통합 로그트리를 이용한 침입분석

김 홍철° , 김 건우, 박 보석, 장 희진, 박 정현, 김 상욱
경북대학교 컴퓨터과학과 컴퓨터언어/멀티미디어 연구실
{hckim, parkbs, janghj, swkim}@cs.knu.ac.kr
kimgw@etri.re.kr, parkjh@kisa.or.kr

Intrusion Analysis using Integrated Log Tree

Hong-Chul Kim° , Gun-Woo Kim, Bo-Seok Park, Hee-Jin Jang,
Jung-Hyun Park, Sang-Wook Kim
Dept. of Computer Science, Kyungpook National University

요 약

최근 전산망 컴퓨터 침입사고를 미연에 방지하고 사고 발생 시 이에 대처하는 전산망 보안 시스템의 필요성이 그 어느 때보다도 높아지고 있다. 그러나 해킹기술의 발달, 컴퓨터 전산망의 복잡성 및 대규모화, 그리고 TCP/IP Internet Protocol Suite가 가지고 있는 근본적인 보안상의 문제점으로 인해 전산망 침입을 미연에 방지하고 대처하는 것이 현실적으로 매우 힘들다. 본 논문에서는 호스트 컴퓨터 시스템의 각 로그파일에 대한 로그트리를 하나의 로그트리로 통합하여 시스템 정보를 수집하고 침입자의 행동을 효율적으로 분석하는 기능을 가지고 있는 서버-에이전트 기반의 침입 분석 에이전트 시스템을 제시한다.

1. 서론

침입의 의미에 대해서 1980년 Anderson은 'Causing DoS(Denial of Service), Creating Backdoor(Trojan Hoarse), Planting Viruses, Exploiting Software Vulnerability 등과 같은 행위를 통해서 컴퓨터의 자원이라 할 수 있는 무결성(Integrity), 기밀성(Confidentiality), 가용성(Availability)을 저하시키는 일련의 행위'라고 정의하고 있다. 컴퓨터 및 전산망에서의 이러한 침입사고는 필연적으로 해당 컴퓨터의 피해를 야기 시키므로 우리는 그러한 피해를 감지하고 분석하며, 추후의 사고를 미연에 방지할 수 있는 조치를 취해야 한다. 이것이 바로 최근 인터넷 전산망 컴퓨터 침입사고에 대처할 수 있는 도구로 여겨지는 전산망 종합 정보보호 시스템의 주된 기능이라고 할 수 있다. 아직까지는 침입사고가 발생했을 경우 그 조치는 대부분 컴퓨터 네트워크 및 보안 전문가의 수 작업을 거치는 경우가 많다. 하지만, 최근 들어 침입사고가 기하급수적으로 늘어나기 때문에 그러한 작업을 자동적이고 효율적으로 수행할 수 있는 자동화된 침입 분석 시스템(Autonomous Intrusion Analysis System)이 그 어느 때보다도 절실히 요구되고 있다.

이에, 본 논문에서는 시스템의 여러 로그 파일에 저장되어 있는 로그 정보를 하나의 로그트리로 통합하여 침입분석에 이용하는 서버-에이전트 방식의 침입 분석 시스템(Intrusion Analysis System)을 보이고자 한다.

앞으로 2절에서는 서버-에이전트 방식의 보안 시스템 아키텍처를 구체적으로 설명하고, 3절에서는 에이전트가 수행하는 침입분석 기능의 핵심기술인 통합 로그트리의 생성 및 이용 기법에 대해서 설명한다. 마지막으로, 4절에서는 결론 및 향후 연구방향을 제시하며 본 논문의 끝을 맺도록 하겠다.

2. 서버-에이전트 기반 보안 시스템의 구조 및 기능

서버-에이전트 기반 보안 시스템은 하나의 서버와 침입피해를 입을 수 있는 다수의 호스트 컴퓨터에 탑재된 에이전트들로 구성된다.

