



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2013-0060042
(43) 공개일자 2013년06월07일

(51) 국제특허분류(Int. Cl.)

G06F 17/00 (2006.01) G06F 11/28 (2006.01)

(21) 출원번호 10-2011-0126327

(22) 출원일자 2011년11월29일

심사청구일자 없음

(71) 출원인

이화여자대학교 산학협력단

서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교)

서강대학교 산학협력단

서울특별시 마포구 백범로 35 (신수동, 서강대학교)

(72) 발명자

최병주

서울특별시 강남구 압구정로 201, 74동 1305호 (압구정동, 현대아파트)

송미경

서울특별시 성북구 종암동 SK아파트 102-2406

(74) 대리인

남정길

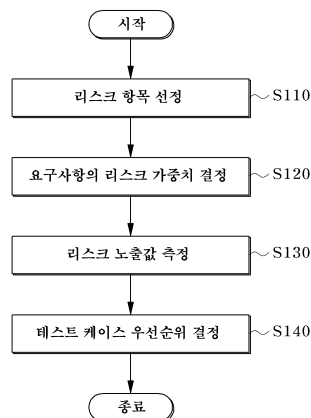
전체 청구항 수 : 총 3 항

(54) 발명의 명칭 **리스크 노출 값 기반의 테스트 케이스 우선순위 결정 방법**

(57) 요약

개시된 기술은 테스트 케이스의 우선순위를 선정하는 방법에 관한 것으로, 보다 상세하지만 제한됨 없이는, 리스크 노출 값을 적용한 테스트 케이스 우선순위 선정 방법에 관한 것이다. 실시예들 중에서, 테스트 케이스 우선순위 결정방법은 요구사항들의 각 항목에 대하여 리스크 가중치를 결정하는 단계; 미리 선정된 리스크 항목들 중 각각의 리스크 항목별로 상기 리스크 가중치를 반영하여 리스크 노출 값을 결정하는 단계; 및 테스트 케이스들 중 각각의 테스트 케이스 마다 상기 리스크 항목들의 발생 횟수 및 상기 리스크 노출 값을 기반으로 각각의 테스트 케이스의 우선순위를 결정하는 단계를 포함한다.

대표도 - 도1



이 발명을 지원한 국가연구개발사업

과제고유번호 C1090-1131-0008

부처명 정보통신연구진흥원

연구사업명 정보통신기술인력양성

연구과제명 융합소프트웨어를 위한 요구 및 품질검증 기술 개발

주관기관 서강대학교 산학협력단

연구기간 2011.01.01 ~ 2011.12.31

특허청구의 범위

청구항 1

요구사항들의 각 항목에 대하여 리스크 가중치를 결정하는 단계;

미리 선정된 리스크 항목들 중 각각의 리스크 항목별로 상기 리스크 가중치를 반영하여 리스크 노출 값을 결정하는 단계;

테스트 케이스들 중 각각의 테스트 케이스 마다 상기 리스크 항목들의 발생 횟수 및 상기 리스크 노출 값을 기반으로 각각의 테스트 케이스의 우선순위를 결정하는 단계를 포함하는 테스트 케이스 우선순위 결정방법.

청구항 2

제1항에 있어서, 상기 리스크 가중치를 결정하는 단계는,

AHP(Analytic Hierarchy Process) 기법을 적용하여 요구사항의 리스크 가능성 가중치 및 요구사항의 리스크 영향력 가중치를 산출하는 단계; 및

상기 요구사항의 리스크 가능성 가중치 및 상기 요구사항의 리스크 영향력 가중치를 기초로 상기 리스크 가중치를 산출하는 단계를 포함하는 테스트 케이스 우선순위 결정방법.

청구항 3

제1항에 있어서, 상기 우선순위를 결정하는 단계는,

상기 각각의 테스트 케이스에 대하여,

$$TCP_i = \sum_{j=1}^n (RE(r_j) \times TM_{ij})$$
 (여기서, TCP_i 는 i 번째 테스트 케이스 TC_i 에 대한 우선순위 값, TM_{ij} 는 테스트 케이스 $i(TC_i)$ 로 실행했을 때, 리스크 항목 $j(R_j)$ 에 해당하는 부분의 발생 횟수, $RE(r_j)$ 는 리스크 항목 r_j 에 대한 리스크 노출값을 나타냄)에 따라 산출되는 TCP 값이 높은 순으로 상기 각각의 테스트 케이스의 우선순위를 결정하는 단계를 포함하는 테스트 케이스 우선순위 결정방법.

명세서

기술분야

[0001] 개시된 기술은 테스트 케이스의 우선순위를 선정하는 방법에 관한 것으로, 보다 상세하지만 제한됨 없이는, 리스크 노출 값을 적용한 테스트 케이스 우선순위 선정 방법에 관한 것이다.

배경기술

[0002] 테스트 케이스를 모두 실행하는 테스트는 몇몇 작은 규모의 특별한 경우를 제외하고는 현실적으로 불가능하다. 일례로, 소스코드의 크기가 20,000라인인 소프트웨어의 경우 전체 테스트 케이스를 실행시키려면 7주가 필요하다고 한다. 따라서 커버리지 선정기준(Criteria)를 적용하여 테스트 케이스들의 일부분을 선정하는 기술이 요구된다. 오류에 대한 효율성과 효과성이 높은 테스트 케이스 선정 기술에 대한 연구 가운데, 테스트 케이스 우선순위 선정 기술은 테스트 케이스들의 우선순위를 결정하고 예산에 따른 테스트 케이스 집합의 규모에 따라 상위 순위의 테스트 케이스들을 선정하여 보다 효율적인 테스트 실행을 유도한다. 본 명세서에서는 테스트 케이스 우선순위 부여 기술을 제안하고, 그의 효과성과 효율성을 실험을 통하여 보인다.

[0003] 기존의 우선순위 부여 방법은 회귀테스트에 적용되는 것으로서 이전 테스트 수행 결과 자료를 근거로 우선순위를 부여하게 된다. 이처럼 주어진 프로그램의 결함이 무엇인지와 어떤 테스트 케이스가 어떤 결함을 발견하는지에 대한 정보를 모두 알고 있다는 가정 하에 테스트 케이스 우선순위를 선정하는 방법으로 최적 우선순위 기법(Optimal prioritization)을 들 수 있다. 이는 이론상으로 가장 최적화된 방법이지만, 결함과 테스트 케이스에 대한 많은 정보를 활용해야 하므로 최악의 런타임이 소요된다. 또한 뮤테이션(mutation) 분석으로 우선 순위를 측정하는 방법인 FEP(Fault Exposing Potential) 방법도 경험적으로 얻어진 결함에 대한 정보가 존재해야 하는

제약을 갖는다. 따라서 위의 방법들은 최초 수행하는 테스트에는 적용할 수 없는 단점을 갖는다.

