

# 사용자 세션을 이용한 효과적인 웹 응용 테스트\*

이춘우<sup>o</sup> 이병정<sup>\*\*\*</sup> 김희천<sup>\*\*</sup> 우치수  
서울대학교 컴퓨터공학부  
서울시립대학교 컴퓨터과학부<sup>\*\*\*</sup>  
한국방송통신대학교 컴퓨터과학과<sup>\*\*</sup>  
{oniguni<sup>o</sup>, wuchisu}@selab.snu.ac.kr  
bjee@venus.uos.ac.kr<sup>\*\*\*</sup>  
hckim@knou.ac.kr<sup>\*\*</sup>

## Effective Web Application Testing Using User Sessions

Chunwoo Lee<sup>o</sup> ByungJeong Lee<sup>\*\*\*</sup> Heechern Kim<sup>\*\*</sup> Chisu Wu  
School of Computer Science and Engineering, Seoul National University  
School of Computer Science, University of Seoul<sup>\*\*\*</sup>  
Dept. of Computer Science, Korea National Open University<sup>\*\*</sup>

### 요 약

웹 응용은 현재 가장 빠르게 성장하고 있고 사회에 미치는 영향이 큰 분야이다. 잘 못 동작하는 웹 응용은 사회적인 큰 손실이 될 수 있다. 이에 웹 응용의 검증 방법의 중요성이 대두되고 있다. 검증 방법으로는 코드 분석과 테스트가 있다. 허나 시장에 제품이 출시되는 시간이 매우 짧고, 관련 기술의 발달이 빠른 웹 응용 개발의 특성상 코드 분석은 꽤 정확한 검증 결과를 얻을 수는 있으나 적절하다고는 할 수 없다. 따라서 테스트를 검증 방법으로 사용하게 된다. 그렇다고 해도, 테스트에 들이는 시간과 노력이 코드 분석과 비슷하다면 이 또한 적절한 검증 방법이라고 할 수 없다. 테스트의 비용을 줄여 전체적인 개발 비용을 줄이는 것이 필요하다. 테스트의 비용은 테스트 케이스를 수행하는 시간과 테스트 케이스 생성, 분석하는 시간으로 측정 가능하다. 이 논문에서는 이 시간들을 줄이기 위해 전체 웹 응용의 경로 중 사용 빈도가 높은 경로를 테스트 케이스로 하여 테스트 케이스 실행 시간을 줄이고, 각 페이지에 맞는 입력 값으로 빈도가 높은 사용자 세션 정보를 이용하여 테스트 케이스의 생성을 자동으로 할 수 있게 하여 테스트가 직접 테스트 케이스를 생성하는 것보다 시간을 줄였다.

### 1. 서 론

웹 응용(Web application)은 현재 가장 빠르게 성장하고 있는 분야 중 하나이다. 이제 웹 응용은 상품을 판매하고 배송하는 사업이나, 엔터테인먼트, 정보의 공유나 전달 등 다양한 분야에서 널리 사용되고 있다. 이렇게 웹 응용이 사회에서 차지하는 비중이 큰 경우, 웹 응용의 오작동은 사회적으로 큰 영향을 끼칠 수 있다. 그러므로 웹 응용이 올바른 동작을 하도록 보장하는 것은 중요하다.

검증 방법으로는 코드의 정적, 동적 분석과 테스트가 있다. 그 중 테스트는 코드의 분석보다 빠른 기간 내에 비교적 정확하게 웹 응용의 오류를 찾는다. 웹 응용의 검증 위해, 많은 웹 응용 검증기술과 도구가 개발되었다. 그러나 대부분의 도구들이 웹 응용의 기능적 요구 사항 검증보다는 링크의 정확성, 로드 테스트(load testing) 등의 부분에 주로 초점이 맞춰져 있다. 몇몇 기능적 요

구 사항에 초점이 맞춰져 있는 도구들도 있으나 테스트 케이스 생성 시 테스트의 입력 값을 사용하는 캡처-리플레이(Capture-replay)방법을 지원한다.

테스트의 비용은 크게 두 가지 방식으로 측정 가능하다. 하나는 각 테스트 케이스를 수행하는데 드는 시간이고 나머지 하나는 테스트가 테스트 케이스를 만들고 그 테스트케이스를 분석하는 데 드는 시간이다[1]. 이 두 측면에서 살펴보면 기존의 웹 응용의 테스트 방법-테스터가 직접 테스트 케이스를 만든다-는 비용이 많이 드는 방식이다. 웹 응용은 점점 더 복잡해지고 이에 따라 테스트 케이스의 개수도 많아진다. 많은 테스트 케이스를 수행하는 데에는 많은 시간이 들고, 테스트가 이를 직접 만들고 분석하는 데에도 시간이 많이 든다.