서버 시스템은 보안 관리자가 네트워크 상의 컴퓨터들을 관리하는데 최대한의 가용성을 제공하기 위해서 웹 브라우저를 통한 관리 인터페이스를 제공한다. 이렇게 함으로써 관리자는 적절한 인증 절차를 거치기만 한다면 어느 곳에서라도 웹 브라우저를 통해서 시스템의 보안 상황을 확인하고 관리할 수 있다. 관리자는 웹을 통해서 해당 호스트 컴퓨터에 탑재된 에이전트를 선택해서 보다 구체적인 보안상황도 확인 할 수 있다. 또한, 에이전트로부터 침입사실을 통보 받은 경우에는 Pager나 Mobile Phone등의 통신수단을 이용해서 관리자에게 침입신호 및 관련 정보를 보고한다.

에이전트 시스템은 호스트 컴퓨터의 침입발생을 감지하고 안전한 통신 프로토콜 채널을 이용한 인터페이스를 통해서 서버로 전달한다. 또한, 여러 가지의 시스템 및 접속 로그 정보를 하나의 로그트리로 통합하여 침입피해 및 침입패턴을 분석한다.

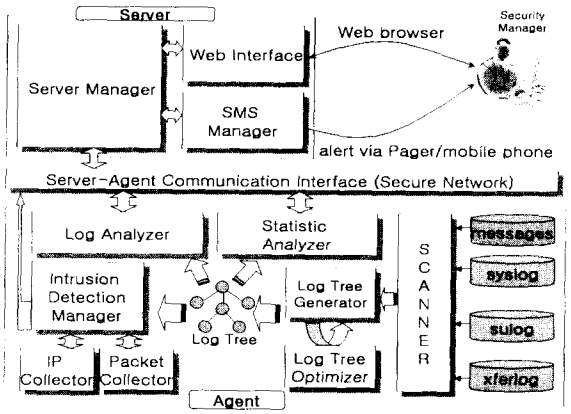

그림 1. 서버-에이전트 보안 시스템의 구조

2.1 서버 시스템

서버 시스템은 각각의 내부적인 기능에 따라서 크게 Server Manager, SMS(Short Message Service) Manager, Web Interface, 그리고 Server-Agent Communication Interface 모듈로 구성되어 있다.

Server Manager는 웹을 통해서 접속하는 전산망 보안 관리자에 대한 인증(authentication)작업을 수행하며, 특정 에이전트를 구분해서 관리하기 위하여 IP address와 System OS type, ID, password 정보를 보관한다. 이때 서버는 호스트 컴퓨터에 에이전트가 설치되어 있지 않을 경우 ftp와 shell script를 통해서 에이전트 설치 및 초기화 작업을 수행할 수 있다. 또한, 서버는 에이전트를 통해서 취약 파일 및 시스템 파일의 변경 여부조사, 각 사용자의 파일에 대한 검색, 파일 보기, 자동 검색, syslog, messages 로그파일조사, 네트워크 서비스 및 access_log 조사, 통합적인 로그조사, 임의의 사용자에 의해서 실행된 프로세스에 대한 조사작업을 수행 할 수 있다.

SMS(Short Message Service) Manager는 특정 시스템에 대한 침입을 감지했다는 정보가 해당 에이전트로부터 전달될 경우 Pager, Mobile phone, 또는 e-mail과 같은 통신 도구를 이용해서 관리자에게 그 사실을 알려주는 기능을 수행한다.

Web Interface는 관리자가 웹 브라우저를 통해서 일련의 인증 과정을 거쳐 서버 시스템에 접속하여 전산망보안 관련 작업을 수행할 수 있도록 해 주는 것으로서, 에이전트과 통신하는 서버측의 CGI program으로 구현되어 있다.

Server-Agent Communication Interface는 서버와 각 에이전트간의 통신을 위한 인터페이스로서, 서로가 신뢰할 수 있는 안전한 보안 프로토콜을 기반으로 구현된다. 위의 그림 1에서 볼 수 있듯이, 에이전트의 Log Analyzer, Statistic Analyzer, 그리고 Intrusion Detection Manager와 직접 통신을 해야하는 서버측의 Server Manager는 이와 같은 안전한 통신 인터페이스를 통해서만 에이전트와 서로 통신한다.

