                          |                 | Undoability   |
| 16 |                                      | Ratio of memory error/time                         | 18                          |                 | Customisability   |
| 17 |                                      | Maximum transmission utilisation                   | 19                          |                 | Operation procedure reduction                                     |
| 18 |                                      | Mean occurrence of transmission error              | 20                          |                 | Attractive interaction  |
| 19 |                                      | Transmission capacity utilisation                  | 21                          |                 | Usability compliance  |
| 1  | Maintainability                      | Audit trail capability                             | 1                           | Efficiency      | I/O loading limits  |
| 2  |                                      | Diagnostic function support                        | 2                           |                 | Media device utilisation balancing                                |
| 3  |                                      | Status monitoring capability                       | 3                           |                 | Mean of transmission error per time                               |
| 4  |                                      | Parameterised modifiability                        | 4                           |                 | Efficiency Compliance   |
| 1  | Portability                          | Adaptability of data structures                    | 1                           | Maintainability | Failure analysis capability                                       |

| No | Characteristic | Metric Name                                |                    |  |   |
|----|----------------|--|--------------------|--|---|
|    |                | High-Classification                        | Sub-Classification |  |   |
| 2  |                | Hardware environmental adaptability        | 2                  | Failure analysis efficiency                                      |   |
| 3  |                | Porting user friendliness                  | 3                  | Change cycle efficiency  |   |
| 4  |                | System software environmental adaptability | 4                  | Change implementation elapse time                                |   |
| 5  |                | Ease of installation                       | 5                  | Modification complexity  |   |
| 6  |                | Available co-existence                     | 6                  | Software change control capability                               |   |
|    |                |  |                    | 7  | Change success ratio                    |
|    |                |  | 8                  | Modification impact localisation (Emerging failure after change) |   |
|    |                |  | 9                  | Availability of built-in test function                           |   |
|    |                |  | 10                 | Re-test efficiency   |   |
|    |                |  | 11                 | Test restartability  |   |
|    |                |  | 12                 | Maintainability compliance                                       |   |
|    |                |  | 1                  | Portability  | Organisational environment adaptability |
|    |                |  | 2                  |  | Ease of Setup Retry                     |
|    |                |  | 3                  |  | Continued use of data                   |
|    |                |  | 4                  |  | Function inclusiveness                  |
|    |                |  | 5                  |  | User support functional consistency     |
|    |                |  | 6                  |  | Portability compliance                  |

마지막으로 이식성의 상위분류 메트릭이다 외부 데이터 Import 기능이 존재하는 모든 소프트웨어의 평가 지표이고 운영환경이 존재하는 모든 소프트웨어의 평가 지표이다. 하위분류메트릭은 조직 환경에서의 사업적인 문제가 없는지 판단하는 특수한 메트릭으로 정량적 측정이 어렵고 소프트웨어 제품의 이전 버전이나 다음 버전이 있을 경우에만 측정이 가능하다. 그리고 법/시행령/시행규칙 등 특수한 규정이 존재하는 경우에만 본 메트릭 평가가 가능하여 하위 분류로 구분된다.

ISO/IEC 9126-2 메트릭을 활용하여 소프트웨어의 정량적 평가 항목 선정 시 위의 사항을 고려하여 품질 특성별로 상위분류로 선정된 메트릭을 먼저 선택할 수 있다

### 3.2 ISO/IEC 9126-2 품질 메트릭 필요도/중요도 분석 설문조사

ISO/IEC 9126-2 품질 메트릭을 분류한 결과의 유효성을 확인하고 소프트웨어 제품 특성별 품질 메트릭 분류를 위하여 2년 이상 GS인증을 통해 ISO/IEC 9126-2 표준을 이해하고, 소프트웨어 시험으로 다양한 소프트웨어 제품을 시험한 경험이 있는 연구원을 대상으로 소프트웨어 제품별 ISO/IEC 9126-2 품질 메

트릭의 필요도 및 중요도에 대한 설문조사를 수행하였다. 소프트웨어 제품은 크게 웹/앱 소프트웨어, 임베디드 소프트웨어, 설치형 소프트웨어 3가지로 분류하였다. 그리고 필요도와 중요도를 각각 5점을 최고 점으로 5점(매우 중요, 매우 필요), 4점(중요, 필요), 3점(보통), 2점(중요하지 않음, 불필요), 1점(매우 중요하지 않음, 매우 불필요)으로 설문조사를 진행하였다. 이 후 필요도와 중요도의 합산한 점수가 높은 순서대로 순위를 구분하였다. 이 중 합산 평균 점수가 6점 이상으로 보통 이상 필요하고 중요하다고 조사된 메트릭을 나열하면 <Table 4>와 같다.

