이미지 로깅서버 기반의 사용자 가용성 보장을 위한 연구

유승희, 조동섭
여자대학교 컴퓨터공학과

A Study for the Availability based on the Image Logging Server

Seung-Hee Yoo, Dong-Sub Cho
Dept. of Computer Science Engineering, Ewha Womans Univ.

Abstract - 인터넷의 급속한 발달로 정보화 사회에서 생활하고 있는 우리는 금융 서비스 등 다양한 웹 서비스를 제공받고 있다. 이와 같이 컴퓨터와 인터넷, 정보기술의 발전은 사용자에게 편리함을 가져다 주었다. 그러나 사이버 범죄라는 역기능을 가지게 되었다. 이를 위하여 이전 논문에서 우리는 이미지 로그파일을 제안하였다. 이미지 로그 파일이란 디지털 증거자료로써 디지털 포렌식으로 사용될 수 있도록 보안상 취약한 기존의 텍스트 파일 형태의 로그 파일을 보완하여 웹 페이지를 이미지로 저장한 파일이다.

본 논문에서는 이러한 이미지 로그 파일이 디지털 증거 자료 뿐만 아니라 웹 페이지의 무결성 문제가 발생하였을 경우 웹 페이지를 복구하기 위한 백업용으로써 사용될 수 있는 이미지 로그 파일에 대하여 기술하여 보고자 한다.

1. 서 론

현재 인터넷은 급속히 발달하였지만, 웹을 이용한 비즈니스 모델이 급속히 확산되고 계속적인 웹 페이지의 운용 및 유지보수로 인하여 링크가 끊어지는 등의 웹 페이지 무결성 문제가 자주 발생하고 있다. 대표적으로 하이퍼텍스트에서 하이퍼링크를 따라가다 보면 자주 깨진 링크들을 발견할 수 있다. 깨진 링크는 크게 두 가지가 있는데 하나는 참조되는 대상이 없는 현수 참조로 웹에서 흔히 볼 수 있는 "페이지를 표시할 수 없습니다"라고 메시지가 나오는 경우이고, 나머지 하나는 링크되는 자원이 자른 내용으로 바뀌어서 원래 링크가 가리키던 내용이 아닌 다른 내용을 참조하게 되는 오문 참조이다. 이런 현수 참조와 오문 참조를 DLP라고 한다[8]. 이러한 DLP가 발생하지 않도록 사전에 예방해야 하지만, 문제가 생겼을 경우에는 빨리 발견하여 수정할 수 있어야 한다. 이때 이미지 로그 파일은 백업용 파일로 아주 유용하게 사용될 수 있다.

또한 같은 URL page의 컨텐츠가 바뀌었을 경우, 이전 컨텐츠가 필요한 상황에서도 이미지 로그 파일이 사용될 수 있다.

본 논문에서는 이러한 경우 이미지 로깅 서버에 저장되어 있는 이미지 로그파일 서치 방법과 구현된 이미지 로깅 서버에 관하여 기술하고자 한다.

본 논문의 구성은 다음과 같다. 2장 본론에서는 관련 연구에 대하여 논하고, 구현된 이미지 로깅 서버에 대하여 알아본다. 그리고 이미지 로깅서버 사용 시나리오와 데이터 검색 방법에 관하여 논하고 마지막 3장에서 결론 및 향후 과제를 기술하고자 한다.

2. 본 론

2.1 관련 연구

관련 연구에서는 기존에 제안한 이미지 로그 파일의 특성과 이미지 로그 파일을 구현하고 사용함으로써 얻을 수 있는 장점에 대하여 기술하고자 한다.

2.1.1 이미지 로그 파일

현재 웹서버에서의 로그 파일은 확장자가 .txt인 텍스트 파일 형식으로 되어있다. 이는 파일 특성상 조작이 쉽고 보안상 취약한 문제점을 가지고 있다. 이를 보완하여 제안한 것이 바로 이미지 로그 파일이다.