2.2 에이전트 시스템

에이전트는 Log Analyzer, Statistic Analyzer, Intrusion Detection Manager, Intrusion Reporting Manager, Log Tree Generator, Log Tree Scanner등으로 구성된다.

Log Analyzer는 이와 같이 구성된 통합 Log Tree를 기반 Detection Manager는 해당 에이전트가 탑재된 호스트 컴퓨터에 대한 세부적인 보안정책을 설정하고 관리한다. 이와 같이 설정된 정책을 바탕으로 에이전트는 네트워크 및 시스템 침입 탐지를 수행한다. 네트워크 레벨의 침입 탐지는 IP collector와 Packet collector를 이용해서 수행하며, 시스템 레벨의 침입 탐지는 Intrusion Pattern Rule Generator를 통해서 수행한다. 침입 감지시 에이전트는 설정되어 있는 정책에 따라서 해당 침입자를 컴퓨터로부터 disconnecting 시키고, 침입에 사용된 ID를 사용중지 시킨다. 또한 침입자의 로그 정보를 저장하고 서버측에 이러한 침입정보를 보고하여 최종적으로는 보안 관리자가 그 사실을 보고 받을 수 있도록 해 준다. 마지막으로, 에이전트는 새로운 침입 패턴이 발견되었을 경우 Pattern Generator를 이용하여 그 정보를 저장하여 이후의 침입 감지에 적용할 수 있도록 한다.

Intrusion Reporting Manager는 침입 발생 시 해당 호스트 컴퓨터의 IP address, ID/password, Platform, 침입자의 신원, 침입 시간 등과 같은 정보를 서버 시스템에 전달한다.

Log Analyzer는 호스트 컴퓨터의 messages, syslog, access_log, sulog, utmp(x), wtmp(x), 등과 같은 시스템의 주요 로그 파일을 읽어내기 위해서 Log Scanner를 이용한다. Log Scanner는 Log configuration file을 참조하여 해당 로그 파일을 읽어들여서 토큰화 한 후 로그 레코드를 생성한다. 이때 읽어들이는 로그 파일은 대략 다음과 같다:

/var/adm/sulog /var/adm/lastlog
/var/log/xferlog /var/log/messages.*
/var/adm/pacct /var/adm/access_log
/var/log/syslog.*

생성된 로그 레코드들은 각각 Linked List의 형태로 메모리에 저장되어 분석된다. 이러한 과정을 거쳐서 생성된 각 로그 파일들은 Log Tree Optimizer에 의해서 수정된 후 Log Tree Generator에 의해서 하나의 로그트리로 통합된다. (그림 2)

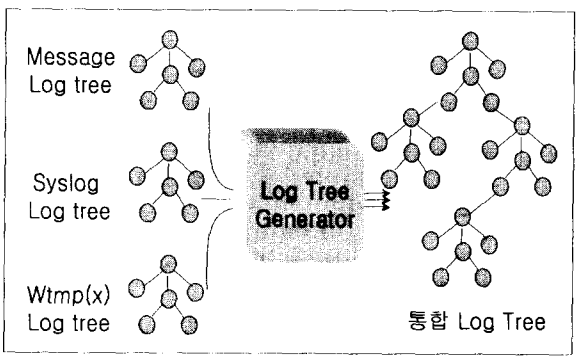

그림 2. 통합 로그트리의 생성

3. 통합 로그트리

Log Scanner에 의해 생성된 각 로그 파일의 리스트는 Log Tree Generator에 의해서 하나의 통합된 Log Tree로 구성된다. 통합 로그트리는 사용자 계정, 시간, 접속 호스트 및 포트, 프로세스와 같은 요소별로 indexing된다. 이와같이 통합 로그트리를 구성하는 이유는, 각 로그별 트리를 따로 관리할 경우 통합적인 로그정보의 접근이 필요한 해킹흔적 분석과 같은 작업을 하는 경우 모든 로그를 따로 처리해야 하므로 각 로그

정보를 연계시키기 힘들다. 반면, 통합 로그트리를 이용하면 그와 같은 로그 정보의 연관성을 쉽게 부여할 수 있다. 아래 그림 3은 통합 Log Tree 자료구조를 보여주고 있다.

