한국컴퓨터정보학회 하계학술대회 논문집 제20권 제2호 (2012. 7)

보안성이 취약한 사용자 계정 관리를 위한 웹 로그 분석기

박기홍°, 이진관*
°*군산대학교 컴퓨터정보공학과
e-mail:{spacepark, leejinkwan}@kunsan.ac.kr

The Web Log Analyser for Managing User Account having Weak Security

Ki-Hong Park°, Jin-Kwan Lee*
°*Dept of Computer Information Engineering, Kunsan National University

● 요 약 ●

인터넷이 확산과 더불어 보안의 문제도 증가하고 있다. 이로 인해 네트워크 보안과 서비스에 대한 관리자의 책임 또한 더욱더 중요시 되고 있다. 본 논문에서는 웹로그를 분석하여 웹호스팅 환경에서 장시간 사용되지 않아 보안성이 약한 사용자 계정을 관리자로 하여금 시스템 보안의 틈새를 찾고 이를 해결할 수 있는 방안을 제시하였다. 이를 위해 WLA(Web Log Analyzer)를 구현하여 웹서버가 수행될 때 기록되는 각각의 로그를 분석한다. 그 결과 웹호스팅을 사용한 계정 이름의 수를 포함한 UUL(Used User List)를 구축하고 일정기간 사용하지 않는 호스팅 서비스 이용자를 찾아내고, 관리할 수 있게 한다.

키워드: 웹로그(Web Log), 웹서비스(Web Services), WLA(Web Log Analyser)

I. 서론

인터넷을 이용하면서 사용자들은 자신의 계정이 있는 서버를 이용하여 웹 서비스, 파일 전송 서비스, 터미널 시비스 등을 수행하기을 원한다. 그러나 특정 업무에 종사하지 않을 경우 일시적인 사용에 그치고 만다. 이로 인해 각 서버에는 사용하지 않는 각 서비스계정이 산재되어 시스템의 부하 및 외부 침입의 경로가 되기도 한다[1]. 특히 장시간 사용하지 않는 홈페이지 계정의 경우 펄(Perl)스크립트나 PHP들 일반적인 언어로 작성된 공개 게시판을 보유하고 있다[2]. 이들은 수시로 보안 업데이트 등의 관리가 필요하지만 사용자는 일시적인 구축만 해놓고 방치하는 경우가 종종 발생한다. 본 논문에서는 이렇게 사용되지 않는 계정을 웹로그 분석기를 통해 사용여부를 추출하고 일정기간 사용되지 않은 것으로 나타날 경우 서비스를 사용하지 못하도록 조치를 취하고자 한다. 단 중요한 개인정보가 계정에 포함되어 있을 수 있기 때문에 계정을 삭제하지 않고 단지 특정 서비스들만을 정지시킨다. 본 논문의 구성은 다음과 같다. 2장에서 관련연구에서는 사용자 사용빈드측정과 문제점을 논하고, 3장에서는 웹로그 분석기에 대해 설명한다. 4장에서는 웹로그 분석기의 성능을 평가하고 마지막 5장에서는 결론 및 향후과제를 논하고자 한다.

II. 관련 연구

1. 관련연구

1.1 사용자 사용빈도 측정

사용자가 홈페이지를 제작하기 위해서는 터미널 서비스(SSH, Telnet) 또는 FTP를 통하여 웹서버에 접근하여 작업을 한다고 가정하자. 우선 처음에 사용자가 TCP프로토콜로 접속을 하였다면 터미널 서비스, FTP 서버는 그림 1과 같이 사용자 인증을 통하여 사용자를 사용자 디렉토리에 접근하게 해준다. 이때 그 접속 과정(어디서 어떤 IP에서 접속을 하였는지 등)을 로그 서비스에게 알려줌으로 로그가 기록되게 된다[3].