발명의 내용

해결하려는 과제

- [0004] 개시된 기술이 이루고자 하는 기술적 과제는 위의 방법이 사용하는 결함정보가 아닌 리스크 분석 결과를 활용하여, 테스트 케이스 실행 결과로 얻어지는 추가 정보가 없는 상황에서 적용할 수 있는 테스트 케이스 우선순위 선정 방법을 제공하는 데 있다. 제안하는 리스크 기반 테스트(Risk based Testing : RBT) 방법은 요구사항에서 식별할 수 있는 제품 리스크 항목을 평가 기준으로 활용하여 테스트 케이스의 우선순위를 선정한다.
- [0005] 기존의 RBT 연구에는 RiteDAP(Risk-based Test case Derivation And Prioritization) 등과 같이 리스크 분석 결과인 리스크 노출값(exposure)을 활용하는 경우와, Safety test처럼 테스트 케이스를 대상으로 측정한 심각도(severity)와 가능성(probability)를 활용하는 경우 등이 있다. RiteDAP는 액티비티 다이어그램(Activity Diagram, AD)에 리스크 정보를 입력하여 테스트 케이스 시나리오의 우선순위를 자동으로 생성한다. 액티비티 다이어그램에서 리스크에 해당되는 엔티티의 액션을 스테리오 타입 <<reAction>>으로 추가하였다. 테스트 케이스 시나리오의 우선순위는 해당 액티비티 다이어그램의 엔티티의 <<reAction>>의 리스크 값들을 모두 더한 리스크 값이 높은 순서에 따라 결정된다. RiteDAP는 리스크 값 측정이 주관적인 점이 단점이다. Safety test는 테스트 케이스를 비용과 심각성(severity probability)을 측정하여 테스트 케이스 우선순위를 측정한다. 테스트 케이스 비용은 고객이나 벤더가 생각하는 결함 중요성으로 측정한다. 심각성은 결함 수와 결함심각도의 곱으로 측정한다. 테스트 케이스 비용 측정이 주관적이며, 결함 정보가 존재해야 수행 가능한 방법이라는 제약을 가지고 있다. 이들 방법들은 리스크 노출값 측정이 주관적이며, 또한 결함 정보를 가지고 있어야 수행 가능한 방법이다. 즉 한번 이상의 테스트 결과가 존재해야 한다는 제약이 있다. 개시된 기술에서는 리스크 노출값 측정의 객관성을 유지하기 위하여 요구사항 항목마다 부여된 가중치를 기준으로 리스크 노출값을 추출하고, 이 값을 이용하여 테스트 케이스 우선 순위를 결정하는 방법을 제안한다.

과제의 해결 수단

- [0006] 상기의 기술적 과제를 이루기 위하여 개시된 기술의 제1 측면은 요구사항들의 각 항목에 대하여 리스크 가중치를 결정하는 단계; 미리 선정된 리스크 항목들 중 각각의 리스크 항목별로 상기 리스크 가중치를 반영하여 리스크 노출 값을 결정하는 단계; 및 테스트 케이스들 중 각각의 테스트 케이스 마다 상기 리스크 항목들의 발생 횟수 및 상기 리스크 노출 값을 기반으로 각각의 테스트 케이스의 우선순위를 결정하는 단계를 포함하는 테스트 케이스 우선순위 결정방법을 제공한다.

발명의 효과

- [0007] 개시된 기술의 실시 예들은 다음의 장점들을 포함하는 효과를 가질 수 있다. 다만, 개시된 기술의 실시 예들이 이를 전부 포함하여야 한다는 의미는 아니므로, 개시된 기술의 권리범위는 이에 의하여 제한되는 것으로 이해되어서는 아니 될 것이다.
- [0008] 개시된 기술의 일 실시예에 따른 리스크 기반 테스트 케이스 선정 방법은 사전에 테스트를 수행한 결과가 없는 상황에서 테스트 케이스의 우선순위를 판단할 수 있도록 하는 장점이 있다. 앞서 언급한 테스트 케이스 우선 순위 선정 기술들, 즉 optimal prioritization, safety test, FEP-Total, 등을 비교 대상으로 하여 본 연구에서 개발한 기술에 대한 실험적 평가를 수행한 결과, 결함 발견 빈도수와 평균 결함 검출률(Average Percentage of Fault Detection : APFD), 결함 심각도 항목에서 개시된 기술이 우수한 결과를 보임을 확인하였다.

도면의 간단한 설명

- [0009] 도 1은 개시된 기술의 일 실시예에 따른 테스트 케이스 우선순위 결정 방법을 설명하기 위한 도면이다.
- 도 2는 제품 리스크 항목을 나타낸 표이다.
- 도 3은 print_tokens의 요구사항과 각 요구사항에 대한 리스크 가중치를 나타낸다.
- 도 4는 리스크 노출값을 측정하는 데 사용되는 리스크 노출 메트릭을 나타낸 표이다.
- 도 5는 도 4의 리스크 노출 메트릭을 참조로 산출된 print_tokens에 대한 리스크 노출 측정 값이다.

도 6은 테스트 케이스의 우선순위를 측정하는데 사용되는 테스트 케이스 우선순위 측정 메트릭을 나타낸 표이다.

도 7은 print_tokens에 대한 테스트 케이스 우선순위를 측정한 예의 일부이다.

도 8 내지 도 13은 개시된 기술의 성능을 평가하기 위한 실험 결과를 나타낸 그래프이다.

