

S/W 벤치마크테스트 평가모델 개발 이슈

박진홍*, 정영은*, 신석규*

*한국정보통신기술협회 시험인증연구소 S/W시험인증센터

The Development Issue of S/W Benchmark Test Evaluation Model

Jin-Hong Park*, Yung-Eun Jung*, Seok-Kyoo Shin*

*TTA Software Quality & Evaluation Center

E-mail : elzk@tta.or.kr, yejung@tta.or.kr, skshin@tta.or.kr

요약

국내 S/W 벤치마크테스트는 일반적으로 각 제품의 기능별, 성능별 비교분석을 통해 제품의 장단점을 파악하는 것이 주된 목적이었다. 기존 벤치마크테스트의 산출물은 시험 항목별 결과와 비교 분석 결과이다. 최근에는 동종 제품들 중에서 특정 목적에 가장 적합한 제품을 선택하기 위한 수단으로 벤치마크테스트 결과가 사용되기도 한다. 이러한 특수한 목적의 벤치마크테스트는 기존 산출물 이외에 특정 목적에 가장 적합한 제품이 무엇인지 판단할 수 있는 순위표가 추가된다. 최근의 새로운 요구사항을 반영하기 위해서는 새로운 S/W 벤치마크테스트 평가모델이 필요하다. 본 논문에서는 기존의 벤치마크테스트와 구분되는 소비자의 새로운 요구를 만족시킬 수 있는 S/W 벤치마크테스트 평가모델 개발 이슈에 대해 연구하고 있다.

1. 서론

국내 S/W 산업이 점차 성숙해감에 따라 유사한 제품의 개발이 많아지고 점점 군을 형성해 가고 있는 추세이다. 다양한 동종 제품의 개발로 인해 사용자는 자신의 사용 목적에 보다 적합한 제품을 선택할 수 있게 되었다. 그러나 많은 동종 제품 가운데 사용자 자신이 다양한 제품을 직접 사용해보지 않고 용도에 맞는 최적의 제품을 선택하는 것은 쉬운 일이 아니다. 이때 객관적이고 공정한 S/W 벤치마크테스트 결과를 확보할 수 있다면 보다 쉽게 좋은 제품을 선택할 수 있을 것이다.

구매자 입장에서 S/W 벤치마크테스트는 객관적인 테스트 결과를 통해 각 제품의 장단점을 올바르게 파악할 수 있게 하여 구매 목적에 보다 근접한 제품을 선택하는데 유용한 지침이 될 수 있다. 반면 개발사 입장에서 S/W 벤치마크테스트는 동종 제품들 간의 객관적인 비교 평가를 통해 각

제품의 강점을 파악하고, 취약점 보완 및 개선방안을 제시할 수 있으므로 S/W 품질 향상에 큰 도움을 줄 수 있다.

또 다른 목적의 벤치마크테스트는 유사 제품들 간의 기능 및 성능의 우선순위를 결정하는데 목적을 둘 수도 있다. 일례로 공공기관 및 대규모 조직에서 도입할 제품을 선정하고자 할 때, 중립의 위치에 있는 전문 기관에 제품의 우선순위를 결정하기 위한 벤치마크 테스트를 의뢰하는 경우도 최근 들어 많은 수요가 발생하고 있다.

기존의 비교 분석을 위한 벤치마크 테스트는 산출물로 각 평가 요소별 평가 값과 비교 분석서를 산출하지만, 요구에 의한 우선순위를 결정하는 벤치마크테스트는 부가적으로 목적 평가 기준에 기반한 제품별 우선순위를 산출하게 된다. 이러한 순위 산출을 위한 벤치마크테스트는 무엇보다 기본 목적이 명확해야 하며, 이러한 벤치마크테스트는 시험에 응하는 제품을 대부분 수용할 수 있

어야 한다.

비교 분석을 위한 벤치마크테스트는 시험 제품을 선정함에 있어 LINUX OS들에 대한 벤치마크 테스트, 백신 S/W들 간의 벤치마크테스트 등과 같이 동종의 국한된 S/W를 대상으로 시험을 진행하는 반면, 순위를 결정하는 벤치마크 테스트는 벤치마크 테스트의 목적에 부합하는 어떠한 제품도 시험에 응할 수 있다.