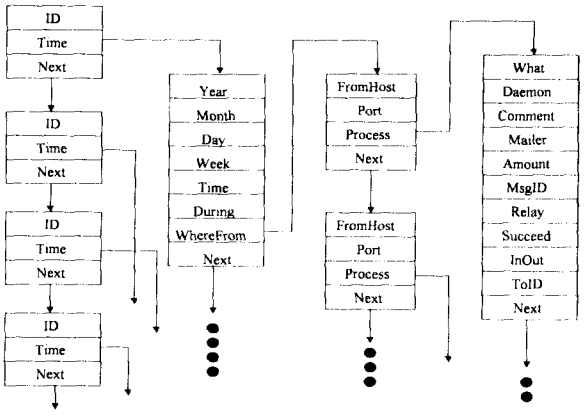

그림 3. 통합 로그트리의 자료구조

3.1 로그트리의 최적화

로그 트리는 정상적인 로그 정보도 가지고 있지만, 반대로 모호하거나 비정상적인 로그 정보도 가지고 있을 수 있다. 따라서, 그러한 부분을 로그트리에서 제거/수정시켜 모든 로그 트리의 내용이 정상적인 정보를 가지도록 할 필요가 있다. 이와 같은 작업을 로그 트리에 대한 최적화라 하며, 이는 곧 트리에서 모호성(Ambiguity)을 제거하고 의미있는 다른 로그 정보들에 대해서는 서로 심볼릭 링크로서 연결시켜 준다.

모든 로그 파일에 확실한 정보만이 존재하는 것은 아니다. 예를 들어, 어떤 로그 레코드는 사용자나 접속 호스트를 알수 없는 경우가 있거나, 또는 접속 포트번호와 호스트를 알 수 없을 경우도 있을 것이다. 이와 같은 경우, 모든 로그 상황을 종합해보면 어떤 사용자가 어떤 호스트로부터 접속을 시도했는가를 유추할 수 있다. 이와 같은 유추의 기준은, 모호한 로그 레코드의 접속 시간이나 포트 넘버 등의 모든 로그 정보를 알아낸 후, 이와 같은 로그 정보를 가지는 노드를 찾으면 된다. 예를 들어, 어떤 로그 레코드의 시간이 2000년 1월1일 12시를 가리키고, 접속 포트가 pts/13이라면 다른 로그 파일들 중에서 같은 시간대에 같은 포트 번호를 가지는 로그 레코드를 찾으면 된다. 이때, 같은 시간대라는 개념이 모호할 수 있으므로 wtpm(x) 로그 파일을 참조하여 모호한 로그가 발생한 시간대를 포함하는 모든 로그 정보를 추출한 후, 이 시간 사이에 발생된 모든 프로세스들을 후보 레코드(Candidate Record)로 만든다. 이렇게 한 후, 해당 포트 번호가 같은 정보가 발견되면 후보 레코드의 범위를 좁혀가면서 관련 프로세스를 검색할 수 있다. 물론, 아무런 노드도 발견되지 않거나 여러 개의 노드가 발견될 수도 있을 것이다. 여러 개의 관련 노드가 발견될 경우에는 실질적으로 트리를 변경시키는 것이 아니라 단순히 가상 링크를 이용해서 연결시켜 줌으로써 침입행위 로그패턴 분석과 모니터링에 사용 가능하게 해 준다. 이는 에이전트 시스템의 Log Tree Optimizer에 의해서 수행되며, 이러한 최적화 과정을 거친 최종적인 로그 트리들을 이용하여 통합 로그트리를 구축한다.