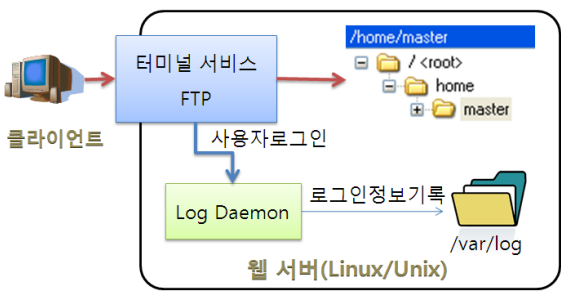

그림 1. 웹서버 시스템 로깅
Fig. 1. Web Server System Logging

그러나 일정 규모의 웹서버의 경우는 상당히 많은 사용자를 보유하고 있어 이 같은 경우 상당한 시간이 흘러 더 이상 사용하지 않는 경우가 정지를 시키거나 삭제를 해야 하는 경우가 생기게 된다. 문제는 어느 기준을 두고 조치를 취할 것인가이다. 가장 간단한 방법은 디렉토리를 검색하여 최근의 데이터가 없을 경우, 그 사용자는 더 이상 사용을 하지 않는다고 가정을 하는 것이다. 그런데 문제는 사용자가 홈페이지나 사용자 디렉토리 내의 파일을 변경하지 않았다고 해서 사용하지 않고 있는 것이 아니라는 것이다.

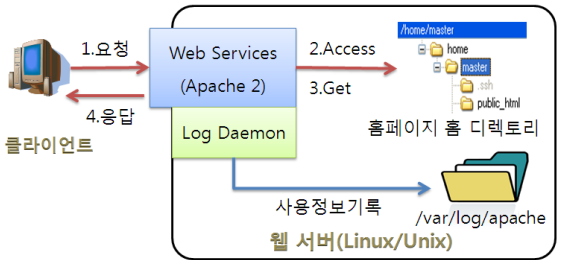

그림 2. 웹서버 로깅
Fig. 2. Web Server Logging

그림 2를 보게 되면먼저 사용자가 웹서버의 웹서비스에게 특정 페이지를 요청하게 된다. 그러면 웹서비스는 요청한 디렉토리를 확인한 후 응답을 돌려주게 되는데 여기서 응답은 특정서비스 페이지이거나 오류 코드 일 수 도 있다[4].

여기서 문제가 나오게 되는데 분명히 사용자 디렉토리는 사용되고 있는 상태이나 어느 파일 하 변경된 것이 없으나 오랜 기간 사용하지 않는 사용자를 추려내는 작업에서 제외 대상이 될 수가 없게 된다. 다행히 그림 2에서 보게 되면 사용내역을 기록하는 것을 볼 수 있는데 이 기록은 특정 사용내역이 아니라 접속 상황이나 서비스의 상태 등을 모두 기록한다. 따라서 이 로그에는 사용자가 요청하고 응답한 내용 또한 들어가게 된다.

III. 본론

1. Web Log Analyzer(WLA)

1.1 로그를 통한 이용자 검출

먼저 로그를 분석하기 위해서는 소스(Source)가 있어야 할 것이다. 정책에 따라 사용자에게 지원하는 서비스가 다루는데 본 논문에서는 다음 서비스가 사용 가능하다는 가정하에 진행을 하도록 하였다.

- 웹서비스(WWW) : 웹서비스는 가장 흔히 사용되고 있는 서버의 정책이라고 할 수 있다. 본 논문에서는 아파치 2.x 버전을 사용하였다.
- 파일전송서비스(FTP) : 계정 내의 파일이나 디렉토리를 관리하기 위한 서비스로 여기서는 홈페이지 내용을 갱신할 목적으로 사용한다고 가정한다.
- 터미널서비스(SSH, Telnet) : 터미널 서비스는 사용자가 서버에 접속하여 원격으로 여러 작업을 수행하고 파일 편집 작업이나 명령을 내리기 위한 텍스트모드의 환경을 말한다.

그림 3은 웹로그분석기(WLA)의 위치를 보여주는데 WLA가 어떠한 로그를 읽어 들여야 하는지는 Log List Properties라는 곳에 관리자가 직접 명시를 해주게 된다[5].

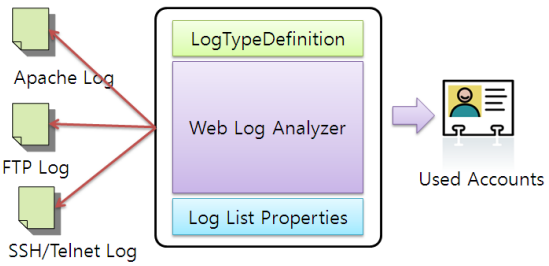

그림 3. 웹로그 분석기
Fig. 3. Web Log Analyzer

Log List Properties는 단순한 텍스트 파일로서 WLA가 시작 시 읽어 들이는 로그 정보 파일이라고 할 수 있다. 각 프로퍼티는 '='을 기준으로 키(key)와 값(value)으로 매칭 되게 된다. 그림 4는 Log List Properties의 내용이다. 여기서는 XXX_LOG형식을 사용한다. XXX는 어떤 서비스인가를 나타내고 있으며 로그 형식 정의(LogType Definition)파일에 이 서비스가 어떤 형식을 사용하여 로그를 남기는지 반드시 정의해야 한다. 로그 정의가 없을 경우, 이 항목은 무시된다.

```
APACHE_LOG = /var/log/apache2/access_log
FTP_LOG = /var/log/auth_log
SSH_LOG = /var/log/auth.log
```