발명을 실시하기 위한 구체적인 내용

- [0010] 개시된 기술에 관한 설명은 구조적 내지 기능적 설명을 위한 실시예에 불과하므로, 개시된 기술의 권리범위는 본문에 설명된 실시예에 의하여 제한되는 것으로 해석되어서는 아니 된다. 즉, 실시예는 다양한 변경이 가능하고 여러 가지 형태를 가질 수 있으므로 개시된 기술의 권리범위는 기술적 사상을 실현할 수 있는 균등물들을 포함하는 것으로 이해되어야 한다.
- [0011] 한편, 본 출원에서 서술되는 용어의 의미는 다음과 같이 이해되어야 할 것이다.
- [0012] “제1”, “제2” 등의 용어는 하나의 구성요소를 다른 구성요소로부터 구별하기 위한 것으로 이들 용어들에 의해 권리범위가 한정되어서는 아니 된다. 예를 들어, 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다.
- [0013] 어떤 구성요소가 다른 구성요소에 “연결되어” 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결될 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 “직접 연결되어” 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다. 한편, 구성요소들 간의 관계를 설명하는 다른 표현들, 즉 “~사이에”와 “바로 ~사이에” 또는 “~에 이웃하는”과 “~에 직접 이웃하는” 등도 마찬가지로 해석되어야 한다.
- [0014] 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한 복수의 표현을 포함하는 것으로 이해되어야 하고, “포함하다” 또는 “가지다” 등의 용어는 실시된 특징, 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0015] 각 단계들은 문맥상 명백하게 특정 순서를 기재하지 않은 이상 명기된 순서와 다르게 일어날 수 있다. 즉, 각 단계들은 명기된 순서와 동일하게 일어날 수도 있고 실질적으로 동시에 수행될 수도 있으며 반대의 순서대로 수행될 수도 있다.
- [0016] 여기서 사용되는 모든 용어들은 다르게 정의되지 않는 한, 개시된 기술이 속하는 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한 이상적이거나 과도하게 형식적인 의미를 지니는 것으로 해석될 수 없다.
- [0017] 도 1은 개시된 기술의 일 실시예에 따른 테스트 케이스 우선순위 결정 방법을 설명하기 위한 도면이다. 도 1을 참조하여, 테스트 장치가 요구사항에 따른 리스크 노출 값을 측정하고, 측정된 리스크 노출 값을 기초로 테스트 케이스 우선순위를 결정하는 방법을 설명한다. 리스크는, 손실이나 피해가 있을 가능성(possibility)으로 정의된다. Boehm은 리스크 노출(Risk Exposure, RE)을 정의하였는데, 이는 리스크의 가능성(risk probability 또는 risk likelihood)과 리스크가 발생했을 때의 손실(risk impact)의 곱으로 나타낸다. 높은 리스크 노출 값을 가지는 소프트웨어 영역에 대해서 더 많은 시간을 투자하라는 것이 리스크 기반 테스트(Risk-based Testing, RBT)이다. 개시된 기술에서는 제품 리스크 항목을 정의하고, 요구사항으로부터 추출한 리스크 노출 값을 측정하여 테스트 케이스 우선순위를 결정하는 방법을 제안한다.
- [0018] S110 단계에서, 테스트 장치는 리스크 항목을 선정한다. 리스크는 보통 프로젝트 리스크, 프로세스 리스크, 제품의 요구사항과 관련된 제품 리스크로 구분한다. 본 실시예에서는 제품 품질을 확인(validate)하기 위하여 필요한 테스트 케이스의 우선순위 측정 과정을 예를 들어 설명한다. 제품 리스크 항목은 여러 객관적 문서들을 통해 도출된다. 제품 리스크 항목에 대하여 IEEE1044.1-1995(Classification for Software Anomalies), Wallace, Sullivan, Jha, J.Bach, Soila에서 제시하는 리스크 항목이 참조된다. IEEE Std 1044.1-1995는 소프트웨어 예외(Anomalies)에 관련한 표준으로, 사용자에게 관련한 증상(symptoms) 및 타입(Type) 리스크를 참고하였다. Wallace에서는 의학 장비의 소프트웨어 실패에 대한 리스크 항목을 참고하였다. Sullivan에서는 데이터베이스 관련 에러 타입(error type), 에러 트리거(error trigger), 결함 타입(defect type) 리스크 항목을

참고하였다. Jha는 모바일 어플리케이션에서 알려진 문제나 가능성 있는 문제들에 대한 리스크 카탈로그(risk catalog)를 Bach's Heuristic Test Strategy Model(Bach, 2006a)를 기반으로 도출하였다. 리스크 카탈로그는 크게 제품(Product Elements), 작동(Operational Criteria), 개발(Development Criteria), 프로젝트 환경(Project Environment)으로 구분되는데, 이 가운데 제품과 관련 리스크 항목을 참고하였다. J. Bach 는 경험적인 분석(Heuristic Analysis)을 통해 분석한 리스크 카탈로그(risk catalog)를 참고하였다. Soila에서는 웹 어플리케이션에서의 소프트웨어 실패 관련 리스크 항목 참고하였다.

[0019] 개시된 기술에서는 상기 리스크 항목들을 참조하여, 제품 리스크 항목을 도 2와 같이 5개의 카테고리로 분류하였다. 도 2의 카테고리를 참조하여 제품 리스크 항목을 선정할 수 있다. 즉, 개시된 기술에서는 상기 리스크 항목들을 참고하여 도 2와 같이 소프트웨어, 데이터, 인터페이스, 갱신, 플랫폼 리스크 군과 그 아래 세부 리스크 항목으로 정리하여 이를 기반으로 제품 리스크 항목을 식별하고, 테스트 케이스 우선 순위를 결정하는 방법을 개발한다. 리스크 항목이 정해지면, 테스트 장치는 S120 단계 및 S130 단계를 통하여 요구사항으로부터 리스크 노출값을 측정한다.

[0020] 우선, S120 단계에서 테스트 장치는 요구사항의 리스크 가중치를 결정한다. 테스트 장치는 요구사항의 각 항목에 대하여 요구사항이 실패하는 리스크 가능성(Threat Likelihood) 과 요구사항이 실패할 때의 영향력(Impact) 측면에서의 가중치를 곱하여 요구사항의 리스크 가중치 $WReq_i$ 를 결정한다. 수학적 1은 요구사항의 리스크 가중치를 구하는 수식을 나타낸다.

수학적 1

$$RE(Req_i) = WL_i * WI_i * 10000$$

[0022] 여기서, Req_i 은 요구사항 i , WL_i 는 요구사항 i 의 리스크의 가능성 가중치, WI_i 는 요구사항 i 의 리스크의 영향력 가중치를 의미한다. 요구사항의 리스크 가능성 및 영향력 측면에서의 상대적 중요도를 산출하기 위해서 AHP(Analytic Hierarchy Process) 기법이 적용된다. AHP 는 의사결정에 영향을 미치는 요인(속성, 특성, 또는 기준)들을 선정하고 점차로 작은 요인으로 분화하여 계층을 구성하고 이러한 요인들의 중요도를 이원비교(pairwise comparison)를 통하여 가중치(우선순위)를 일관성 있게 결정하는 의사결정 과정이다. AHP는 단위와 무관하게 서로 다른 특성의 비교가 가능하며, 두 특성 간의 상호비교를 통하여 선호도만을 판단하면 되므로 요구사항 항목을 리스크 관점에서 우선순위를 산정하는 데 적합하다. 요구사항을 분석하는 사람마다 상대적인 중요도를 다르게 부여할 수 있는데, 이를 조정하기 위해 요구사항 분석가가 부여한 값의 평균값을 제시하고, 요구사항 항목 간의 리스크 측면에서의 이원비교 값을 확정한다

[0023] 예를 들어 print_tokens는 총 18 개의 요구사항 항목이 있다. 도 3은 print_tokens의 요구사항과 각 요구사항에 대한 리스크 가중치를 나타낸다. 요구사항의 리스크 가능성과 리스크 영향력은 AHP 기법을 적용하여 측정된 수치이다. 첫 번째 요구사항에 대한 리스크의 가능성, 영향력은 AHP를 적용하여 산출한 값으로써 각각 0.080, 0.102이며, 리스크 가중치 $RE(Req_1)$ 은 수학적 1에 따르면 82이다.