3.2 통합 로그트리를 이용한 로그 분석

Log Analyzer는 이와 같이 구성된 통합 Log Tree를 기반으로 서버에 전송할 로그 정보를 구축한다. 이때, 네트워크 보안 관리자가 용이하게 관찰할 수 있도록 계정, 시간, 접속호스트/포트, 프로세스 순으로 로그 정보를 구축한 다음, 설정되어 있는 로그패턴에 대한 정책에 따라서 이를 검색한다. 예를 들어, 특정 침입 패턴을 적용한 침입이 발생했다면 통합 로그트리 내에서 침입자에 해당하는 Log Tree뿐만 아니라, 그 Log Tree에 Hacking pattern을 적용해서 생성해 낸 Hacking Tree도 서버에게 전송해 준다. Hacking Tree는 통합 Log Tree가 가지고 있는 포맷과 유사한 형태를 가진다.

3.3 통합 로그트리를 이용한 통계 분석

통계 분석기(Statistic Analyzer)는 시스템 관리자가 서버를 통해서 요청하는 특정 항목에 대한 로그 정보를 생성해서 다시 서버에게 전달해주는 기능을 가지고 있다. 즉, 통계 분석기는 계정별, 시간별, 접속 호스트/포트별, 프로세스별로 로그 트리를 분석하여 관리자에게 전달해 준다.

여기서, 통계 분석기와 위의 로그 분석기의 차이점은, 로그 분석기는 침입이 발생했을 경우, 에이전트가 설정한 정책에 따라서 해당 침입 정보를 Log Tree에서 추출해내는 것이며, 통계 분석기는 네트워크 보안 관리자가 원하는 항목에 대한 로그 분석정보를 Log Tree에서 추출해내는 것이다.

4. 결론

본 논문에서는 전산망에 연결된 호스트 컴퓨터의 시스템 로그 파일을 각각의 로그트리로 만들고, 그 트리들을 최적화한 다음 생성한 통합 로그트리를 이용한 효율적인 서버-에이전트 기반의 침입 분석 시스템을 기술하였다. 이와 같이 함으로써 침입분석에 사용되는 각 로그 정보들의 유기적인 결합이 가능하여, 침입흔적 분석과 세밀한 분석 작업이 가능하였다.

하지만, 오늘날 전산망 및 컴퓨터 침입은 그 형태가 계속 변화하고 있으며 그 기법 또한 기술적으로 매우 진보되고 있다. 컴퓨터 바이러스와 백신과의 관계에서도 볼 수 있듯이 침입을 감지하고 조치하는 기술 역시 일단 침입사고가 발생한 이후에만 가능하다는 근본적인 문제점을 가지고 있다.

최근의 보안 시스템은 과거의 시스템과 같이 하나의 기능만을 가지고 있는 것이 아니라, 침입을 탐지(Detection)하고 차단(Protection)하며, 침입 피해를 분석하고, 침입자의 행동 및 침입패턴을 지능적으로 축적시키고, 아울러 취약점 및 피해를 복구하는 기능을 모두 가지고 있어야 한다. 본 논문에서 기술한 보안 시스템 역시 이와 같은 기능을 어느 정도 가지고 있기는 하지만 로그 분석에 대한 기법에 치중한 결과 다른 부분이 미흡한 상태이다. 향후에는 침입탐지와 침입 패턴을 분석하는 부분을 좀 더 중점적으로 보완 개발하고, 통합 로그트리를 보다 더 유연하고 확장성 있게 만들 수 있는 통합트리 최적화 알고리즘을 고안하는 것이 우선이라고 생각된다.

5. 참고문헌

[1] 한국정보보호센터, "전산망 보안관리 통합시스템 개발에 대한 연구", 1998,12,pp1

[2] 한국정보보호센터, "공개 보안 도구", 정보 보호 현황 1998,10,pp229-247

[3] 정보통신부, "정보 시스템 침해 사고 방지기술 개발에 관한 연구", 1999.1.

[4] Larry J. Hughes, Jr., Internet Security Techniques, New Riders Publishing, 1995