확장자가 .jpeg인 이미지 로그 파일은 파일 특성상 텍스트 파일보다 조작이 훨씬 어려워 보안상 뛰어나기 때문에 디지털 증거자료로써도 좋은 디지털 포렌식이 될 수 있다.

클라이언트가 웹 서버에게 정보를 요청하고, 웹서버는 그에 대한 응답을 할 때마다 서버 소프트웨어는 로그 파일로 그 기록을 남긴다.

보통의 웹서버는 클라이언트의 요청이 들어오고 오류가 나는 등의 모든 상황을 로그 파일에 기록한다. 로그 파일에는 처리상황(status), 접근한 사용자의 아이피 또는 도메인(host), 사용자가 접근한 날짜와 시간(time), 요청사항(request)등이 기록된다. 즉, 로그 파일은 웹서버를 통해 이루어지는 모든 작업들에 대한 기록이라고 할 수 있다. 그리하여 컴퓨터에 불법으로 침입한 공격자[6]가 남긴 흔적이 저장되는 곳이 로그 파일이고, 그렇기 때문에 디지털 증거로 사용되고 활용되어 지는 것이 로그 파일이라 할 수 있다.

이미지 로그 파일은 파일 이름이 사용자가 접근한 날짜와 시간으로 접근한 사용자의 아이피 주소 폴더에 저장된다. 그리고 요청사항인 웹 URL 페이지의 이미지를 로깅하여 저장한 것이므로 텍스트 파일 형식인 로그 파일보다 많은 정보를 가지고 있다.

2.2. 구현된 이미지 로깅 서버 시스템

이미지 로깅 서버는 객체 독립적으로 웹 URL 페이지의 이미지를 저장하고 보존하며 제 3의 신뢰기간의 요청에 따라 이미지를 제공해 주는 데이터 베이스 서버이다.

본 서버는 제 3의 신뢰기간에서 받은 정보를 바탕으로 이미지 로그 파일을 만들어서 이미지 로그 DB에 저장한다.

이미지 로깅 서버 시스템의 구현환경은 다음과 같다.

- 운영체제 : Windows XP Sp2
- CPU : 3.2GHz
- 메모리 : 2GB
- 하드 디스트 : 200GB
- 구현 툴 : Visual Studio.net
- 구현 언어 : VC++, MFC(Microsoft Foundation Class)

이미지 로깅 기법에는 Guangming Software(US)사의 HTML SnapShot이라는 소프트웨어를 사용하였다[5]. 본 소프트웨어는 웹 URL 페이지를 이미지(jpeg)파일로 만들어주는 소프트웨어로써 Visual C++로 구현된 프로그램이다.

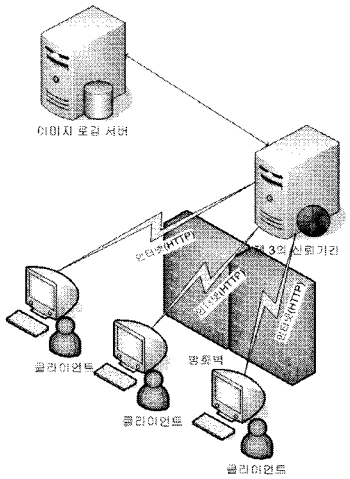
구현된 시스템에서는 제 3의 신뢰기간에서 받은 정보인 신뢰기간 클라이언트 IP Address와 요구사항으로 웹 URL 페이지를 이미지 파일로 만들어서 이미지 로그 DB에 저장시킨다.

이미지 파일의 이름은 제 3의 신뢰기간으로부터 받은 정보인 웹 로그 파일 생성시간으로 저장되며 신뢰기간 클라이언트의 IP 주소 폴더에 저장된다.

구현된 이미지 로깅 서버의 장점으로는 보존성, 보안성, 용이성, 독립성 등을 들 수 있다.

보존성은 웹 페이지의 DLP 문제가 발생하였을 경우나 컨텐츠의 내용이 변경되었을 경우에도 웹 페이지를 확인할 수 있고, 복구할 수 있다.

보안성은 조작이 어려운 이미지파일인 .jpeg파일 형태로 로그 파일을 저장함으로써 보안상 취약한 텍스트파일 형식의 로그파일의 문제점을 개선할 수 있다.



