시한폭탄 기반의 악성 웹 콘텐츠들의 효율적인 탐지 방법*

김동진○ 조성제† 김홍근††
단국대학교 컴퓨터학과○†, 한국인터넷진흥원††
kdjorang@dankook.ac.kr○, sjcho@dankook.ac.kr†, hgkim@kisa.or.kr††

An Efficient Method to Detect Malicious Web Contents Based on Time-Bomb

Dongjin Kim○ Seongje Cho† Hongkun Kim††
Department of Computer, Dankook University○†, Korea Internet & Security Agency††

1. 연구의 필요성

방화벽, IDS/IPS 및 안티바이러스 도구의 도입, 서버 보안의 강화 등으로 서버나 웹서버 측 공격이 어려워지자, 상대적으로 보안이 약한 클라이언트 (PC) 측 공격이 급증하고 있다[1,2]. 대표적으로 웹 브라우저의 취약점을 악용하여, 사용자가 인지하지 못한 상태에서 악성 웹 콘텐츠를 다운로드하는 "Drive-by-Download" 공격 등을 이용한 개인 정보 탈취, 웜 및 트로이목마의 전파 등의 악성 행위들이 증가하고 있다. 2007년과 비교하여 2010년 발표된 OWASP Top 10에 의하면, 사용자를 속이기 위해 정상적인 웹사이트에서 사용자 모르게 악성 콘텐츠 유포 사이트로 유도하는 "Unvalidated Redirects and Forwards" 취약점이 새롭게 추가되는 등, 새로운 악성 웹 콘텐츠에 대한 방어가 중요해졌다. 하지만 이를 방화벽 등의 전통적인 방법으로 방어하는 것은 불가능하기 때문에 직접 웹브라우저로 웹 서버를 방문하고, 시스템에 허가되지 않은 상태 변화를 분석하여, 악성행위를 탐지하는 고 상호작용 클라이언트 허니팟(high-interaction client honeypot)에 대한 연구가 활발히 이루어지고 있다[1, 3].

최근에는 고 상호작용 클라이언트 허니팟을 우회하기 위해 허니팟의 웹 서버(페이지) 방문시간 보다 더 긴 시간이 지난 후 공격이 수행되도록 하는 시한폭탄(time-bomb)공격 등의 공격방법이 등장하였다. 동적분석인 고 상호작용 클라이언트 허니팟으로는 시한폭탄 기반의 악성 페이지를 탐지하는 것이 어렵고, 탐지하더라도 탐지 시간 증가 및 False-Negative 발생 등의 오버헤드가 발생한다[4, 5].

본 논문에서는 시한폭탄 기반의 공격을 탐지하는 오버헤드를 최소화하고, 탐지 정확도를 높이기 위해 기존의 동적분석에 정적분석을 결합한 모델을 제안하고, 실제 실험을 통해 평가하고 검증하였다.

2. 시한폭탄 기반 악성 웹 콘텐츠 탐지 시스템 구축 및 실험

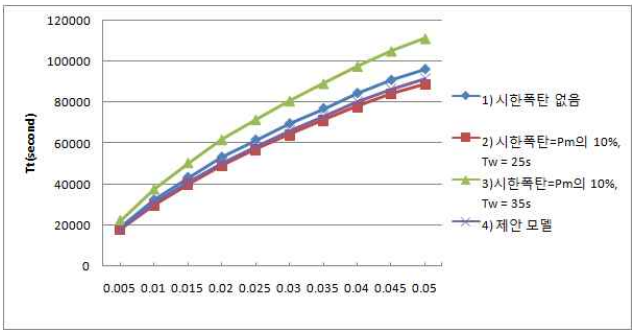
효율적인 시한폭탄 기반 악성 웹 콘텐츠 탐지를 위해 동적분석 이전에 정적분석을 통해 시한폭탄 공격 패턴을 가진 시한폭탄 공격 의심 웹 페이지들을 분류하여, 시한폭탄 공격으로 분류된(의심되는) 페이지들에 대해서는 방문대기시간을 늘리고(35초), 악성 웹페이지 비율이 높을 경우에는 효율적인 순차 방문 알고리즘을 적용하였다. 시한폭탄 공격으로 분류되지 않은 웹 페이지들은 기존 연구를 통해 확인된 일반적인 공격 대기시간, 즉 일반적인 고 상호작용 클라이언트 허니팟의 방문 대기시간(25초)을 적용하고, 악성 웹 페이지 비율이 낮은 경우 효율적인 분할정복 방문 알고리즘을 적용하는 탐지 모델을 제안한다[6]. 제안 모델의 성능을 분석하기 위해 "자원비용(전력)", "속도(탐지시간)", "탐지 정확도(True Positive, TP)", "악성 웹서버 비율(Pm)" 등을 이용하여, 고 상호작용 클라이언트 허니팟 시스템을 평가하는 방법인 비용 기반 평가방법(True Positive Cost Curve, TPCC)을 통해 검증하였다[4].