[0024] S130 단계에서, 테스트 장치는 리스크 항목별로 요구사항의 리스크 가중치를 반영하여 리스크 노출 값을 측정한다. 도 3은 리스크 항목의 리스크 노출 값 측정 메트릭이다. RM_{ij} 는 리스크 항목 r_i 에 요구사항 항목 j 가 매칭되면 1, 아니면 0 을 나타낸다. $RE(r_i)$ 는 리스크 항목 r_i 에 대한 리스크 노출값을 뜻한다. $RE(r_i)$ 값은 수학적 2와 같이 측정될 수 있다.

수학적 2

$$RE(r_i) = \sum_{j=1}^n (RE(Req_j) \times RM_{ij})$$

[0026] 도 4는 리스크 노출값을 측정하는 데 사용되는 리스크 노출 메트릭을 나타낸 표이다. 도 5는 도 4의 리스크 노출 메트릭을 참조로 산출된 print_tokens에 대한 리스크 노출 측정 값이다. 예를 들어 print_tokens의 도 4의 첫 번째 요구사항, R_1 은 소프트웨어에 관한 리스크 군 가운데 입력문제, 출력문제, 시작/정지, 에러 핸들링 리

스크 등이 매칭되어 있음을 알 수 있다. 리스크 항목의 리스크 노출값은 수학적 2를 이용하여 계산될 수 있는데, 입력 문제 리스크에 대한 리스크 노출 값은 수학적 2에 의해 $(82*1) + (68*1) + \dots + (14*0) = 150$ 이 된다.

- [0027] S140 단계에서, 테스트 장치는 테스트 케이스의 우선순위를 결정한다. 테스트 장치는 리스크 항목의 리스크 노출 값을 이용하여 테스트 케이스 우선순위(Test Case Priority, TCP)를 측정한다. 도 6은 테스트 케이스의 우선순위를 측정하는데 사용되는 테스트 케이스 우선순위 측정 메트릭을 나타낸 표이다. TM_{ij} 는 테스트 케이스 $i(TC_i)$ 로 실행했을 때, 리스크 항목 $j(R_j)$ 에 해당하는 부분의 발생 횟수를 나타낸다. TCP는 테스트 케이스가 리스크 항목에 해당하는 부분을 얼마나 많이 실행하는가를 나타내는 메트릭으로써, i 번째 테스트 케이스 TC_i 에 대한 TCP_i 는 수학적 3과 같다.

수학적 3

[0028]
$$TCP_i = \sum_{j=1}^n (RE(r_j) \times TM_{ij})$$

- [0029] 테스트 장치는 테스트 케이스 우선순위를 나타내는 TCP 값이 높은 순서대로 테스트 케이스를 선정한다. 예를 들어 도 7은 print_tokens에 대한 테스트 케이스 우선순위를 측정한 예의 일부이다. 도 7을 참조하면, 테스트 케이스가 총 4130 개가 있다. 도 7에서는 테스트 케이스들이 TCP 값이 가장 높은 순으로 정렬되어있다. TC_1 을 살펴보면, 출력문제에 관한 리스크를 4 회, 계산 리스크를 7 회, 시작/정지 리스크를 3 회, 에러 핸들링 리스크를 9 회, 상호작용 리스크를 50 회를 실행하게 되며, TCP_1 은 수학적 3에 따라 $(150*0)+(182*4) + (191*7) + (196*3) + (228*9) + (265*50) + \dots = 25477$ 이 된다.

- [0030] 개시된 기술의 일 실시예에 따른 테스트 케이스 우선순위 결정방법의 효과를 확인하기 위한 실험을 수행하였다. 실험 대상 프로그램은 테스트 커버리지와 테스트 케이스의 우선순위 기법의 비교 평가 실험에 실험 대상 프로그램으로 널리 사용되고 있는 Siemens 프로그램으로 선정하였다. 개시된 기술과 비교 대상이 된 종래의 테스트 케이스 우선순위 결정 기법은 No prioritization(테스트 케이스 작성 순서대로 선정), Safety Tests(테스트 케이스의 리스크 노출 값이 높은 순서대로 선정), Optimal prioritization(결함 발견율을 높일 수 있는 테스트 케이스 순서대로 선정), FEP-Total(뮤테이션 분석을 통한 결함 존재 가능성이 높은 순서대로 선정) 기법들을 포함한다. 개시된 기술에 따른 우선순위 결정 기법은 $RE(r_i)$ 로 표시한다.

- [0031] 테스트의 목적은 가장 심각한 결함을 먼저 찾고자 하며, 릴리즈 전에 결함이 고쳐지기를 원하며, 가능한 많은 결함을 찾기를 바란다. 결국, 결함을 얼마나 빨리 찾는 지와 이상이면 제품의 심각한 결함이 있는 곳을 얼마나 빨리 찾는지가 중요한 목적이 된다고 할 수 있다. 개시된 기술에서 제안한 기법이 결함 검출 면에서 타 기법과 비교할 때 얼마나 우수한가를 판단하기 위하여, 결함발견빈도수와 평균결함검출률에 근거한 결함 발견율 관점 및 결함 심각도 관점에서 실험 결과를 살펴본다.

- [0032] 첫 번째로 테스트 케이스 우선순위기법을 적용하여 결정한 테스트 케이스의 순서로 테스트했을 때 발견되는 결함에 관하여 실험 결과를 분석한다. 이를 위하여 각 테스트케이스에 대한 결함 발견 빈도수를 측정한다. 그런 후 Average Percentage of Fault Detected(APFD)를 측정하여 타 기법과 비교한다. 테스트 되는 프로그램에 포함된 결함의 수를 m , 전체 테스트 케이스 수를 n , 테스트 케이스 순서 풀에서 결함 i 를 노출시키는 첫 번째 테스트 케이스를 TF_i 라 할 때, APFD는 수학적 4와 같이 계산한다.

수학적 4

[0033]
$$APFD = 1 - \frac{\sum_{i=1}^m TF_i}{n*m} + \frac{1}{2n}$$

- [0034] 두 번째로는 발견되는 결함의 심각도 측면에서 실험 결과를 분석한다. 결함의 종류에는 시스템을 다운시키거나 주요 기능을 동작하지 못하게 하는 심각한 결함이 있을 수 있고, 단지 시스템의 속도를 느리게 하는 결함이 존

제할 수 있다. 주어진 기간 내에 테스트를 수행할 경우 심각한 결함을 먼저 발견할 수 있도록 테스트가 수행된다면 매우 효과적일 것이다. 본 실험에서는 결함에 대하여 AHP 기법을 적용하여 결함의 비중을 측정하고, 수학적 5를 이용하여 테스트 케이스가 발견한 결함들의 평균 심각도를 측정한다.