중요도와 필요도 측면에서 보통 이상의 등급으로 소프트웨어 제품의 공통으로 적용되는 품질 메트릭은 기능성의 1)기능 구현의 완전성, 2)기능 적합성, 3)기능 구현의 정도 사용성의 4)기능 이해성, 5)사용 중 운용의 일관성 효율성의 6)응답시간, 7)처리율, 8)작업 완료시간, 9)최대 응답시간 비율. 유지보수성의 10)감사 추적성. 이식성의 11)시스템 소프트웨어 환경 적응성, 12)데이터 구조 적응성으로 분석되었다. 따라서 위 12개의 메트릭은 모든 소프트웨어에 정량적 평가 지표에 대한 항목으로 활용될 수 있다. 그리고 이 메트릭 모두 제3.1절의 ISO/IEC 9126-2 메트릭 분류에서 상위분류로 구분된 항목이다

**Table 4** Web/app, embedded, installation software survey result table

| Web/App software    |  |               | Embedded software   |  |               | Installation software |  |               |
|---------------------|--|---------------|---------------------|--|---------------|-----------------------|--|---------------|
| Main-characteristic | Metric name                            | Average-grade | Main-characteristic | Metric name                            | Average-grade | Main-characteristic   | Metric name                                    | Average-grade |
| Functionality       | Functional implementation completeness | 10.0          | Functionality       | Functional adequacy                    | 9.2           | Functionality         | Functional adequacy                            | 10.0          |
|                     | Functional adequacy                    | 9.9           |                     | Computational accuracy                 | 8.1           |                       | Functional implementation coverage             | 8.4           |
|                     | Access auditability                    | 7.8           |                     | Functional implementation completeness | 8.0           |                       | Functional implementation completeness         | 8.2           |
|                     | Functional implementation coverage     | 7.7           |                     | Precision                              | 7.9           |                       | Computational accuracy                         | 6.4           |
|                     | Data exchangeability                   | 7.6           |                     | Functional implementation coverage     | 7.4           |                       | Precision                                      | 6.0           |
|                     | Computational accuracy                 | 6.4           |                     | Data corruption prevention             | 7.2           | Reliability           | Failure resolution                             | 6.3           |
|                     | Access controllability                 | 6.2           |                     | Data exchangeability                   | 6.8           |                       | Operational consistency in use                 | 8.2           |
| Reliability         | Fault removal                          | 6.2           |                     | Access controllability                 | 6.4           |                       | Message understandability in use               | 8.0           |
| Usability           | Function understandability             | 9.0           | Reliability         | Fault removal                          | 9.6           | Usability             | Function understandability                     | 7.9           |
|                     | Evident functions                      | 8.9           |                     | Failure resolution                     | 9.4           |                       | Evident functions                              | 7.8           |
|                     | Operational consistency in use         | 7.5           |                     | Fault density                          | 8.8           |                       | Self explanatory error messages                | 7.4           |
|                     | Message understandability in use       | 7.2           |                     | Failure density against test cases     | 8.4           |                       | Undoability                                    | 6.6           |
|                     | Self explanatory error messages        | 7.2           |                     | Break-down avoidance                   | 7.0           |                       | Understandable input and output                | 6.5           |
|                     | Undoability                            | 6.7           |                     | Test coverage                          | 6.8           |                       | Error correction                               | 6.4           |
|                     | Physical accessibility                 | 6.6           |                     | Availability                           | 6.8           | Efficiency            | Response time                                  | 9.7           |
|                     | Help accessibility                     | 6.5           |                     | Mean time between failures (MTBF)      | 6.6           |                       | Throughput                                     | 8.6           |
|                     | Understandable input and output        | 6.4           |                     | Restartability                         | 6.3           |                       | Response time (Mean time to response)          | 7.9           |
|                     | Interface appearance                   | 6.3           |                     | Estimated latent fault density         | 6.1           |                       | Maximum memory utilization                     | 7.7           |
| Efficiency          | Response time (Mean time to response)  | 8.4           | Usability           | Error correction                       | 7.3           |                       | Response time (Worst case response time ratio) | 7.6           |
|                     | Response time                          | 8.2           |                     | Operational consistency in use         | 6.6           |                       | Turnaround time                                | 7.4           |
|                     | Turnaround time                        | 7.8           |                     | Function understandability             | 6.2           |                       | Mean occurrence of memory error                | 6.5           |



