웹 사이트 취약점 분석: A 논문 투고 사이트를 기반으로

서영민[1], 이재혁[2], 이경률[1]*
[1]대구가톨릭대학교 컴퓨터소프트웨어학부
[2]대구가톨릭대학교 컴퓨터소프트웨어학과
e-mail : dudals1003@cu.ac.kr, gurtmggg@cu.ac.kr, carpedm@cu.ac.kr

Vulnerability Analysis of Web Sites: Based on Manuscript Submission Site A

Yeongmin Seo[1], Jaehyuk Lee[2], Kyungroul Lee[1*]
[1]School of Computer Software, Daegu Catholic University
[2]Department of Computer Software, Daegu Catholic University

요 약

지식의 진보 및 연구분야의 발전을 위하여, 많은 연구자들이 자신의 연구를 활발하게 진행하고 있으며, 의미있는 연구 결과를 논문으로 작성하여 관련 분야의 학회에서 제공하는 논문 투고 사이트에 제출한다. 투고된 논문의 공정한 심사를 위하여, 심사위원에게는 저자정보를 공개하지 않는 것이 일반적이며, 저자와 관련된 정보가 공개되는 경우에는 공정하지 못한 심사나 부정청탁과 같은 불법적이거나 비윤리적인 행위가 발생할 수 있다. 하지만 대부분의 학회에서는 논문 투고를 위하여 웹 사이트를 활용하며, 웹 사이트가 가지는 취약점으로 인하여 저자정보가 노출되는 등의 불법접근이 가능할 것으로 판단된다. 이에 본 논문에서는 A 학회의 논문 투고 사이트를 대상으로, 웹 사이트에서 발생 가능한 취약점을 분석하였으며, 분석 결과를 기반으로 심사위원에게 공개되지 않는 저자정보 및 소속을 확인하는 웹 페이지의 불법접근 취약점을 실증하였다. 실증한 취약점을 기반으로, 불법적인 접근이 불가능하도록 대응방안을 적용한다면, 더욱 공정하게 심사할 수 있는 논문 투고 사이트를 제공할 것으로 사료된다.

1. 서론

연구자들은 과거부터 현재까지 다양한 분야에서 자신의 연구를 활발하게 진행하고 있으며, 의미있는 연구의 결과물들은 지식의 진보 및 연구분야의 발전에 기여하기 위하여 논문으로 작성한다 [1, 2].

논문 제출과 관련하여, 일반적으로 연구자들은 작성된 논문들을 다양한 학회에서 제공하는 논문 투고를 위한 전용의 웹 사이트를 통하여 제출하며, 제출된 논문을 심사하기 위하여 편집위원이 임의의 심사위원을 선정하고, 선정된 심사위원이 제출된 논문을 심사하여 최종적으로 심사결과를 통보한다. 이 과정에서 선정된 심사위원이 심사를 진행하는 과정 중, 논문 투고자와의 사적인 관계로 인하여 공정하지 못한 심사를 하거나 논문이 게재되도록 부정청탁을 하는 등의 불법적이거나 비윤리적인 행위들을 방지하기 위하여, 심사위원이 투고한 논문의 저자명 및 소속과 같은 저자정보를 확인하거나 유추할 수 없도록 논문 투고 사이트에서 관련된 정보를 비공개하고 해당 정보를 삭제한 논문 파일을 심사위원에게 제공한다.

현재 대부분 학회에서는 웹 사이트를 기반으로 논문을 투고하도록 제공하고 있으며, 만약 어떠한 이유로든 웹 사이트가 취약점을 내포한다면, 심사위원이 저자정보나 심사 중인 논문을 확인할 가능성이 존재한다. 이는 결국 연구 윤리적인 측면에서의 문제가 발생할 수 있을 뿐만 아니라, 연구자의 연구 의욕을 감소시킬 수 있으며, 향후 연구 발전을 늦추는 계기가 될 것이다. 따라서 웹 사이트에서 내포한 취약점을 비윤리적인 악용을 방지하기 위하여, 논문 투고 사이트의 취약점을 분석할 필요성이 있다[3].

이러한 취약점이 발생하는 원인으로는 대부분의 웹 사이트에서 제공하는 웹 페이지의 데이터 전송 방식으로 인하여 발생할 가능성이 높으며, 본 논문에서는 A 학회의 논문 투고 사이트를 대상으로, POST[4], GET[5]과 같은 웹 페이지의 데이터 전송 방식의 취약점을 분석하고 그 결과를 도출한다.

본 논문의 구성은 다음과 같다. 2장에서는 A 학회의 논문 투고 사이트의 취약점을 분석하기 위하여 HTTP 프로토콜의 데이터 전송 방식인 GET 방식과 POST 방식의 특징, 그리고 각 방식에서 발생할 수 있는 문제점을 서술한다. 3장에서는 분석한 문제점을 기반으로 실제 A 학회에서 제공하는 웹 사이트에서 발생 가능한 취약점을 실증하며, 결론 및 향후 계획을 4장에 서술한다.