수학적 5

$$\text{평균 결함 심각도} = \frac{\text{발견되지 않는 faults severity의 합} - \text{발견된 faults severity의 합}}{\text{발견되지 않은 fault severity의 합}} * 100$$

[0035]

[0036] 테스트케이스 우선순위 기법을 적용한다는 의미는 현재 알지 못하는 결함을 더 효과적으로 발견할 수 있는 테스트 케이스 풀을 선정하도록 하는 것이다. 본 실험을 위하여 알고 있는 결함 정보를 다양하게 변화를 주도록 실험 설계를 한다. 예를 들어 Siemens program의 print_tokens의 경우 결함 버전이 총 7 개이다. 가능한 실험 방법은 두 개의 결함 버전 1 과 2 에 대한 결함에 대한 정보를 이용하여 테스트 케이스 우선순위를 측정하는데 이용하고, 나머지 5 개의 결함 버전 들의 결함을 발견하는지를 측정하게 되는 것이다. 7 개 결함 버전 가운데 무작위로 2 개를 결함정보를 아는 버전으로써 택하도록 하여 총 7 셋트 구성하였다. 아는 결함 버전 개수를 3, 4, 5 개까지 확대하여 각각 7 셋트씩 구성하여 print_tokens의 경우 총 28 개의 실험세트를 구성하여 실험을 수행하였다. 나머지 실험 대상 프로그램에 대한 실험 셋트도 이와 동일한 방법으로 구성하여 총 947 개 실험 세트에 대하여 본 실험을 수행하였다.

[0037] 결함 발견 빈도수 관점에서의 실험 결과, 즉, 테스트 케이스 우선 순위가 높은 순으로 발견된 결함 수 현황은 다음과 같다. 빠른 시간 안에 많은 결함을 발견하는지 여부를 측정하기 위해서는 가장 처음 실행하는 테스트 케이스 그룹이 결함을 발견할 수 있는 테스트 케이스 개수가 높으면 테스트 효과성이 높다는 것을 보여주는 것이다. 총 947 개 실험 세트에 대하여 비교대상 테스트 우선 순위 선정 기법을 적용했을 때 발견한 결함 수를 도 8의 그래프로 표시하였다. 도 8에서, x 축은 테스트 케이스 우선순위가 높은 순서로 된 테스트 케이스며, y 축은 테스트 케이스가 발견한 결함 개수이다. 첫 번째 그래프는 Siemens 전체 프로그램의 결함 개수를 나타낸 그래프이고, 나머지는 각 프로그램 별로 결함 개수를 나타낸 그래프이다. 그래프를 살펴보면, 개시된 기술에서 제안한 방법인 RE(ri)가 다른 기법에 비해서 많은 결함을 먼저 발견하는 것을 알 수 있다.

[0038] 도 9는 각 기법에 의해 선정된 테스트 케이스에 대하여 평균 결함 검출률 (APFD) 값을 나타낸 표이다. 도 10은 테스트 진행률에 따른 결함 발견률을 도식화 한 것으로 X 축은 테스트케이스 수행률, Y 축은 결함 발견률을 나타낸다. 굵은 실선이 RE(ri)에 해당한다. RE(ri) 기법은 Optimal prioritization 기법보다는 APFD 가 낮다. 하지만, 결함의 정보를 이용하여 테스트 케이스 우선순위를 측정한 Safety Tests, FET-total 에 비교했을 때 해서도, 결함 정보 없이 리스크 카테고리의 노출 값만을 이용함에도 RE(ri)의 결함 발견율이 좋은 편임을 알 수 있다. 도 11은 함수 별로 기법의 APFD 값에 대한 box-plot 그래프이다. X 축은 테스트케이스 우선순위 선정 기법이고, Y 축은 결함 발견률이다. 우선 순위 상위 30%까지의 테스트 케이스의 오류발견을 범위를 아래 짙은 박스로, 50%까지는 옅은색 박스로, 70%까지는 하얀 박스로 오류발견을 범위를 나타낸다. 중간 선의 아래는 테스트 케이스 한 개로 수행했을 때의 Minimum 결함발견율을 나타내며, 위는 Maximum 결함 발견율을 나타낸다. 예를 들어 첫 번째 그래프는 전체 함수에 대한 평균에 대한 것인데, 이에 대한 상세한 데이터를 표로 다시 나타내면 도 12와 같다. Min의 결함 발견율은 Optimal prioritization 기법 다음으로 RE(ri)가 19%로 높다. 상위 30%의 테스트 케이스들로는 RE(ri)의 기법이 높은 발견율을 보이는 것을 확인할 수 있다. 테스트 수행 초기에 결함 발견율이 높음을 의미한다.

[0039] 도 13은 수학적 5에 의한 심각한 결함의 발견 정도를 나타내는 그래프이다. X 축은 테스트 케이스 우선순위에, Y 축은 x까지의 테스트 케이스들에 의해 발견하지 못한 결함들의 심각도를 나타낸다. 결함 심각도의 기울기는 심각한 결함이 얼마나 남아있는가를 의미하는 것이다. 즉 기울기가 낮을수록 심각한 결함이 적게 남아 있음을 의미한다. 첫 번째 그래프는 Siemens의 전체 프로그램에 대한 심각도를 나타낸 그래프이다. RE(ri)의 그래프가 굵은 선으로 나타난 것인데, 다른 기법의 그래프에 비해서 기울기가 낮은 것을 확인할 수 있다. 즉 테스트 초기에 심각한 결함을 많이 발견할 수 있음을 뜻한다. 또한 RE(ri)의 그래프는 모든 프로그램에 대하여 기울기가 일정한 것을 확인할 수 있다. 따라서 RE(ri) 그래프의 기울기를 분석하면 테스트 종료시점을 어느

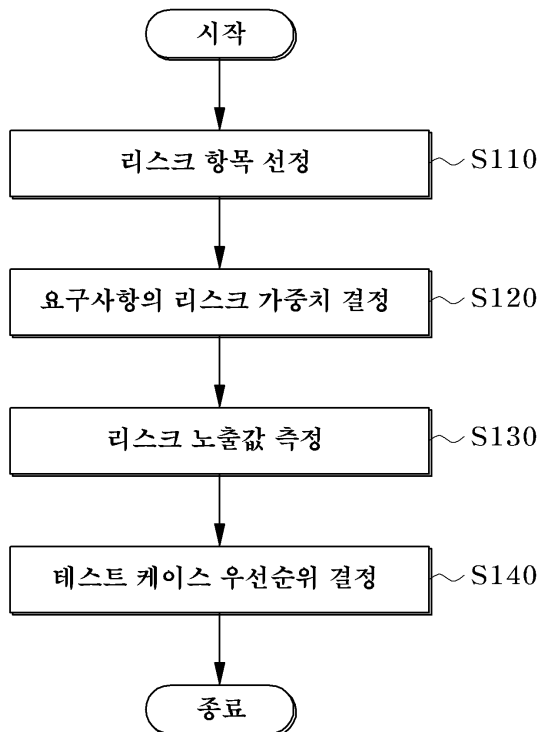
정도 예측할 수 있는 장점이 있다.