| Web/App software    |  |               | Embedded software                |  |  | Installation software               |  |               |  |     |
|---------------------|--|---------------|----------------------------------|--|--|-------------------------------------|--|---------------|--|-----|
| Main-characteristic | Metric name  | Average-grade | Main-characteristic              | Metric name  | Average-grade                              | Main-characteristic                 | Metric name                              | Average-grade |  |     |
|                     | Throughput   | 7.1           | Efficiency                       | Response time                                      | 9.2  |                                     | Throughput (Mean amount of throughput)   | 6.4           |  |     |
|                     | Maximum memory utilization                         | 6.8           |                                  | Response time (Mean time to response)              | 8.0  |                                     | Throughput (Worst case throughput ratio) | 6.4           |  |     |
|                     | Response time (Worst case response time ratio)     | 6.7           |                                  | Throughput   | 7.4  | Maintainability                     | Audit trail capability                   | 7.1           |  |     |
|                     | Turnaround time (Mean time to turnaround)          | 6.4           |                                  | Response time (Worst case response time ratio)     | 7.2  |                                     | Change success ratio                     | 6.6           |  |     |
|                     | Turnaround time (Worst case turnaround time ratio) | 6.3           |                                  | Turnaround time                                    | 6.2  |                                     | Diagnostic function support              | 6.4           |  |     |
| Maintainability     | Audit trail capability                             | 7.1           |                                  | Turnaround time (Mean time to turnaround)          | 6.1  |                                     | Status monitoring capability             | 6.0           |  |     |
|                     | System software environmental adaptability         | 7.9           |                                  | Turnaround time (Worst case turnaround time ratio) | 6.1  |                                     | Modification impact localization         | 6.0           |  |     |
| Portability         | Porting user friendliness                          | 7.2           |                                  | Maintainability                                    | I/O devices utilization                    |                                     | 6.0                                      | Portability   | System software environmental adaptability | 9.3 |
|                     | Adaptability of data structures                    | 7.0           |                                  |  | Audit trail capability                     |                                     | 7.6                                      |               | Porting user friendliness                  | 8.7 |
|                     | Available co-existence                             | 6.9           |                                  |  | Diagnostic function support                |                                     | 7.1                                      |               | Ease of installation                       | 8.6 |
|                     | Ease of installation                               | 6.1           | Failure analysis capability      |  | 7.1  | Adaptability of data structures     | 7.7                                      |               |  |     |
|                     |  |               | Modification impact localization |  | 6.9  | Ease of setup retry                 | 7.3                                      |               |  |     |
|                     |  |               | Change success ratio             |  | 6.2  | Hardware environmental adaptability | 7.1                                      |               |  |     |
|                     |  |               | Portability                      |  | Hardware environmental adaptability        | 7.9                                 |  |               |  |     |
|                     |  |               |                                  |  | System software environmental adaptability | 6.6                                 |  |               |  |     |
|                     |  |               |                                  |  | Adaptability of data structures            | 6.2                                 |  |               |  |     |

제품의 특성별로 웹/앱 소프트웨어 제품의 경우 사용성 품질특성의 메트릭, 임베디드 소프트웨어 제품은 신뢰성 품질특성의 매트릭이 정량적 평가 지표에

대한 항목으로 추가적으로 활용될 수 있다. 그리고 설치형 소프트웨어 제품의 경우 웹앱 소프트웨어 제품과 유사하나 추가로 유지보수성 품질특성의 메트릭

이 정량적 평가지표 항목으로 활용될 수 있다

#### 4. 사례 연구

본 연구는 ISO/IEC 9126-2 메트릭을 활용하여 소프트웨어에 대한 정량적 평가 지표를 선정하는데 그 목적을 두고 있다. 따라서 제3장에서 소프트웨어 제품별(웹/앱 소프트웨어, 임베디드 소프트웨어, 설치형 소프트웨어