이러한 소비자의 새로운 요구사항을 반영하기 위해서는 기존의 벤치마크테스트 평가모델과 구별되는 특수 요소가 반영된 새로운 S/W 벤치마크테스트 평가모델의 개발이 절실히 필요하다.

본 연구에서는 이러한 요구사항을 만족시킬 수 있는 벤치마크테스트 평가모델 개발 및 프로세스에 대해 연구하고 있으며, 향후 산업계에서 벤치마크테스트 수행 시 지침으로 활용할 수 있도록 하고자 한다.

2. 프로세스 정의

효율적인 벤치마크테스트 수행을 위해서는 벤치마크테스트 프로세스 정의가 선행되어야 한다. [그림 1]은 일반적인 벤치마크테스트 수행을 위한 프로세스를 단계, 활동, 작업으로 구분하여 도식화한 것이다.

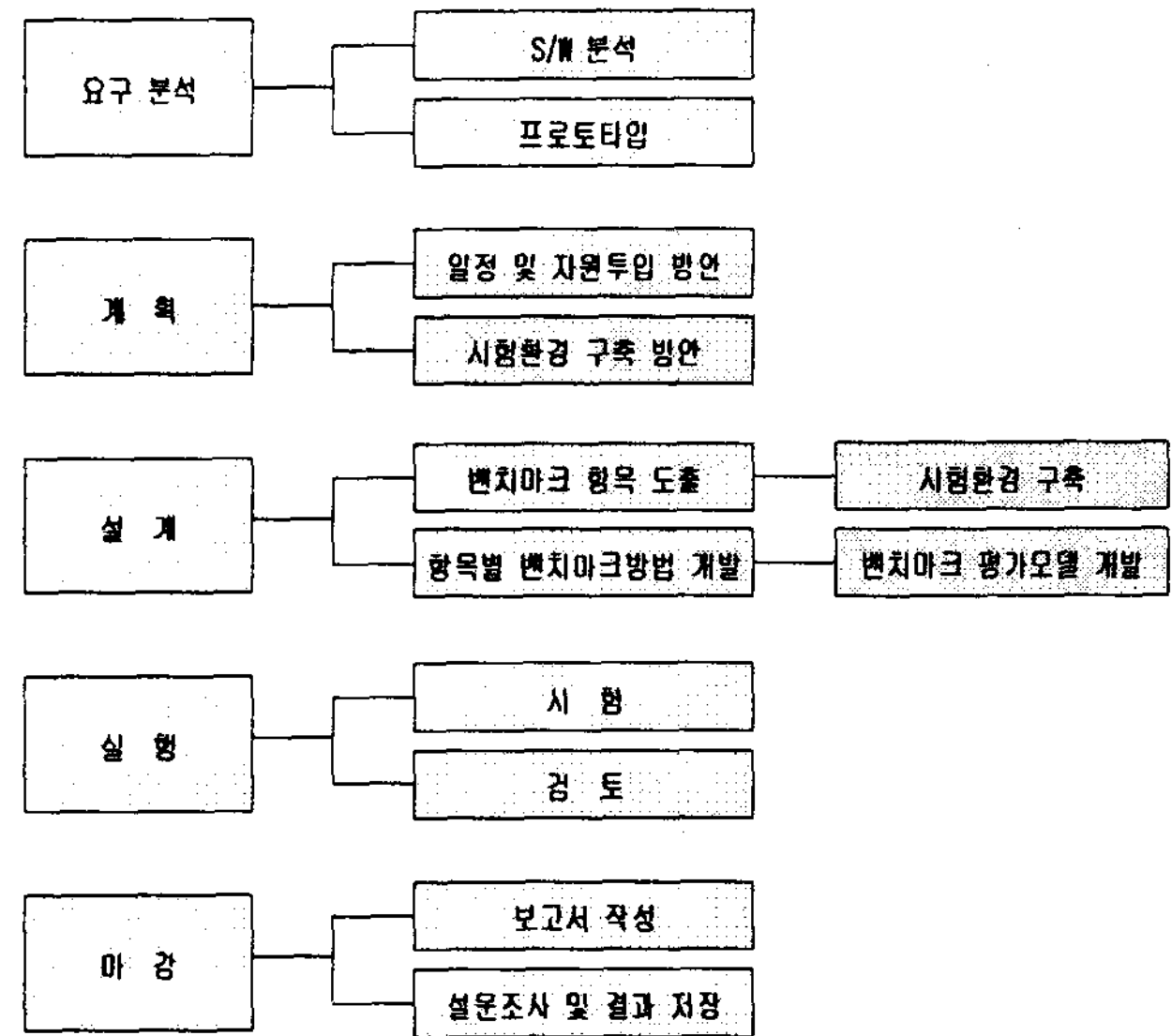
2.1 벤치마크테스트 프로세스

2.1.1 요구분석 단계

시험 대상 항목과 벤치마크테스트 항목 리스트를 도출하기 위해 대상 S/W에 대한 철저한 분석을 실시한다. 분석 이후 항목 도출 시에는 벤치마크테스트에 참여하는 제품의 업체 개발자 및 담당자들과의 미팅을 통해 도출 항목에 대해 합의하고, 실제 벤치마크테스트 가능 여부를 검증하기 위해 프로토타입 단계를 포함하여 설계단계에 앞서 준비해야 할 조건을 파악한다.

2.1.2 계획단계

요구분석 단계에서 도출한 결과물들을 토대로, 일정 및 자원 투입 계획을 구체적으로 산정한다.



[그림 1] S/W 벤치마크테스트 프로세스

별도의 교육이 필요한 경우 교육계획을 수립하며, 전문가 그룹을 구성하도록 한다. 시험환경 구축을 위한 환경설정과 시험도구 및 S/W 선정도 계획단계에서 이루어진다.

2.1.3 설계 단계

계획 수립 이후 도메인별 전문가 그룹을 구성하여 시험 대상 주요 항목 및 부가항목을 선정한다. 벤치마크테스트 항목 확정 이후 항목별 측정기준 및 공식을 도출하여 메트릭을 작성하고, 벤치마크테스트에 참여하는 업체들과 항목에 대한 합의를 진행한 후 벤치마크테스트 평가모델을 완성한다.

4) 실행 단계

개발된 평가모델의 각 항목 별 벤치마크테스트를 수행한다. 비교 대상 제품별로 동일한 조건에서 시험을 실시하고 시험 결과가 도출되면 결과에 대한 문제점을 정리 검토 후, 필요한 경우 보완 시험을 수행하고 결과를 정리한다.

5) 마감 단계

벤치마크테스트의 절차, 방법 및 비교분석 결과를 정리 검토하고 결과에 대한 고객 의견을 조사한다. 최종적으로 벤치마크테스트 과정 및 결과를 저장한다.

[표 1] 유형별 벤치마크테스트 프로세스

단 계	활 동	비교분석 벤치마크테스트 작업	순위도출 벤치마크테스트 작업
요구분석	S/W분석 및 업체 상담	<ol style="list-style-type: none"> 1. 대상 S/W 분석(프로그램, 매뉴얼, 기타 관련자료) 2. 시험 대상항목 및 벤치마크테스트 항목 리스트 도출 3. 해당 업체들과 상담(벤치마크테스트 항목 등) 4. 항목별 시험 절차 및 방법 정의 5. 투입공수 및 수수료 산정 	<ol style="list-style-type: none"> 1. 대상 S/W 분석(프로그램, 매뉴얼, 기타 관련자료) 2. 시험 대상항목 및 벤치마크테스트 항목 리스트 도출 3. 벤치마크테스트 의뢰기관과 상담(벤치마크테스트 의뢰기관이 요구하는 항목 및 일반적인 벤치마크테스트 항목 등) 4. 항목별 시험 절차 및 방법 정의 5. 투입공수 및 수수료 산정