[0040] 개시된 기술에서는 요구사항 항목마다 산출한 리스크 노출값을 활용하여 테스트 케이스 우선순위를 부여하는 기술을 제안하고, 실험을 통해 기존 방법들과 비교 분석하여 개시된 기술에서 제안하는 우선순위 부여 기술의 특징을 설명하였다. 개시된 기술의 특징은 다음과 같다. 첫째, 본 명세서에서 제안한 기술은 기존의 방법들이 활용하는 선행된 테스트 결과를 필요로 하지 않는다. 대신, 리스크 항목 평가 메트릭을 개발하여 활용한다. 요구사항에 가중치를 부여하고 이를 활용하여 리스크 노출값을 측정하는 방법을 개발하였고, 나아가 리스크 노출값을 기반으로 테스트 케이스 우선순위 부여 메트릭을 개발하였다. 이는 선행 테스트 수행 결과 없이 실행 가능한 방법으로서, 그 활용도가 높을 것으로 기대한다. 둘째, 제품 리스크 항목을 구체적으로 정의하여 활용하였다. 제품 리스크 항목 정의를 위하여 IEEE std 1044.1-1995를 비롯한 의미 있는 문서들을 참조하였고, 이를 기반으로 요구사항의 가중치를 평가하기 위한 메트릭도 개발하였다. 셋째, 제안한 테스트 케이스 우선순위 부여 기술을 결함 발견률 관점에서 결함 발견 빈도수와 평균 결함 검출률 (Average Percentage of Fault Detection : APFD), 그리고 결함 심각도로 분석하였다. 실험은 Siemens Program을 대상으로 이루어졌다. 개시된 기술의 우선순위에 따른 테스트 실행 결과, 심각한 결함에 대한 검출 효과성이 나타났으며, 동시에 소요 시간 및 비용 면에서의 효과성도 상대적으로 높게 측정되었다.

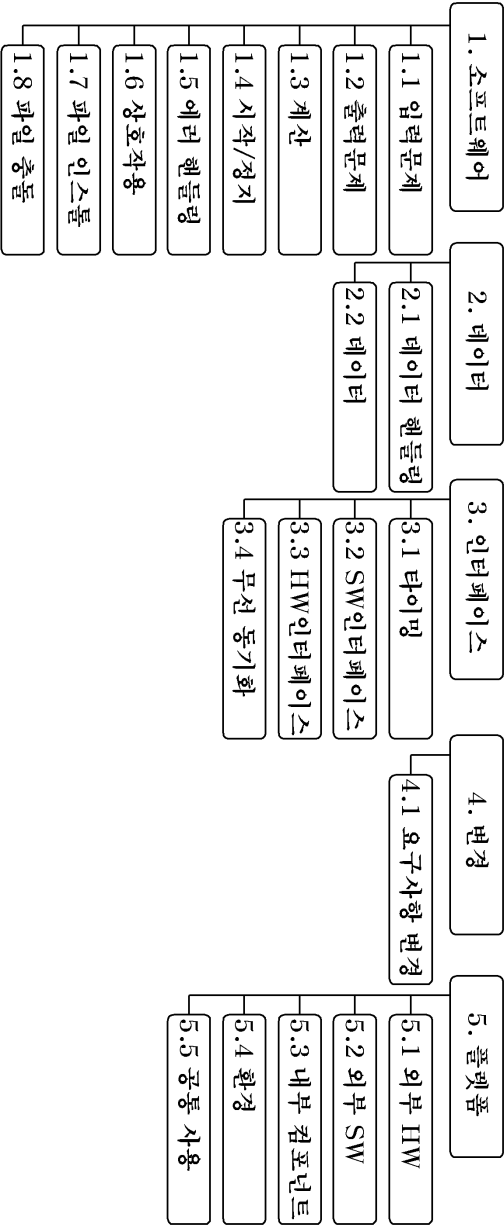
[0041] 이러한 개시된 기술인 시스템 및 장치는 이해를 돕기 위하여 도면에 도시된 실시 예를 참고로 설명되었으나, 이는 예시적인 것에 불과하며, 당해 분야에서 통상적 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시 예가 가능하다는 점을 이해할 것이다. 따라서, 개시된 기술의 진정한 기술적 보호범위는 첨부된 특허청구범위에 의해 정해져야 할 것이다.

도면

도면1



도면2



도면3

요구사항 Req _i		WL _i	WI _i	RE(Req _i)
R ₁	어휘 분석을 위한 입력이 되어야 하며, 잘못된 문자를 입력했을 때 에러 메시지를 출력하고 종료된다.	0.080	0.102	82
R ₂	입력된 파일 이름이 존재하지 않을 때, 에러 메시지를 출력하고 종료된다.	0.073	0.094	68
R ₃	입력 값이 공백이거나 다음 라인에 입력 문자가 있을 경우에 문자의 끝인지 여부를 측정한다.	0.066	0.094	62
R ₄	입력된 문자의 끝인지 여부를 측정하여 문자의 길이를 측정한다.	0.0354	0.0500	18
R ₅	읽어들인 문자의 위치를 알 수 있어야 한다.	0.0354	0.0560	20
R ₆	입력된 파일 이름에 대한 크기에 대한 정보를 알 수 있어야 한다.	0.0460	0.0690	32
R ₇	읽어들인 문자에 대한 토큰을 찾는다. 이때 문자의 끝인지 여부를 측정하여 토큰을 구분한다.	0.0730	0.0980	72
R ₈	토큰이 숫자인지 여부를 측정한다.	0.0210	0.0300	6
R ₉	문자의 끝인지를 측정하고 실제 토큰을 찾아낸다.	0.0489	0.0200	10
R ₁₀	알파벳과 숫자가 연달아 입력 되었는지 여부를 측정한다.	0.1620	0.0400	65
R ₁₁	키워드를 찾아내고 키워드가 아닌 경우에는 에러메시지를 출력한다.	0.0386	0.0360	0.0014
R ₁₂	특별 문자들을 찾아내고 특별 문자에 해당되지 않을 경우 에러메시지를 출력한다.	0.0386	0.0370	14
R ₁₃	문자의 끝에 도달하기 전까지 공백과 같은 문자들을 삭제한다.	0.0247	0.0310	8

도면4

Risk Item	요구사항					RE(ri)
	Req ₁		Req _i		Req _n	
	WL ₁		WL _i		WL _n	
	WI ₁		WI _i		WI _n	
	RE(Req ₁)		RE(Req _i)		RE(Req _n)	
r ₁						RE(r ₁)
r _j			RM _{ij}			RE(r _i)
r _m						RE(r _m)

도면5

Risk Item ri		요구사항 risk exposure				RE(ri)
		Req ₁	Req ₂	...	Req ₁₃	
		0.080	0.073		0.0412	
		0.102	0.094		0.0340	
		82	68	...	14	
1.소프트웨어	입력 문제	1	1	...	0	150
	출력 문제	1	1	...	0	182
	계산	0	0	...	1	191
	시작/정지	1	1	...	0	196
	에러 핸들링	1	1	...	0	228
	상호작용	0	0	...	0	365
	파일 인스톨	0	0	...	0	0
	파일 충돌	0	0	...	0	0
...	...	...	...	...	...	...