2. 웹 사이트의 데이터 전송방식

대부분의 웹 사이트는 웹 브라우저와 웹 서버와의 데이터를 전달하기 위한 통신 규약인 HTTP(Hyper Text Transfer Protocol) 프로토콜을 사용한다[5]. HTTP에서는

데이터를 전송하기 위하여 POST 방식[3], GET 방식[4], PUT 방식[6], DELETE 방식[7]을 제공한다.

일반적으로는 데이터 조회 및 추가를 위하여, POST 방식과 GET 방식을 사용하며, GET 방식은 주로 데이터를 읽거나 검색할 때 사용하는 방식으로, URL(Uniform Resource Locator)에 쿼리 스트링을 첨부하여 서버에 전송하는 특징을 가진다. 반면, POST 방식은 새로운 리소스를 생성할 때 사용되는 방식으로, 데이터들을 URL 뒤에 첨부하여 서버로 전달하는 것이 아닌, Body에 담아서 전달하는 특징이 있다. GET 방식은 서버에 요청하는 데이터들이 URL에 첨부되기 때문에, 해당 정보가 노출되는 취약점이 존재하며, A 학회에서 제공하는 논문 투고 사이트는 GET 방식을 사용한다. 따라서 본 논문에서는 A 학회에서 제공하는 논문 투고 사이트의 취약점을 분석하기 위하여, GET 방식으로 전달되는 데이터를 분석하며, 분석 결과를 기반으로, 불법접근을 통한 저자정보 확인과 같은 악용 가능성을 검증한다.

3. A 논문 투고 사이트 취약점 분석

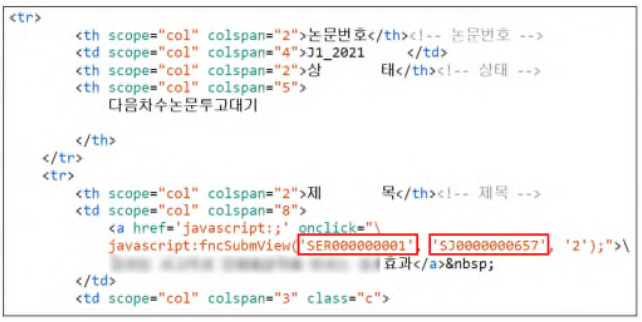
A 학회의 논문 투고 사이트는 논문 제출 및 심사 등과 같은 다양한 웹 페이지들로 구성되어 있으며, 이러한 웹 페이지들 중 논문의 저자정보를 확인할 수 있는 페이지는 GET 방식을 사용한다. 이러한 이유로, 해당 웹 페이지의 URL을 확인하고 전달되는 데이터를 분석한다면, 접근이 허용되지 않은 페이지에 불법적으로 접근이 가능한 취약점이 발생한다. 이러한 취약점으로 인하여, 공정하게 심사하여야 하는 논문의 저자정보가 외부로 노출되는 문제점이 발생할 것으로 판단된다.

상기 취약점을 확인하는 과정은 크게 3단계로 나누어지며, 첫 번째 단계는 심사하는 논문의 정보를 획득하는 단계, 두 번째 단계는 페이지 소스코드를 통한 논문 식별 번호 확인 단계, 마지막 단계는 수집된 정보를 토대로 논문의 저자정보를 확인하는 취약점 검증 단계이다.

첫 번째 단계는, 심사하는 논문의 정보를 획득하는 단계로, 획득 과정을 그림 1에 나타내었다. 그림을 살펴보면, 심사 현황 조회 페이지를 통하여, 심사할 논문의 제목, 논문번호, 분야, 투고일과 같은 정보를 확인할 수 있으며, 저자정보는 "저자: SECRET (SECRET: SECRET: 투고자)" 와 같이 심사위원이 확인할 수 없도록 구성된다. 즉, 일반적인 방법으로는 심사위원이 논문을 투고한 저자정보를 확인할 수 없다.

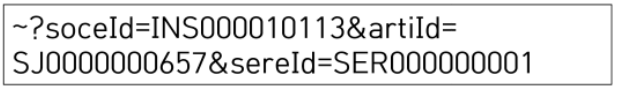
두 번째 단계는 심사 현황 조회 페이지의 소스코드를 확인하여, 심사 대상 논문의 식별 번호를 확인하는 단계이며, 여기에서 획득한 논문의 식별 번호는 세 번째 단계인 취약점 검증 단계에서 활용된다. 임의의 논문에 대한 식별 번호를 확인하는 소스코드를 그림 2에 나타내었다.

(그림 1) 심사 대상 논문의 정보를 출력하는 웹 페이지



(그림 2) 페이지 소스코드로부터 획득한 논문 식별 번호

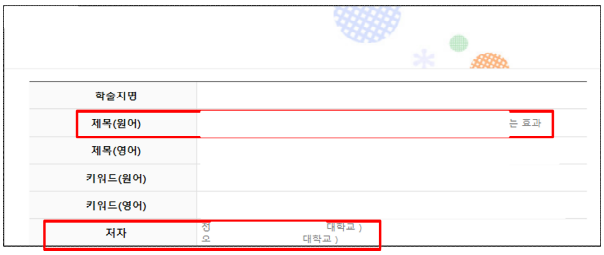
그림에서의 "SER00000007"과 "SJ0000000657"은 논문을 식별하기 위한 정보인 논문 식별 번호로 판단되며, 해당 번호가 가지는 의미를 파악하기 위하여, 식별 번호를 포함하는 URL의 일부분을 다음 그림에 나타내었다.