계획	일정 및 자원 투입 방안	<ol style="list-style-type: none"> 1. 인력 및 자원 투입 계획 2. 일정 계획 3. 교육 계획 4. 전문가 그룹 구성(Optional) 	<ol style="list-style-type: none"> 1. 인력 및 자원 투입 계획 2. 일정 계획 3. 교육 계획 4. 전문가 그룹 구성
	시험 환경 구축 방안	<ol style="list-style-type: none"> 1. 시험 환경 설정 2. 시험 도구 및 S/W 선정 	<ol style="list-style-type: none"> 1. 시험 환경 설정 2. 시험 도구 및 S/W 선정
설계	벤치마크테스트 항목 도출	<ol style="list-style-type: none"> 1. 전문가 그룹 운영(Optional) 2. 벤치마크 항목 확정 3. 메트릭 작성(항목별 측정기준 및 공식 도출) 4. 해당 업체들과 합의 	<ol style="list-style-type: none"> 1. 전문가 그룹 운영 2. 벤치마크 항목 확정 3. 항목별 가중치 부여 4. 항목별 배점 부여 5. 과락 항목 선정 6. 메트릭 작성(항목별 측정기준 및 공식 도출) 7. 전문가 그룹 및 의뢰기관과 합의
	항목별 벤치마크테스트 방법 개발	<ol style="list-style-type: none"> 1. 벤치마크테스트 절차 개발 2. 벤치마크테스트 방법(시나리오, 스크립트, 테스트케이스 등) 개발 	<ol style="list-style-type: none"> 1. 벤치마크테스트 절차 개발 2. 벤치마크테스트 방법(시나리오, 스크립트, 테스트케이스 등) 개발
	시험환경 구축	<ol style="list-style-type: none"> 1. 시험 환경 구축 2. 시험 도구 및 S/W 설치 	<ol style="list-style-type: none"> 1. 시험 환경 구축 2. 시험 도구 및 S/W 설치
실행	벤치마크테스트 수행	<ol style="list-style-type: none"> 1. 벤치마크테스트 항목별 시험 진행 	<ol style="list-style-type: none"> 1. 벤치마크테스트 항목별 시험 진행
	검토	<ol style="list-style-type: none"> 1. 문제점 정리/협의 2. 보완 시험 3. 결과 정리 	<ol style="list-style-type: none"> 1. 문제점 정리/협의 2. 보완 시험 3. 결과 정리
마감	보고서 작성	벤치마크 절차, 방법 및 비교분석 결과 정리	벤치마크 절차, 방법 정리 및 순위표 산출
	검토	벤치마크 결과 검토	벤치마크 결과 검토