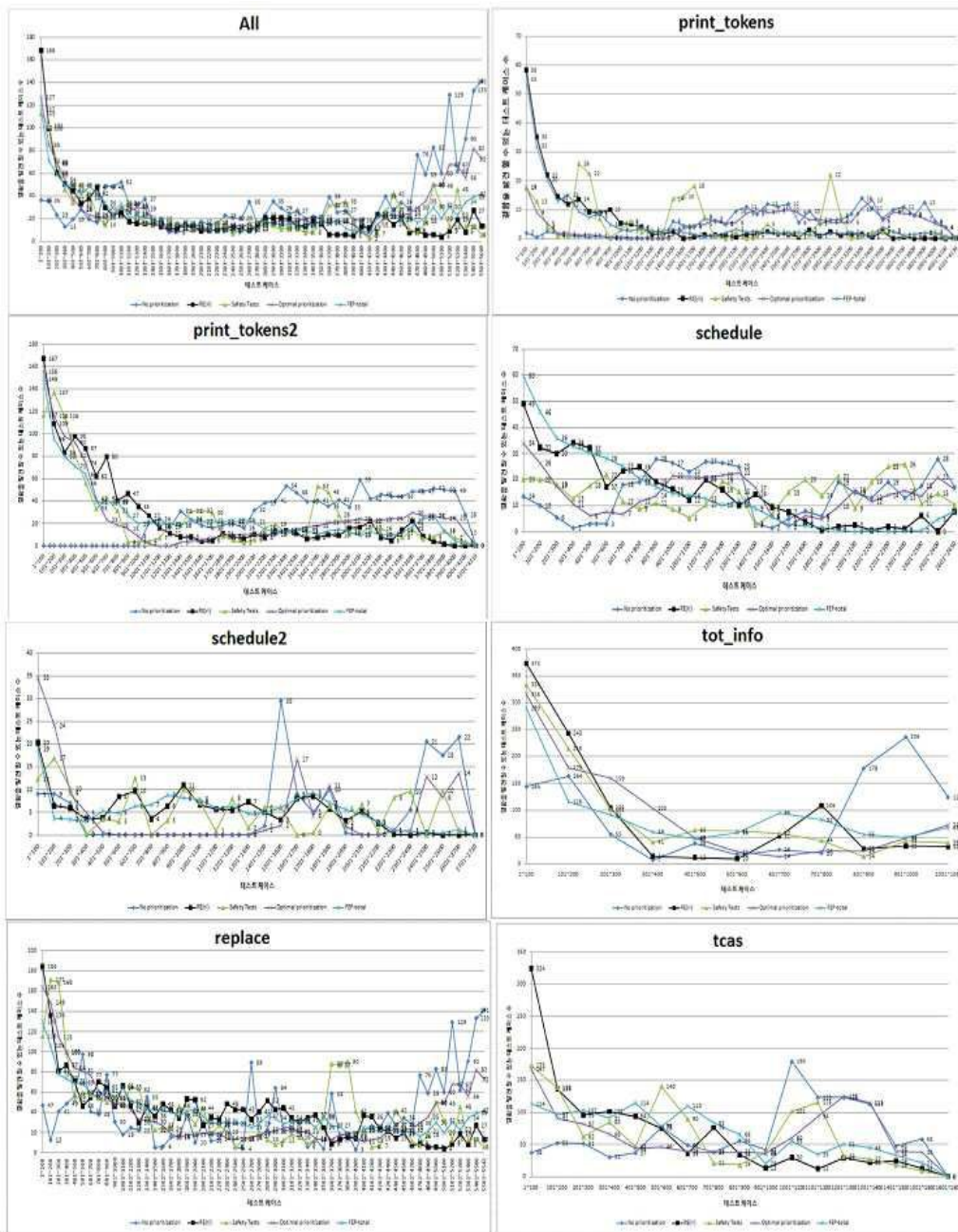
도면6

Risk Item	RE(ri)	Test Case				
		TC ₁		TC _i		TC _n
r ₁	RE(r ₁)					
r _j	RE(r _i)			TM _{ij}		
r _m	RE(r _m)					
TCP		TCP ₁		TCP _i		TCP _n

도면7

Risk Item		RE(ri)	Test Case			
			TC ₁	TC ₂	...	TC ₄₁₃₀
1.소프트웨어	입력 문제	150	0	0	...	1
	출력 문제	182	4	2	...	1
	계산	191	7	4	...	0
	시작/정지	196	3	3	...	1
	에러 핸들링	228	9	2	...	1
	상호작용	365	50	16	...	0
	파일 인스톨	0	0	0	...	0
	파일 충돌	0	0	0	...	0
...	...	...	...	...	...	...
TCP			25477	8494	...	756