〈그림 1〉 시스템 구성도

본 논문은 한국 학술 진흥 재단의 BK21 2단계 프로젝트의 지원으로 작성되었습니다.

용이성은 파일 시스템에 디렉토리 관리 기능을 추가하여 디렉토리를 만들고 파일을 저장함으로써 데이터 관리 및 사용을 용이하게 한 것이고 독립성은 이미지 로깅 서버는 독립적으로 동작하며, 웹 서버에는 아무런 영향을 주지 않고 이미지 로그파일을 생성, 관리한다. 그리하여 높은 처리율을 가지고 용량 문제에 자유로우며 확장성이 효율적이다.

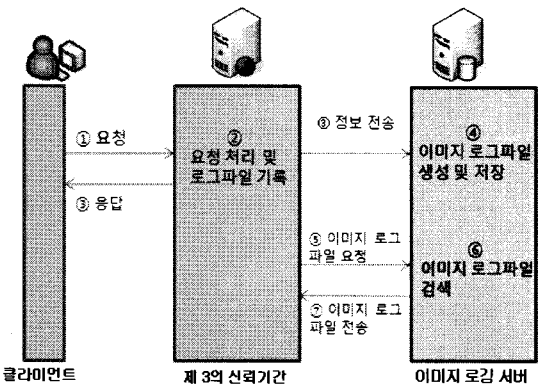
2.3 이미지 로그 파일 요청 시나리오

이미지 로깅 서버에서 이미지 로깅 서비스를 받고 있는 제 3의 신뢰기간에서는 이미지 로그 파일을 요청할 수 있다. 본래의 목적은 사이버 범죄가 발생하였을때 디지털 증거자료로써 사용하기 위함이지만, 이 외에도 DLP 문제가 발생하였을 경우나 이전 웹 페이지를 확인하고자 할 때 백업용으로 요청할 수 있다.

다음은 제 3의 신뢰기간에 클라이언트가 접속하기부터 이미지 로그 파일 요청까지의 시나리오이다.

1. 2008년 1월 1일 11시 26분 54초에 아이피 주소가 203.255.177.XXX인 클라이언트가 제 3의 신뢰기간에 접속하였다.
2. 제 3의 신뢰기간은 클라이언트의 요청을 처리하여 클라이언트에게 응답한 후 로그 파일을 기록한다.
3. 로그파일 기록과 동시에 제 3의 신뢰기간은 이미지 로깅 서버에게 클라이언트의 시간, 아이피 주소, 요구사항(URL page)등의 정보를 보낸다.
4. 이미지 로깅 서버는 그 정보를 바탕으로 클라이언트가 방문 중인 URL page를 이미지 로깅 하여서 이미지 로그 DB에 저장한다.
5. 제 3의 신뢰기간에서는 2007년 12월 22일의 웹 페이지를 확인하고자 하여서 이미지 로깅 서버에 그때의 이미지 로그 파일을 요청한다.
6. 요청을 받은 이미지 로깅 서버는 요청받은 날짜의 이미지 로그 파일을 검색하여 제 3의 신뢰기간에 이미지 로그 파일을 전송한다.

이와 같은 시나리오는 아래의 (그림 2)에 나타나 있다.



<그림 2> 시나리오

2.3.1 이미지 로그 파일 검색 방법

이미지 로깅 서버는 다음과 같은 방법으로 이미지 로그 파일을 검색할 수 있다. (표 1)은 검색을 위한 이미지 로그 파일 데이터 베이스 테이블이고 다음은 이미지 로그파일 검색에 필요한 파라미터들이다.