2.2 순위도출 형 프로세스 고려사항

[표 1]은 비교분석 형 벤치마크테스트와 순위도출 형의 벤치마크테스트의 프로세스를 비교한 도표이다. 도표에서 나타나듯이 일반적인 비교분석 형 벤치마크테스트와 순위도출 형 벤치마크테스트는 동일한 프로세스를 준수하지만 각 단계별로 약간의 차이를 갖는다.

2.2.1 요구분석 단계

비교분석 형에서 벤치마크테스트 항목 협의 주체는 벤치마크테스트에 응하는 업체, 벤치마크테스트 수행기관, 전문가집단이다. 항목을 도출하고 확정하는 단계에서 특정 제품 또는 특정 기능에 편중되지 않도록 모든 참여 업체들과 협의 조율 과정이 수반된다.

반면 순위도출 형에서의 항목 협의 주체는 벤치마크테스트 의뢰기관, 벤치마크테스트 수행기관, 전문가집단이다. 벤치마크테스트에 응하는 업체는 협의 주체 대상에 포함될 수 없다. 항목을 도출하고 확정하는 단계에서 의뢰 기관의 요구사항을 정확히 파악하고 전문가 집단 및 수행기관과의 협의 조율 과정이 수반된다.

2.2.2 설계 단계

설계단계의 최종 산출물은 벤치마크테스트 평가모델이다. 최종 확정된 평가 항목들은 요구분석 단계와 마찬가지로 각각 참여 업체들과의 합의(비교분석 형), 의뢰 기관과의 합의(순위도출 형)를 필요로 한다.

[표 1]의 순위도출 형 설계단계의 작업 영역을 살펴보면 평가모델 개발 시 부가적으로 고려해야 할 사항으로 항목별 가중치, 항목별 배점, 과락 항목에 대해 언급되어 있다. 벤치마크테스트 목적에 따라 평가모델 개발 요소가 크게 달라짐을 알 수 있으며, 각 사항의 자세한 내용에 대해서는 3장에서 다루고 있다.

2.2.3 마감 단계

벤치마크테스트 수행 종료 후 마감단계에서는 각각 벤치마크테스트 항목 별 비교 분석 결과와 순위 표(순위도출 형)를 산출하게 된다. 결과에 대

한 공표는 전자의 경우 각 참여업체들에게 공개되고 일반에 공표 여부는 또 다른 프로세스를 거쳐 공표 여부가 결정되게 된다. 하지만 후자의 경우 벤치마크테스트의 결과가 의뢰 기관에만 공개되며, 참여업체 및 일반에 공개하지 않는다.