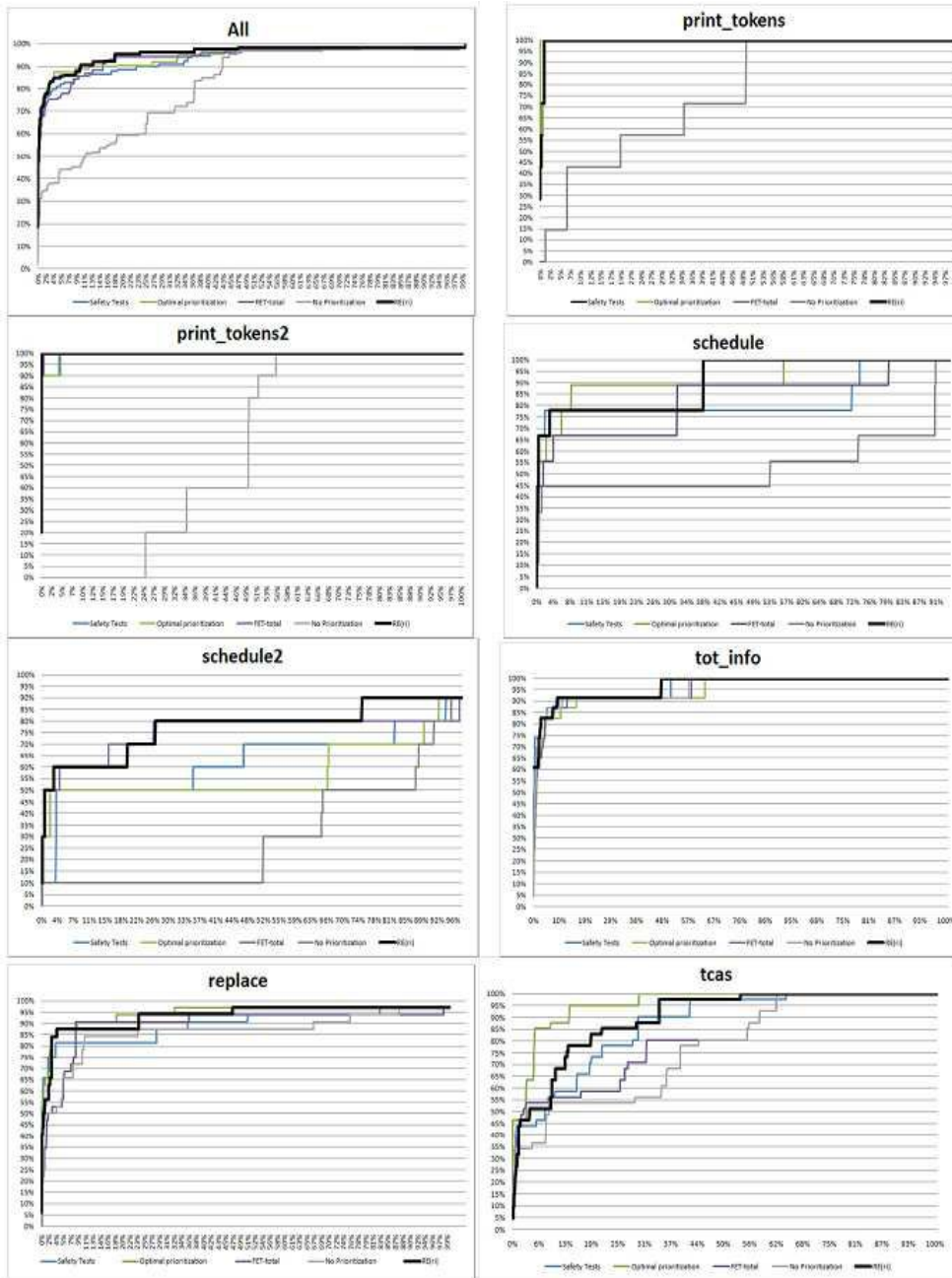
도면8



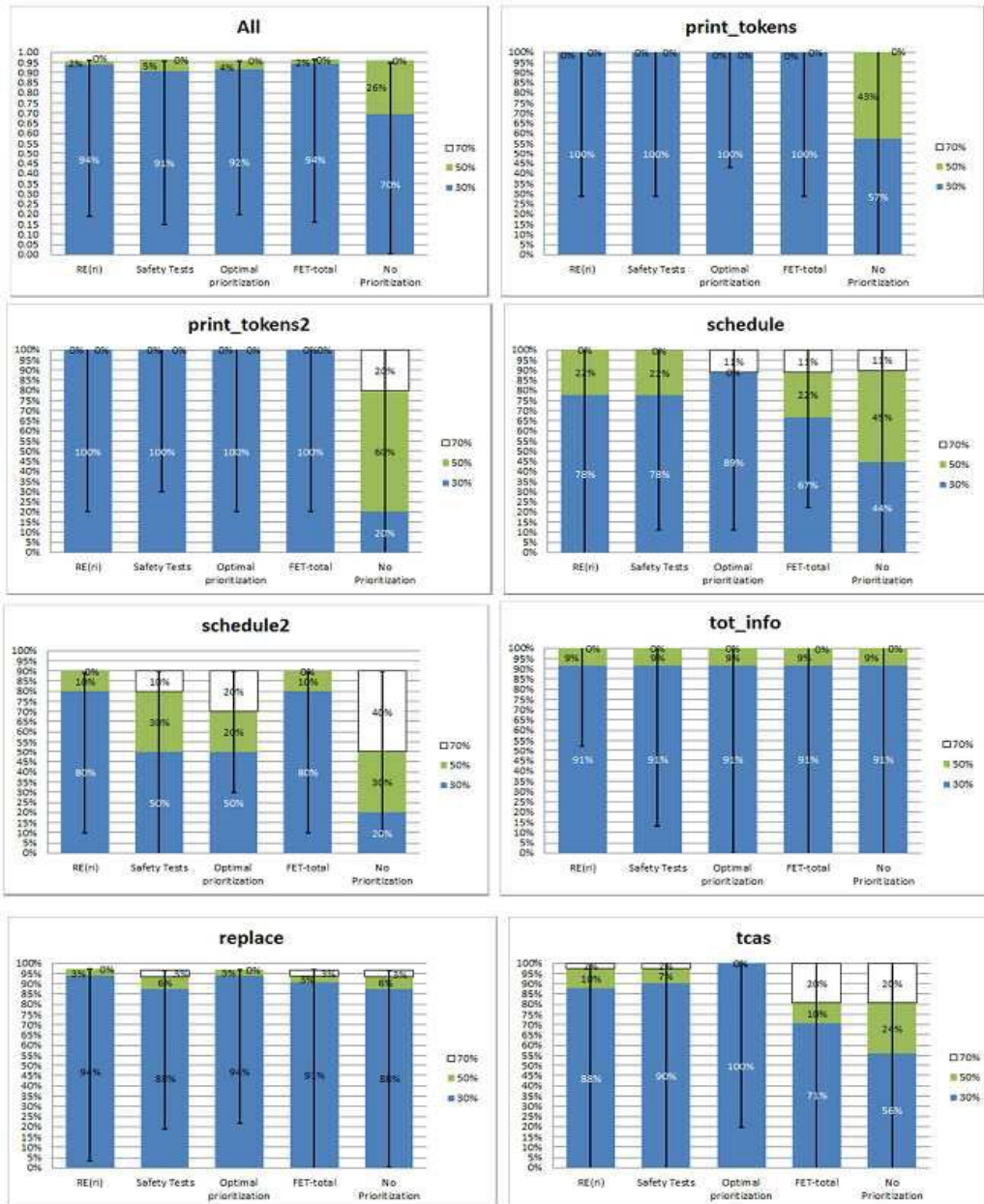
도면9

	No prioritization	RE(ii)	Safety Tests	Optimal prioritization	FET-total
print_tokens	76	99	99	99	91
print_tokens2	57	99	99	99	99
Schedule	56	98	86	99	99
schedule2	30	77	63	90	75
tot_info	94	95	95	92	93
Tcas	90	88	87	96	81
Replace	89	89	89	89	90
평균	70	92	88	95	90

도면10



도면11



도면12

	No prioritization	RE(i)	Safety Tests	Optimal prioritization	FET-total
Maximum	0.97	0.97	0.97	0.98	0.97
70%	0.96	0.96	0.96	0.96	0.96
50%	0.96	0.95	0.96	0.96	0.96
30%	0.70	0.94	0.91	0.94	0.92
Minimum	0.01	0.19	0.17	0.20	0.16

도면13