실험을 위해 기존 연구 [4]를 따라 시간 당 소모 전력은 0.125 US dollars로 고정하고, 탐지 정확도는 측정할 수 없으므로 악성웹페이지를 정상 페이지로 잘못 판단하는 False-Negative(FN)를 1에서 뺀 값(1-FN)을 사용하며, Pm은 0.005~0.05, 분할정복 방문 알고리즘의 웹서버 그룹의 크기는 기존 연구를 따라 Pm당 각 80~8개(그룹 크기)를 사용하였다[1]. 전체 탐지시간(Tt)을 결정짓는 요소 중 시한폭

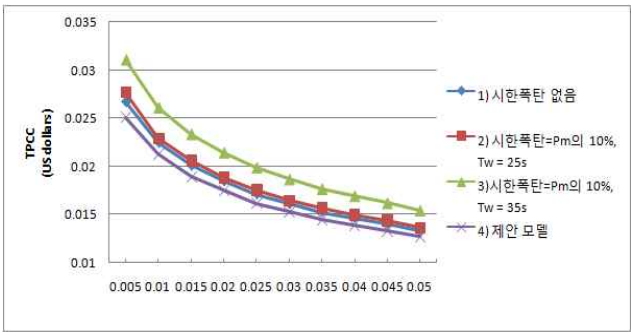
* 본 연구는 단국대학교 대학연구비의 지원, 2010년 한국인터넷진흥원 위탁과제 "홈페이지 은닉형 악성코드 유포 패턴 분석방법 연구"의 지원을 받아 수행되었음.

탄 공격 탐지 성능에 영향을 주지 않는 요소인 “시스템 초기화 시간(Tq)”, “가상머신 리버팅 시간(Tr)”, “웹 서버 응답 시간(Ti)”, “테스트 대상 웹 서버 개수(n)”는 Tq=500s, Tr=60s, Ti=1.2s, n=5000으로 고정하였으며, 시한폭탄 시간은 35초로 하였다. 탐지 성능에 영향을 주는 요소인 “방문대기시간(Tw)”과 “Pm에 대한 시한폭탄 공격비율”에 따라 4개의 테스트 모델을 시뮬레이션하여, (그림 1)과 같이 Tt를 기준으로 성능을 비교하고, (그림 2)와 같이 TPCC를 기준으로 성능을 비교하여, 제안모델에 따른 고 상호작용 클라이언트 허니팟 시스템의 성능 및 탐지율의 향상을 검증한다.

Tt 기준으로 모델별 성능을 평가 및 비교한 (그림 1)을 보면 제안모델인 4)보다 시한폭탄 웹 페이지를 포함하였지만, 이를 탐지하지 못하는 2)번 모델이 더 좋은 성능인 것으로 보인다. 하지만 (그림 1)은 Tt를 기준으로 비교한 것이기 때문에 시한폭탄 웹 페이지에 대한 탐지 정확도를 고려한 것이 아니므로 정확하지 못한 성능평가이다. 전체 탐지시간(Tt) 이외에 탐지 정확도(TP) 등을 고려한 (그림 2)의 TPCC를 이용한 성능 평가 비교를 보면 2)번 모델에 비해 탐지 정확도가 높고, 3)번 모델에 비해 Tt가 더 빠른, 본 논문의 제안모델 4)가 모든 Pm에서 약 10% 더 성능이 좋은 것을 알 수 있고, 이를 통해 본 논문에서 제안한 시한폭탄 기반 악성 웹 콘텐츠 탐지 모델의 효율성과 성능을 검증하였다.



(그림 1) Tt 기준의 모델별 성능 평가(X축:Pm)



(그림 2) TPCC 기준의 모델별 성능 평가(X축:Pm)

3. 결론 및 향후 방향

웹을 통해 일반 사용자들을 공격 대상으로 하는 클라이언트 측 공격이 크게 늘어나면서 새로운 공격 패턴을 찾아내고, 이를 효과적으로 예방 및 대응할 수 있는 고 상호작용 클라이언트 허니팟 시스템이 주목 받고 있다. 하지만, 고 상호작용 클라이언트 허니팟의 경우 실행 기반의 동적 테스팅으로 인해 전체 테스트 시간이 오래 걸리고, 시한폭탄 기반의 공격을 탐지할 수 없다는 단점이 있다.

이를 개선하기 위해, 본 논문에서는 동적분석 이전에 패턴 매칭 기반으로 시한폭탄 공격이라고 의심되는 웹 서버 페이지들을 정적분석하여 분류하는 기법을 제안하였다. 즉, 의심스러운 웹 페이지들을 동적분석하기 전에 시한폭탄 공격이 아닌 웹 페이지들과 시한폭한 공격의 웹 페이지들로 분류하였다. 또한, 서로 분류된 웹 페이지들에 서로 다른 동적분석 기법을 적용하여 악성 웹 페이지 탐지율을 높이고 탐지 속도를 높여서 전체적으로 약 10%의 시스템 성능 향상을 보였고, 결론적으로 고 상호작용 클라이언트 허니팟의 성능을 개선하였다.

실제 정적분석은 대부분의 동적분석의 단점 및 한계를 개선시키는데 매우 효과적이다. 따라서 본 연구진은 앞으로, 분석 대상 웹 서버를 줄이기 위해 악성이라고 의심되는 웹 서버들만을 분류하기 위해 정적분석을 도입하여 연구하려고 한다.

참고문헌

[1] Christian Seifert, Ian Welch, Peter Komisarczuk "Application of divide-and-conquer algorithm paradigm to improve the detection speed of high interaction client honeypots", SAC, March, 2008.
[2] Marco Cova, Christopher Kruegel, Giovanni Vigna, “Detection and Analysis of Drive-by-Download Attacks and Malicious JavaScript Code”, the International World Wide Web Conference Committee, April 2010.
[3] The Open Web Application Security Project, “OWASP Top 10 2010”, 2010.
[4] Christian Seifert, Peter Komisarczuk, Ian Welch, "True Positive Cost Curve: A Cost-Based Evaluation Method for High-Interaction Client Honeypots", "2009 Third International Conference on Emerging Security Information , Systems and Technologies", 2009.
[5] Christian Seifert. Personal Communication, 2006.
[6] Y.-M. Wang. Personal Communication, 2006.