(그림 3) 논문 식별 번호를 포함하는 URL 일부

취약점 분석을 위하여 상기 그림의 URL을 토대로, soceId는 학회 식별 번호, artiID는 논문의 식별 번호인 것으로 가정하였다. 이 정보를 기반으로, GET 방식을 사용하여 논문 저자를 확인하도록 제공하는 웹 페이지의 URL에서 논문의 식별 번호를 심사하는 논문의 식별 번호로 조작한다면, 논문 저자를 확인할 수 있을 것으로

판단하였으며, 그 결과를 그림 4에 나타내었다.



(그림 4) 심사 대상 논문의 저자이름 및 소속 노출 결과

그림 4를 살펴보면, 심사하는 논문과 동일한 제목을 가지는 논문의 저자이름과 소속을 확인할 수 있다. 그림 1과 같이, 정상적인 접근을 통해서는 심사위원이 저자정보를 확인할 수 없도록 SECRET으로 출력하였지만, 본 논문에서 분석한 웹 사이트 취약점을 활용한다면, 저자이름과 소속을 확인할 수 있으며, 이를 통하여 심사위원의 불공정한 심사와 같은 비윤리적인 행위가 가능할 것으로 판단된다.

4. 결론

본 논문에서는 HTTP 데이터 전송 방식 중, 상대적으로 보안에 취약한 GET 방식을 사용하는 A 학회의 논문 투고 사이트를 대상으로, 제공되는 웹 페이지의 소스코드에서 노출된 논문 식별 번호를 활용하여 심사 중인 논문의 저자정보 및 소속을 확인하는 웹 페이지 불법접근 취약점을 분석하고 실증하였다. A 학회의 논문 투고 사이트에서 사용하는 방식은 A 학회뿐만 아니라, 다른 학회에서도 동일한 방식으로 사용하는 것을 확인하였으며, 이는 동일한 방식을 사용하는 학회의 논문 투고 사이트 모두 본 논문에서 검증한 취약점을 활용하여 불법접근이 가능할 것으로 판단된다.

A 학회에서 사용하는 논문 투고 시스템은 2021년 9월 27일 기준, A 학회를 포함한 1,345개의 학술단체에서 동일하게 적용되어 있으며, A 학회를 제외한 다른 학회 논문 투고 사이트에서의 취약점을 향후 분석할 예정이다.

Acknowledgements

이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. NRF-2021R1F1A1050542).

참 고 문 헌

[1] 이동국. “텍스트 마이닝을 통한 교사의 테크놀로지 통합에 관한 연구 동향 분석”, 『교육정보미디어연구』 제26권 제1호, 2020. pp.261-282.

[2] 남은숙, 김이활, 김주현, “간호직무열의에 대한 개념분석(간호 직군 중심으로)”, 『근관절건강학회지』 제27권 제2호, 2020. pp.112-121.

[3] 최병하, 최승교, 조경산. “HTTP Outbound Traffic 감시를 통한 웹 공격의 효율적 탐지 기법”, 『한국컴퓨터정보학회논문지』 제16권 제1호, 2011. pp.125-132.

[4] 천우봉, 박원형, 정태명. "HTTP Get Flooding 기술을 이용한 APT(지능적 지속 위협)공격 도구의 설계와 구현", 『컴퓨터교육학회 논문지』 제14권 제6호, 2011. pp.65-73.

[5] 신민철, 조인휘. “효율적인 웹 브라우징을 위한 지능적 HTTP 압축 방식”, 『한국통신학회 학술대회논문집』, 2007. pp.36-39.

[6] 박채금, 노봉남, 김정일. “GET과 POST정보를 이용한 웹 어플리케이션 보안 시스템 설계”, 『한국정보과학회 학술발표논문집』 제32권 제2호, 2005. pp.142-144.

[7] 임종혁, 박재철, 김동국, 노봉남. “스크립트 파일 기반의 효율적인 웹 공격 탐지 프로파일링”, 『한국정보과학회 학술발표논문집』 제32권 제2호, 2006. pp.511-514.

[8] 오채연, 이준, 문형우, 노희준, 김규일, 권태웅. “효율적인 보안관제 지원을 위한 휴리스틱 기반 HTTP 웹 공격 자동 분류 방법”, 『한국정보과학회 학술발표논문집』, 2020. pp.992-994

[9] Pedro J. Roig, Salvador Alcaraz, Katja Gilly, and Carlos Juiz, "Algebraic Form al Modelling for HTTP Main Methods using ACP", 2019 23rd International Conference Electronics, June 2019. pp.1-6.