3. 평가모델 개발 이슈

순위도출을 위한 벤치마크테스트는 비교분석을 위한 벤치마크테스트와 근본적으로 목적이 다르기 때문에 의뢰 기관의 요구가 정확히 반영되고 평가모델 개발 시 순위를 도출할 수 있는 별도의 요소가 추가되어야 한다. [표 1]의 설계단계를 살펴보면 순위도출 벤치마크테스트 작업 영역에 평가모델 개발 시 고려해야 할 요소로 항목별 가중치 부여, 항목별 배점 부여, 과락항목 설정을 나타내고 있다.

본 장에서는 상기 언급한 순위도출 형 벤치마크테스트 평가모델 개발 시 고려해야 할 요소들에 대해 정의하고 각각의 적용 방법에 대해 자세히 설명하도록 한다.

3.1 가중치 부여

가중치는 시험 대상 제품의 평가항목 중요도에 따라 차등적으로 부여해야 하며 벤치마크테스트 의뢰기관, 벤치마크테스트 수행기관, 전문위원 3자의 충분한 의견 수렴을 거친 후 결정되어야 한다.

세부항목을 기준으로 벤치마크테스트 항목의 중요도를 2단계(필수항목, 부가항목)로 구분하여 가중치를 부여하는 방법, 3단계(필수항목, 중요항목, 부가항목)로 구분하는 방법 또는 더욱 세밀하게 구분하여 가중치를 부여할 수도 있다. 하지만 통상적으로 2단계 또는 3단계 구분이 일반적인 가중치 부여 방법으로 사용되고 있다.

항목의 중요도를 2단계로 구분할 경우 평가모델이 간결하고 명확해 지는 반면 정밀도는 보다 세밀한 구분방법 보다 덜하게 된다. 반면 3단계로 구분할 경우 보다 정밀한 평가가 가능하지만 제품의 순위에 변동을 줄 만큼의 차이를 발생시키는 일은 드물다. 순위를 선정하기 위한 벤치마크

테스트에서는 제품을 도입하려는 벤치마크테스트 의뢰기관에서 중요하게 여기는 필수항목에 통상 50% 이상의 배점을 할당하기 때문에 더욱 그러하다.

[표 2]와 [표 3]은 TTA에서 기 수행한 순위 선정을 위한 벤치마크테스트 사례 중 일부 데이터이며, 가중치 단계와 항목 수를 나타내고 있다. 해당 벤치마크테스트에서는 총 40개의 세부항목에 대해 2단계 가중치를 4:1 비율로 부여하여 적용하였다.

[표 2] 2단계 가중치 및 항목 수

	기능성	사용성	효율성	이식성	합계
필수항목	16	5	3	5	29
부가항목	5	2	2	2	11
합계	21	7	5	7	40

[표 3] 3단계 가중치 및 항목 수

	기능성	사용성	효율성	이식성	합계
필수항목	13	3	3	3	22
중요항목	5	2	1	2	10
부가항목	3	2	1	2	8
합계	21	7	5	7	40

3.2 배점 기준

3.2.1. 하향식 및 상향식 배점 부여

품질특성, 분류, 평가항목, 세부항목, 시험항목, 테스트케이스의 6단계의 레벨로 구성된 평가모델을 예로 들면 품질특성별로 큰 점수를 부여한 후 하향식으로 점수를 배분하는 Top-Down 방식과, 테스트케이스마다 작은 점수를 부여하여 상향식으로 점수를 합산하는 Bottom-Up 방식이 존재한다. 언급한 두 가지 방식이 S/W 벤치마크테스트의 일반적인 배점방식이다. 상기 가중치 부여 결과를 그대로 점수화하여 다음 하위 레벨에 배점

하는 경우가 Top-Down 방식에 해당할 수 있다. 이렇게 Top-Down 방식을 적용할 경우 최하위 레벨의 테스트케이스들 간의 형평성이 결여될 수 있는 문제점이 발생한다. 이러한 문제점은 Bottom-Up 방식의 장점을 반영함으로써 어느 정도 해소할 수 있다.

Bottom-Up 방식을 적용할 경우에는 가중치 부여 시 구분했던 필수항목, 중요항목, 부가항목들 각각에 대해서는 항목 내부에서 동일한 배점을 적용해야 한다. 이렇게 함으로써 Top-Down 방식을 적용함으로써 발생하는 형평성 결여 문제점을 일부 해소할 수 있다.