어)로 상위분류와 하위분류로 구분한 정량적 평가 지표 항목이 실제로 유효한지 여부와 ISO/IEC 9126-2 메트릭을 활용하여 정량적 평가 항목과 기준을 선정 시에도 동일한 결과도 출되는지에 대한 검증을 실시하였다. 한국산업기술시험원에서 수행한 신청자 제시 기준 소프트웨어 시험 제품 중 웹/앱 소프트웨어, 임베디드 소프트웨어, 설치형 소프트웨어 제품에 대하여 기업에서 제시한 시험항목 및 시험기준, 시험결과와 ISO/IEC 9126-2 품질 메트릭을 활용한 항목, 기준, 결과를 비교하였다.

Table 5 Web/App software case

| Applicant suggest                                       |  |                                  | Utilize ISO/IEC 9126-2 metrics             |   |   | Comparison |
|---|--|----------------------------------|--|---|---|------------|
| Item  | Criteria   | Results                          | Item                                       | Criteria  | Results   |            |
| Processing of simultaneous connections                  | Keep the web login concurrent connected users more than XX people  | XX people connection is possible | Transmission capacity utilization          | Does it perform the operation properly when more than XX concurrent connected users are connected?  | when XX people are connect, perform properly                  | match      |
| connection response speed                               | Web login response time should be within XX s  | Maximum XX s                     | Response time                              | How long is the response time of software web login?  | Maximum XX s  | match      |
| Control of the loss rate of the encrypted communication | When sending the encoded message more than XX cases in transitive server, receiving rate of encode message is more than XX % | XX %                             | Throughput                                 | When sending the encoded message more than XX cases in transitive server, how many encode message is received successfully?                   | XX EA success received messages/XX EA sending messages (XX %) | match      |
| Mobile access control                                   | When unregistered mobile phone try to login, failure should be more than XX%   | XX %                             | Access controllability                     | When unregistered mobile phone try to login, How much access can be blocked?  | XX %  | match      |
| Ensure the validity of the browser                      | Web browser the login function is working properly is XX EA or more  | XX EA                            | System software environmental adaptability | Login function is work properly with XX or more of the Web browser of the operating environment?  | Work properly with XX EA Web browser                          | match      |
| Ensure the validity of the XML                          | Among the XX kinds of message in software, more than XX % message should be able to verify availability using XML Schema     | XX %                             | Data corruption prevention                 | Through verifying availability using XML Schema, How many messages are able to verify availability among the XX kinds of message in software? | XX %  | match      |
| Ensure the validity of the CSS                          | Web browser should be more than XX that it can be printed in main and important screen                                       | XX EA                            | System software environmental adaptability | Does printing function of main and important screen work properly in more than XX kinds of web browser operating environments?                | Work properly with XX EA Web browser                          | match      |

먼저 신청자가 제시한 항목 및 기준을 분석하여ISO/IEC 9126-2 메트릭의 어떠한 항목에 해당할 수 있는지를 판단하였다. 신청자가 제시한 항목 및 기준은 각 소프트웨어 제품에 맞게 상세히 서술되어 있고, ISO/IEC 9126-2 메트릭 이름과 메트릭의 목적은 모든 상용 소프트웨어에 적용될 수 있게 광범위하게 서술되어있다. 따라서 ISO/IEC 9126-2 메트릭의 목적을 변형하여 신청자가 제시한 기준 수준으로 상세히 기술하였다

첫째, 웹/앱 소프트웨어이다. B2B 업무용 비즈니스 플랫폼 소프트웨어로 웹 소프트웨어와 앱 소프트웨어가 모두 존재하는 제품이다. 신청자 제시 시험 항목과 기준, 시험 결과에 대한 내용과 ISO/IEC 9126-2 품질 메트릭을 활용한 항목, 기준, 결과를 비교한 사항은 <Table 5>를 참고한다.

먼저 신청자가 제시한 7개의 항목을 ISO/IEC 9126-2 메트릭으로 활용 시 6개의 항목으로 축소가 가능하였다. 신청자가 제시한 항목인 브라우저 유효성 확보와 CSS 유효성 확보 2개의 항목이 ISO/IEC 9126-2 메트릭의 시스템 소프트웨어 환경 적응성 한 개의 항목 하나에 해당하였다. 그리고 6개의 항목 모두 웹/앱 소프트웨어 필요도, 중요도에 따른 선정 메트릭에서 중요하고 필요하다고 판단되는 항목이었다.