웹 응용은 다른 응용 프로그램과 달리 시장에 제품이 출시되기 까지 개발 기간이 매우 짧다. 짧은 개발 기간은 정확한 웹 응용을 개발해야 하는 것과 함께 웹 응용 개발자들에게 부담을 준다. 이 때 테스트에 많은 시간을 소비한다면 개발에 드는 전체적인 비용이 올라간다. 앞에서 살펴본 두 가지 측면에서 테스트의 비용을 줄인다면 전체적인 개발비용을 줄일 수 있다. 이 논문에서는 효율적인 테스트 케이스 생성 방법을 제안하여 테스트의

\* 본 연구는 한국과학재단 목적기초연구(R01-2002-000-00135-0)지원으로 수행되었음.

비용을 줄이고자 한다. 테스터가 직접 페이지의 순서를 정하고 그에 맞는 이름-값 쌍을 정하는 대신 시제품에서 얻을 수 있는 로그를 분석하여 그 결과를 참고로 테스트 케이스를 만든다면 필수적인 테스트 케이스를 자동적으로 만들 수 있다. 이 때, 테스트 케이스의 수를 줄일 수 있다면 비용을 결정하는 방법 중 첫 번째 요소에 관해 줄일 수 있고 자동으로 생성한다면 두 번째 측면에서 비용을 줄일 수 있다.

2장에서는 관련 연구를 살펴보고 웹 응용의 특징과 웹 테스트 케이스의 구성에 대한 정들은 3장에서 알아본다. 그리고 다음 장에서 테스트 케이스의 효율적인 생성 방식에 대해 기술한다. 마지막 5장 6장에서는 이에 대한 연구들을 해보고 결론을 맺도록 하겠다.

2. 관련 연구

웹 응용의 테스트 케이스는 일정한 경로(path)를 통해 방문할 페이지의 순서와 각 페이지에 관련된 이름-값(name-value) 쌍, 즉 테스트 입력 값으로 구성된다. 웹 응용의 테스트 케이스 생성도 이 두 부분에 초점을 맞춰서 이루어진다. 이 장에서는 관련된 연구 몇 가지를 간략하게 살펴보고 다음 장에 각각의 장단점에 대해 살펴본다.

웹 응용의 테스트 기술에 관련된 연구는 몇 가지가 있으나 최근에 발표된 방식으로는 Ricca의 방식이 있다 [2]. [2]에서 Ricca는 UML을 이용하여 웹 응용 모델을 만들어 웹 응용의 분석과 테스트 케이스 생성을 한다. 경로 표현(path expression)을 통해 가능한 경로를 표현하고 이를 이용해서 임시 테스트 케이스(quasi test case)를 생성한다. 이 방식은 UML을 사용하여 표현하여 표준적인 테스트 케이스를 생성한다는 장점을 가지고 있으나, 경로 표현에서 가능한 모든 경로를 다 사용한다는 점에서 크기가 큰 웹 응용의 경우 그 경로가 너무 많아져 전체적인 테스트 케이스의 수가 많아진다. 그리고 이름-값 쌍을 테스터가 직접 만들어야 하기 때문에 테스트 케이스 생성 시간이 많이 걸릴 수 있고, 그 입력 값의 정확도를 예상하기 힘들다.

Ricca의 문제점을 경로의 측면에서 보완하기 위해 Lee가 웹 응용의 항해 구조를 이용하여 테스트 케이스를 줄이는 방법을 제안하였다 [3]. 이 방법은 웹 응용에서 방문 횟수가 많은 링크로만 이루어진 경로로 경로 트리(path tree)를 만들고 이를 이용하여 테스트 케이스를 생성하는 방식이다. 이 방법은 빈도가 높은 경로를 이용하고 자동으로 생성이 가능하다는 점에서 Ricca의 방법보다는 테스트 케이스 생성 시간 및 테스트 수행 시간 단축의 효과가 기대되는 방식이었으나 테스트 입력 값에 대한 고려가 없기 때문에 보다 효율적인 테스트 케이스 생성 방법의 여지가 있다.

Ricca의 방법과 달리 테스트 케이스를 자동으로 만들려는 시도가 Eibaum에 의해 시도되었다 [4]. Eibaum은 테스트 케이스 생성 시간에 초점을 맞추었다. 테스터가 경로를 정하고 이름-값 쌍을 직접 정하는 대신 웹 응용 사용자가 웹 응용을 사용하면서 남긴 로그 정보를 이용하여 테스트 케이스를 생성했다. 이 방법은 테스트 케이

스 생성 시간이라는 측면에서는 좋은 결과를 보였지만 사용자의 세션을 그대로 테스트 케이스로 사용하는 방법으로는 세션의 선택 방법이 부채하고 얼마나 많은 양의 세션을 사용할 것인지 그리고 사용자 세션에 나타나지 않는 입력 값에 대한 대처가 없었다.