그림 4. 로그 목록 속성
Fig. 4. Log List Properties

log_type_def.properties 파일에 들어가는 항목은 두가지 형식으로 나누었다. Standalone은 서비스에 종속된 로그 시스템을 이용하는 서비스이다. 아파치 웹서버가 그 예이다. Syslog는 시스템에서 제공하는 로그 서비스를 이용하는 서비스로 SSH, FTP, Telnet이 이에 해당한다.

WLA는 이 형식을 프로퍼티 항목으로 넣어 놓고 로그 리스트에 정의된 서비스들의 타입을 알아내도록 되어있다.

범용으로 쓰이고 있는 아파치 서버의 경우 그림 5의 로그 형태를 가지고 있다.

먼저 각 서비스는 정의된 형식을 사용하여 로그를 저장하는데 이 형식을 정규식 패턴을 사용하여 항목들을 뽑아내도록 한다[6,7].

그림 5대로 정규표현식을 이용하게 되면 그룹화가 가능하게 된다. 차후에 필요한 항목만 따로 모아 여러 가지 분석을 할 수 있게 된다. 본 논문에서 필요한 사항은 사용 기록이므로 G3, G4 그룹

의 사용자 - '~'로 시작하는 부분이 사용자이다. 사용자가 요청하는 G4 항목의 경우 여러 가지 형식으로 저장되어 있을 수 있는데 본 논문에서는 GET, POST, HEAD 만을 검색에 사용하였다. 모든 로그가 분석대상이 될 수 없다. 왜냐하면 모든 로그를 분석하게 되면 그만큼 비용이 들기 때문에 가급적 필요없는 로그는 버리는 것이 작업 시간을 줄이는데 도움이 되기 때문이다. 또한 포함되는 않는 항목은 해킹이나, 사용자의 입력오류 등에 기인하고 있으므로 차후에 다른 목적으로 이용하면 될 것이다.

1.2 사용자 객체 및 사용자 리스트 객체의 생성

WLA가 처리하고 난 최종 결과물로 본 논문에서는 UUS(User Usage Shifter)로 보내기 위한 일종의 임시 장소라고 할 수 있다. 이는 MVC(Model, View, Controller)모델과 같이 View와 Control사이에서 쓰이는 것고 비슷하다고 할 수 있다.

모든 로그 분석이 끝났을 경우 UUS에게 파일 형태의 UsedUserList를 넘기게 된다. 여기서 파일 형태로 저장하는 이유는 단순한 파라미터 형식으로 넘길 경우, 단지 UUS만이 사용하게 되지만 오브젝트 파일 형태로 저장할 경우 다른 관리프로그램에서도 불러 들여 사용 가능하기 때문이다.

1.3 시스템 구현

실험 평가를 위해 Linux PC를 구성하고 다중 도메인 환경에서 웹서버를 실험 운영하였다.

실질적으로 디렉토리 서버와 통신하는 부분은 UNIX와 Linux가 약간씩 차이가 있으나 구현부분에 대해서는 언급하지 않았다. 왜냐하면 LDAP프로토콜을 사용하여 통신하는 부분은 어느 플랫폼에서나 동일하기 때문이다[8].

표 1. 시스템 환경
Table. 1. System environment

분류	세부내용	
H/W	Intel Celeron CPU E3400 Memory 2G HDD 500G	
S/W	OS	Gentoo Linux
	LDAP	OpenLDAP 2.4.30
	LDAP 모듈	nss_ldap v264 pam_ldap v183
	Web 서버	Apache 2.2.22
	SSH 서버	OpenSSH 5.9
	FTP 서버	PureFTP 1.0.35

표 1은 가상의 환경을 구현하기 위해 사용된 시스템이다. 실제 로그 데이터는 운영 중인 서버를 테스트하기는 어렵기 때문에 가상으로 구성하였고 또한 사용자 정보를 분석하기 위해 사용자 파일에 관한 사항들은 직접적으로 서버에 영향을 미치지 않으므로 운영중인 서버 내에서 작업을 하였다. 다음으로 서비스 로그 분석을 위해 실제 운영중인 서버에서 로그 파일을 평가 시스템에 복사하여 분석하고 평가하였으며 로그 파일은 최근부터 3개월 전의 자료를 가지고 분석하였다.