다음으로 임베디드 소프트웨어이다. 어뢰경보체계 시스템으로 선체에 부착된 시스템이 어뢰를 탐지 시

이에 대한 알람을 제공하기 위해 만들어진 제품에 탑재되는 소프트웨어이다. 신청자 제시 시험 항목과 기준, 시험 결과에 대한 내용과 ISO/IEC 9126-2 품질 메트릭을 활용한 항목, 기준, 결과를 비교한 사항은 <Table 6>을 참고한다.

신청자가 제시한 3개의 항목을 ISO/IEC 9126-2 메트릭으로 활용 시 2개의 항목으로 축소가 가능하였다. 신청자가 제시한 정적분석과 정적분석메모리 항목이 ISO/IEC 9126-2 메트릭의 결합 밀도 항목으로 활용이 가능하였다. 그리고 2개의 항목 모두 임베디드 소프트웨어 필요도, 중요도에 따른 선정 메트릭에서 중요하고 필요하다고 판단되는 항목이었다.

마지막으로 설치형 소프트웨어이다. 차량에 설치되는 LED를 검사하는 검사기 하드웨어 장치를 제어 및 모니터링하기 위한 PC에 설치되는 소프트웨어이다. 신청자 제시 시험 항목과 기준, 시험 결과에 대한 내용과 ISO/IEC 9126-2 품질 메트릭을 활용한 항목, 기준, 결과를 비교한 사항은 <Table 7>을 참고한다.

신청자가 제시한 4개의 항목을 ISO/IEC 9126-2 메트릭으로 활용 시 2개의 항목으로 축소가 가능하였다. 신청자가 제시한 항목 가운데 파라미터 설정 연계1, 2와 모델 수량 관리 항목이 ISO/IEC 9126-2 메트릭으로 활용 시 계산 정확성 한 개의 항목 하나에 해당하였다. 또한 모두 설치형 소프트웨어 필요도, 중요도에

Table 6 Embedded software case

| Applicant suggest      |  |   | Utilize ISO/IEC 9126-2 metrics |  |   | Comparison |
|------------------------|--|---|--------------------------------|--|---|------------|
| Item                   | Criteria   | Results                                       | Item                           | Criteria   | Results                                       |            |
| Static Analysis        | Confirm the result of applying the Code Rule List using LDRA 9.1.0 tool about XX kinds of software project                         | Violation XX EA                               | Fault density                  | How many faults are discovered when applying the Code Rule List using LDRA 9.1.0 tool about XX kinds of software project?                            | Violation XX EA                               | match      |
| Static Analysis-Memory | Confirm the result of applying the Code Rule List using Sparrow4.6_t10 tool about XX kinds of software project                     | Violation XX EA                               | Fault density                  | How many faults are discovered when applying the Code Rule List using Sparrow4.6_t10 tool about XX kinds of software project?                        | Violation XX EA                               | match      |
| Unit Test              | Confirm the Statement Coverage, Branch Coverage of each project functions using LDRA 9.1.0 tool about XX kinds of software project | Statement Coverage XX %, Branch Coverage XX % | Test coverage                  | How much does the Statement Coverage, Branch Coverage of each project functions implement, using LDRA 9.1.0 tool about XX kinds of software project? | Statement Coverage XX %, Branch Coverage XX % | match      |

**Table 7** Installation software case

| Applicant suggest               |  |              | Utilize ISO/IEC 9126-2 metrics |   |               | Comparison |
|---------------------------------|--|--------------|--------------------------------|---|---------------|------------|
| Item                            | Criteria   | Results      | Item                           | Criteria  | Results       |            |
| Parameter setting cooperation-1 | Maximum and minimum of LDB quantity input into model screen should be applied  | Is applied   | Computational accuracy         | How much is it accurate compared between model screen of maximum and minimum of LDB quantity and input into model?                                  | Is consistent | match      |
| Parameter setting cooperation-2 | Implementing the number module features of model and inputting the quantity of module number as should apply and output input into inspection condition screen | Is output    | Computational accuracy         | How much is it accurate compared between module quantity result from the number module features of model and output in inspection condition screen? | Is consistent | match      |
| EMG button click                | From implementing EMG function to complete, It takes only within XXs   | Maximum XX s | Response time                  | How long does it take time to complete the EMG function?  | Maximum XX s  | match      |
| Model quantity management       | Model of good data and defective one should be as the last date that implemented the manual/automatic inspection   | Is stored    | Computational accuracy         | How much is it accurate between model of good data and defective one and the last date that implemented the manual/automatic inspection ?           | Is consistent | match      |