3. 통계적 데이터를 사용한 웹 응용 테스트 기술

2장에서 살펴본 바와 같이 아직 웹 응용의 테스트 케이스 생성 방법에는 개선의 여지가 남아있다. 웹 응용 테스트의 비용을 줄이기 위해 첫째, 테스트 케이스가 수행되는 시간을 단축하면 되고 둘째, 테스트 케이스 생성 및 분석 시간을 줄이면 된다. 이를 이루는 방법은 테스트 케이스를 자동으로 생성하되 그 케이스들이 테스터가 직접 만든 것보다 테스트 효과가 떨어지서는 안 된다. 이를 위해 테스트 케이스의 두 구성 요소인 경로(즉, 페이지의 순서)와 입력 값(이름-값 쌍)의 개수에서 살펴본다.

3.1 임시 테스트 케이스로부터 링크 제거

임시 테스트 케이스는 웹 응용에서 가능한 모든 항해 경로이다. 웹 응용이 크고 복잡해짐에 따라 이 가능한 항해 경로 또한 많아진다. 웹 응용의 특성 상 한정된 개발 시간 또는 짧은 보수 시간 이내에 능한 모든 경로를 다 테스트하는 것은 어려운 일이다. [3]에서 제안된 방법은 방문 횟수가 많은 링크를 중심으로 신장 트리(spanning tree)를 생성하는 것으로써 사용 빈도가 높은 테스트 케이스를 생성하여 전체적인 테스트 경로의 수를 줄이는 방법이다. 링크의 방문 횟수를 알 수 있는 도구는 개발되어 있다 [5]. 페이지와 링크를 그래프로 모델링할 수 있기 때문에 방문 횟수를 가중치로 갖는 여러 개의 신장 트리를 생성할 수 있고, 생성된 신장 트리를 가중치를 기준으로 순위를 매길 수도 있다. 생성된 신장 트리 중 몇 가지를 선택하면 임시 테스트 케이스보다 적은 수의 경로가 생성되지만 높은 빈도의 테스트 케이스를 생성하는 것이므로 테스트의 정확도 크게 영향을 주지 않을 수 있다.

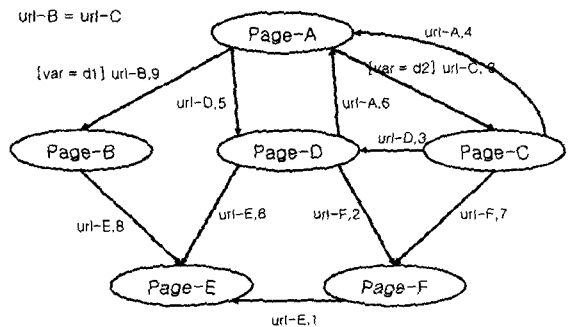


그림 1 페이지 항해 다이어그램의 예.

그림 1은 페이지 항해 다이어그램의 한 예이다. 앞에서 설명한 신장 트리를 만들면 그림 2가 된다. 그림 2의 신장 트리의 크기는 32가 된다. 여러 개의 신장 트리를 만들면 다양한 크기의 신장 트리가 생기게 된다. 신장 트리의 크기가 큰 것이 빈도가 높은 경로로 이루어진 것이다.

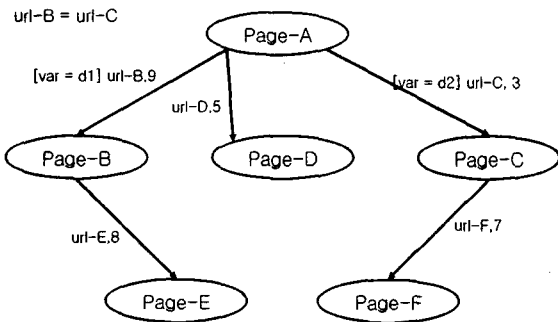


그림 2 페이지 간의 항해 테스트 케이스 추출을 위한 신장 트리의 예. (가중치 합 : 32)

### 3.2 사용되지 않는 입력 값 제거

자주 사용되지 않는 경로를 줄이는 방법으로 테스트 케이스의 개수를 줄일 수 있지만 입력 값의 개수가 많고 이를 테스터가 직접 입력한다면 테스트의 비용이 증가한다. 테스트 입력 값의 크기를 최적화 하는 것은 NP-complete이다[5]. 따라서 완벽하게 최적화할 수는 없지만 어느 정도는 줄일 수 있다.