O : 원본 데이터
D : 데이터(URL 페이지의 이미지 로그 파일)
L : 손실된 데이터
T_0 : 현재 시간
T_{-n} : 과거 시간
T_n : 미래 시간
$D(n) = 0$, 데이터 존재
$D(n) = 1$, DLP 발생 데이터
$T_{-n} < T_0 < T_n$
$D(t_0) = O$

<표 1> 이미지 로그파일 데이터베이스 테이블

	T_{0-n}	...	T_{0-3}	T_{0-2}	T_{0-1}	T_0
$D(0)$	0	...	1	0	1	1
$D(1)$	1	...	1	0	0	1
$D(2)$	0	...	1	1	1	1
$D(3)$	1	...	1	1	1	1
...	...	...	...	...	...	...
$D(n)$	1	...	1	1	1	1

제 3의 신뢰기간에서 T_{0-1}의 $D(2)$ 요구 시,

T_{0-1}의 $D(2)$를 탐색한다. $D(2) = 1$이면 데이터가 보존되고 있는 것이므로 요청사항에 따라 $D(2)$의 데이터를 전송한다.

제 3의 신뢰기간에서 T_{0-1}의 $D(1)$ 요구 시,

T_{0-1}의 $D(1)$을 탐색한다. 이 경우 $D(1) = 0$이므로 이때의 데이터는 존재하지 않는 것이다. 이럴 경우 T_{0-1}과 가장 가까운 ΔT의 값을 검색한다.

$\Delta T_{0-1} \cong T_{0-1}$

그리고 $D(\Delta T_{0-1})$

이 때, $D(\Delta T_{0-1}) \neq 0$이면, $D(\Delta T_{0-1}) = O$이므로, $D(\Delta T_{0-1})$의 데이터를 전송한다.

3. 결 론

웹은 양적으로나 복잡도에 있어서 놀라운 속도로 성장하고 있다. 이에 따라 웹사이트를 이용한 홍보나 기업의 이익을 목적으로 하는 웹사이트들이 급속히 증가하면서 웹 페이지의 추가, 삭제, 갱신 등의 끊임없는 작업이 필요하다. 이로 인하여 링크가 끊어지는 등의 문제가 발생할 수 있는데, 이런 경우 큰 손실을 가져올 수 있다.

본 논문에서는 디지털 포렌식으로 활용되었던 이미지 로깅 서버를 바탕으로 웹 페이지 가용성 보장에 관하여 연구해 보았다. DLP 문제나 이전 웹 페이지 컨텐츠를 복구하여야 할 문제가 발생 하였을 때 이미지 로그 파일은 사이트 무결성을 유지하여 줄 수 있다. 이를 위해서 본 논문에서는 제 3의 신뢰기간과 이미지 로깅 서버가 어떻게 작동하는지 살펴보았고, 제 3의 신뢰기간의 이미지 로그 파일 요청 시 이미지 로깅 서버의 이미지 로그 파일 검색 및 전송에 대하여 살펴 보았다.

향후에는 중복되는 이미지 로그 파일을 알고리즘 등을 통하여 분류 할 수 있는 마이닝 기법, 디지털 포렌식 자료의 추출에 사용되는 포렌식 알고리즘 등의 구체적인 방법을 고안해야 할 것이고 지속적인 연구가 필요하다.

[참 고 문 헌]

[1] A.R Arasteh, M. Debbabi, A. Sakha, M. Saleh, "Analyzing multiple logs for forensic evidence", Digital Investigation, Elxtronic Edition, Volume 4, pp.82-91, 2007

[2] Rahul Bhaskar, "State and local law enforcement is not ready for a cyber Katrina", Communications of the ACM, 49(2), pp81-83. 2006.

[3] Linda Volonino, "Computer forensics and electronic discovery: The new management challenge", Computers & Security, 25(2), pp91-96, 2006.

[4] R. Finlayson, D. Cheriton, "Log Files: An Extended File Service Exploiting Write-Once Storage", ACM, 21(5), pp137-148, 1987

[5] http://www.guangmingsoft.net/

[6] http://news.netcraft.com/

[7] H. Custer, Inside Windows NT, Microsoft Press, 1993

[8] 김원중, "웹페이지의 무결성 유지를 위한 WPMS(Web Page Management System)시스템", 정보통신기초기술연구과제, 2004.06.30

[9]김선우, "윈도우 네트워크 프로그래밍 : TCP/IP 소켓 프로그래밍", 한빛미디어, 200