평가방법은 Unix/Linux 검색 툴과 WLA를 이용한 사용자 계정 검색 및 속도를 비교하였다.

사용자 계정이 사용되었는지 알아내기 위해서는 파일 시스템이 변경되었는지 알아내기 위해서는 파일 시스템이 변경되었는가를 확인하면 된다.

검색도구로는 이 파일 시스템을 검사하기 때문에 상당한 I/O시간을 요구하게 되는데 각 3, 7, 14, 30, 90일 간 사용된 사용자 계정을 검색하는데 걸린시간은 대략 45초에서 120초가 소요 되었다.

```
time find /home -mtime 3 -exec ls -lrt {} ₩;
```

위와 같은 방법으로 특정 기간 동안 사용된 파일 만을 검색, 어떤 사용자의 계정의 데이터가 새로 갱신되었는지 파악할 수 있다. 하지만 위의 명령을 사용하게 되면 상당한 파일 리스트가 출력되므로 필터링을 거쳐 사용자 정보를 뽑아내야 한다. Unix/Linux의 기본 명령을 사용하여 사용된 계정을 찾는 부분에서 최대의 단점은 파일 시스템 전부를 검사해야 하기 때문에 사용되는 않은 파일도 모두 검사를 해야 한다는 것이다.

그림 5. 검색시간 비교
Fig. 5. Compare Search Time

WLA를 통해 최근 3개월간의 서비스 사용자 수를 검색한 결과 전체 계정 중 18%도 못미치는 사용률을 확인하였다.

WLA와 Unix/Linux 기본 툴을 이용한 검색 시간 비교를 한 결과는 그림 5와 같다. 하루 동안 사용된 사용자 계정을 알아보는 속도를 비교한 것이다. 파일 시스템을 검색한 것보다 로그를 통하여 사용된 사용자 계정을 검색하는 시간이 훨씬 더 빠른 것을 볼 수 있다. 이는 파일 시스템을 사용하는 서버의 부하에도 영향을 미치는 것으로 짧은 시간에 검색을 마치는 것이 서버 사용률에도 효과적이다.

Ⅳ. 결론

본 연구에서는 사용되지 않아 보안성이 취약해진 사용자 계정의 보안 위협을 해결하기 위해 사용된 사용자 계정을 웹로그 분석기를 통해 로그를 분석하여 일정기간 사용되지 않는 계정을 추출

하고 해당 서비스를 사용하지 못하도록 하였다. 이를 위해 로그의 패턴에서 웹서비스의 사용되는 영역을 추출한 결과 약 18%만 일정기간 사용되는 것으로 파악되어 나머지 82%의 계정에 대해 보안을 유지할 수 있었다. 또한 유닉스 관리툴을 이용할 경우 사용자의 사용정도를 파악하는데 많은 시간이 소요됐으나 제안한 분석기를 통해 분석할 경우 약 75%의 속도 증진을 볼 수 있었다. 향후 로그 패턴 분석기를 다양한 서비스에 접목시키고 보다 쉽게 활용하능하도록 GUI환경의 시스템을 구축되어야 할 것이다. 또한 최근의 홈페이지 제작시 주로 이용되는 제로보드나 그누드의 사용유무를 체크하여 버전이 낮은 공개게시판을 사용할 때 사용자에게 안내할 수 있어야 할 것이다.

참고문헌

[1] “Analysis and Diagnosis for Unix System Security Vulnerability”, Korea National Computerization Agency, 1995.

[2] Young-Ho Park, “Analysis and Improvement for Incheon Area Campus Network Security”, MS Thesis, University of Incheon, 2002.

[3] HP, “Distributed Systems Administration Utilities”, pp.70-73, 2005.

[4] “Unix Desk Reference”, Peter Dyson, Sybex, 1997.

[5] Dae-hyung Lee Et.5, “LDAP-based user account management system implementation”, Proceedings of Society for Internet Information, 2006.

[6] R. BrandtSteven, “Regex Recipes v1.0”, “Regular Expressions in Java http://www.javaregex.com, 2012.

[7] “Linux-Pam”, Kernel.org, http://www.kernel.org/pub /linux/libs/pam, 2012

[8] Rob Weltman Dahbura Tony, “LDAP Programming with JAVA”, Addison-Wesley, pp.4-31, 2000.