3.1에서 테스트할 페이지들의 순서를 정했다. 이제 각 페이지에 맞는 이름-값 쌍을 정하기 위해 로그의 세션들 중에서 테스트 케이스의 페이지 순서와 같은 순서의 세션을 찾는다. 3.1에서 정한 페이지의 순서는 가장 빈도가 높은 링크로 되어 있기 때문에 로그에서도 여러 세션이 있을 것이다. 하지만 그 세션들의 이름-값 쌍은 다를 수도 있고, 같을 수도 있다. 이 세션들 중에서 선택하는 방법에는 몇 가지 방법이 있다. 하나는 중복되는 것들을 제외한 모든 세션을 사용하는 것이다. 두 번째로 여러 세션 중에서 빈도가 높은 링크들을 고른 후 이들을 섞어서 새로운 입력 값을 만드는 것이다. [4]에 따르면 세션만을 이용한 방법에는 시험 사용자들의 행동이 다양하지 않다는 한계가 나온다. 하지만 빈도가 높은 세션을 사용하여 입력 값을 만든다면 공통적으로 발생하는 상황에 대해서 매우 빠른 시간 안에 테스트 케이스를 생성할 수 있다.

### 4. 결론

빠르게 발전하는 웹 관련 기술로 인해 현실적으로 웹 응용을 개발하는데 테스트는 코드 분석보다 좋은 검증 방법이다. 하지만 웹 응용의 크기가 커지고 복잡해 짐에 따라 테스트 케이스도 커지게 되었고 이제 테스터가 직접 테스트 케이스를 만드는 것은 비효율적이 되었다. 이

에 우리는 테스트 케이스를 자동으로 생성하는 방법을 생각하게 되었다. 테스트 케이스를 자동으로 생성하기 위해서 시험 사용자들이 남긴 로그를 분석하였고 이를 통해 사용 빈도가 높은 링크로 이루어진 페이지들의 순서를 찾아내고 각 페이지에 관련된 이름-값 쌍을 정하기 위해서 빈도가 높은 세션과 빈도가 낮은 세션을 사용하였다.

이 논문에서 기술한 방법을 통해 테스트 케이스를 자동으로 생성할 수 있다고 생각한다. 그리고 세션을 테스트 케이스로 직접 사용하는 방법이나 가능한 모든 경로를 다 테스트 하는 방법보다 더 빈도가 높은 케이스가 생성되므로 더 효율적인 테스트를 할 수 있다.

테스팅을 할 때 짧은 시간으로 비용을 줄이는 것과 정확한 테스트는 서로 다른 방향을 향해가는 두 마리의 토끼이다. 여기서 제한한 테스트 방법은 그 중 비용이라는 측면에 초점을 맞춘 것이다. 반대 측면에서는 빈도가 높은 경로와 세션의 선택은 결국 가능한 모든 경우 중에서 몇 가지를 선택한 것이므로 테스트의 정확성에 문제가 생길 수도 있다. 만에 하나 시험 사용자가 행하지 않았던 입력이 테스트에 필수적인 입력인 경우가 생길 수도 있기 때문이다. 이런 경우 지금으로서는 달리 방법이 없다. 이런 문제는 보다 많은 수의 세션 정보를 얻는 것으로 어느 정도 해결 할 수 있을 것이다.

새로운 웹 응용을 만들 때 어느 정도는 기존의 웹 응용과 많은 부분에서 비슷한 점을 가지고 있다. 기존에 사용하던 웹 응용의 경우 그 사용 기간에 따라 많은 양의 사용자 세션 정보가 존재 한다. 이 세션 정보를 사용하면 보다 풍부한 정보를 가지고 테스트 케이스를 생성할 수 있다.

또 한 가지 정확성과 관련된 요인으로는 테스트 스위트의 크기를 들 수 있다. 빈도가 높은 경로들 중 몇 가지를 선택하느냐와 세션을 선택할 때 빈도가 높은 것의 수와 빈도가 낮은 것의 수가 정확도를 높이는 데 영향을 미칠 것이다. 앞으로 이 방법을 적용한 테스트 도구들 만들어서 이 요인들에 대해 실험을 하여 적절한 경로의 종류, 세션의 수를 정하는 것이 필요하다.

### 5. 참고 문헌

- [1]Y. A. Zuev, "A set-covering problem: The combinatorial-local approach and the branch and bound method," U.S.S.R Computational Mathematics and Mathematical Physics, 19,(6), 1979
- [2]F. Ricca and P. Tonella, "Analysis and Testing of Web Applications," Proc. of the 23rd International Conference on Software Engineering, pp. 25-34, 2001.
- [3]Lee Chunwoo, "가중치에 기반한 웹 응용 항해 테스트 모델," 정보과학회, 2003
- [4]S. Elbaum, "Improving Web Application Testing with user session data," ICSE, 2003
- [5]B. Berrendt and M. Spiliopoulou, "Analysis of navigation behavior in web sites integrating multiple information system," VLDB, 2000.
- [6]M. R. Gary and D. S. Johnson, "Computers and Intractability," Freeman, New York, 1979.