다른 선정 메트릭에서 중요하고 필요하다고 판단되는 항목이었다.

그리고 세 가지 소프트웨어 분류별 신청자 제시 기준에 대한 시험 결과 모두 ISO/IEC 9126-2 메트릭을 활용하여 시험 기준을 제품의 용도나 확인하고자 하는 목적에 맞게 상세히 서술 시 동일한 결과를 얻을 수 있음을 확인하였다.

### 5. 결론 및 향후 연구

소프트웨어 정량적 평가지표에 대한 항목과 시험 기준 및 방법을 도출하기 위해 ISO/IEC 9126-2의 메트릭을 분석하여 공통적으로 적용될 수 있는 메트릭과 특수한 소프트웨어에만 적용될 수 있는 메트릭으로 구분하고 상위분류와 하위분류로 나누었다. 그리고 소프트웨어 제품을 특성에 따라 웹/앱 소프트웨어, 임베디드 소프트웨어, 설치형 소프트웨어로 구분하여 제품의 특성별 중요도와 필요도에 따른 메트릭을 선정하기 위하여

설문조사를 실시하였다. 마지막으로 소프트웨어 제품별 소프트웨어 R&D 프로젝트에서 기업이 제시한 시험 항목과 기준, 결과와 ISO/IEC 9126-2 메트릭을 활용한 항목과 기준과 결과를 비교하였다.

소프트웨어 R&D 과제에서 요구하는 정량적 평가표의 평가 항목과 평가 기준을 작성하는데 ISO/IEC 9126-2 품질 메트릭을 활용할 수 있다. 그리고 소프트웨어 제품의 분류별(웹/앱 소프트웨어, 임베디드 소프트웨어, 설치형 소프트웨어)로 ISO/IEC 9126-2 품질 메트릭 가운데에서도 더 중요하거나 필요하다고 판단되는 메트릭이 구분되었다. 이를 활용하여 현실성을 반영하고, 좀 더 쉽게 정량적인 결과를 도출하는 것이 가능하다.

개인정보보호법이나 예측의 정확성과 같은 특수한 항목에 대해서는 ISO/IEC 9126-2 품질 메트릭 항목과 기준으로 활용이 어려운 부분이 존재한다 따라서 향후 연구에서는 이러한 특수한 항목에 대하여 정량적 평가 항목과 기준을 제시할 수 있는 방법에 대한 연구가 필요하다.

## References

- [1] Telecommunications Technology Association (2010). "Software requirements quality evaluation guideline".
- [2] National Science & Technology Council (2013). "National research and development project standard performance index-Performance goal-index set guideline-".
- [3] Kim, M. J. (2015). "Software Reliability Evaluation based on Quality Evaluation Metrics", Chungnam National University.
- [4] Ministry of Knowledge Economy. (2008). "KS X ISO/IEC TR9126-2: 2008 Information technology-Software engineering- Product quality-Part 2: External metrics".
- [5] Kwon, W. I. and Chung, C. S. (2003). "International Standardization about the quality of the Software Product - Focusing on ISO/IEC9126, 14598 change". TTA Journal, Vol. 85, p. 209
- [6] ISO/IEC (2001). "ISO/IEC TR 9126-1 Software engineering - Product quality - Part 1: Quality model".
- [7] ISO/IEC (2003). "ISO/IEC TR 9126-2 Software engineering - Product quality - Part 2: External metrics".
- [8] Ministry of Knowledge Economy, National IT Industry Promotion Agency. (2010). "Software technical evaluation criteria application guide".
- [9] Telecommunications Technology Association (2006). "A study on software quality evaluation framework development".
- [10] LIG Nex1. (2013). "A Study on Weapon Systems Embedded Software Test Performance Indicators Selection and Application".