3.2.2. 기능구현여부와 기능구현 정확성

순위를 선정하는 벤치마크테스트에는 기존 벤치마크테스트 보다 다양한 제품들이 벤치마크테스트에 참여한다. 이때 기능 평가 항목이 모든 제품을 수용하기는 현실적으로 불가능하다. 어떠한 항목에 대해 해당 제품의 모호한 구현은 Pass/Fail을 판단하는데 어려움이 있다. 따라서 명확한 기준을 마련하기 위해 기능성에 속한 벤치마크테스트 항목은 실제 해당 기능이 구현되어 있는지를 시험하는 “기능구현”과 구현된 기능이 의도한대로 정상 동작 하는지 여부를 시험하는 “정확성”으로 구분하여 배점한다. 이렇게 구분된 배점을 통해 모호하게 구현된 기능에 대해 부분 점수를 부여할 수 있으며, 부분 점수의 적용을 통해 벤치마크테스트를 수행하면서 발생할 수도 있는 오 측정의 역효과를 부분적으로 감소시킬 수 있다.

3.3 과락 항목 설정

구매를 위한 우선순위 도출을 목적으로 하는 벤치마크테스트 수행 시 대상 제품은 시험수행기관 또는 의뢰기관에서 선정하는 것이 아니라 신청에 의해 접수된 제품을 대상으로 벤치마크테스트를 수행하게 된다. 이러한 경우 목적에 부합하지 않는 제품들도 벤치마크테스트에 신청하는 경우도 있다.

과락 항목의 범위는 벤치마크테스트 의뢰자가 ‘반드시 구현되어야 하며 정상적인 동작을 보장해

야한'다고 여기는 항목과, 간단한 예로 "제품은 정상적으로 설치 가능해야한다." 와 같은 시험 진행을 위한 필수 기본 항목으로 구성한다.

과락 항목 채택 여부는 실질적으로 벤치마크테스트 의뢰자와 협의 하에 채택 하여야 하며, 이러한 과락 항목에 대해서는 모집공고 시 명확하게 공시하여야만 한다.

과락 항목에 대한 모집공고를 명확하게 공시하여도 해당 요구사항을 만족하지 못하는 제품들이 실제 벤치마크테스트에 응하는 사례가 다수 존재하며 현재까지 TTA에서 수행한 순위 선정을 위한 벤치마크테스트 참여 업체 중 37.5% 가량이 해당 케이스로써 실격된 전례를 보인다.

일반적으로 과락 항목은 배점을 부여하지 않으며 Pass/Fail 형태의 특수 항목으로 분류한다.

4. 결론 및 향후 연구

본 논문에서는 각 제품의 기능별 성능별 비교분석을 통해 제품의 장단점을 파악하는 것을 목적으로 하는 기존의 벤치마크테스트 외에 최근 수요가 늘고 있는 순위 도출을 목적으로 하는 벤치마크테스트에 대해 연구하고, 각 단계별로 주요 점검 요소를 도표로 정리하여 나타내었으며, 세부적으로 상이한 요소에 대해서 자세한 설명을 통해 해당 벤치마크테스트 프로세스 및 평가모델 개발과 관련한 지침을 제시하였다.

향후 보다 다양한 가중치 부여 방법 개발에 관한 연구와, 보다 공정한 배점기준에 관한 연구를 지속적으로 진행하여 평가모델에 반영하여야 하며, 이러한 연구를 통해 벤치마크테스트 목적에 부합하는 적합한 평가모델 개발이 가능할 것이라 예상된다.

[참고문헌]

- [1] Benchmark handbook, <http://www.benchmarkresources.com/handbook>, 2004
- [2] 소프트웨어 벤치마크테스트(BMT) 현황, 김재웅, 신석규, 2005년 3월, 정보처리학회지

[3] S/W 벤치마킹 테스트 및 품질 표시제 시행 방안에 관한 연구 보고서, 한국정보통신기술협회(TTA), 2002

[4] S/W분야별 벤치마킹 테스트 모델 개발 및 시험서비스에 관한 연구 보고서, 한국정보통신기술협회(TTA), 2002

[5] 벤치마킹 테스트 보고서, 한국정보통신기술협회(TTA), 2004

[6] 소프트웨어 품질 벤치마킹을 위한 평가기술에 관한 연구, 한국정보통신기술협회(TTA), 2001