리눅스 분산 시스템 로그 검색 및 추적 시스템 설계

박준형° 송춘환 김민수 노봉남
전남대학교 멀티미디어 협동과정
{werther, chsdhs, phoenix, bongnam}@athena.chonnam.ac.kr

A Design of Audit Retrieval and Trail System For Distributed Linux

Jun-Hyoung Park° Choon-Hwan Song Min-Soo Kim Bong-Nam No
Dept. of Multimedia, Chonnam National University

요 약

최근의 침입이나 공격들은 광범위한 망의 이용과 긴 시간을 두고 공격을 행하는 추세로 발전하고 있다. 이러한 공격들에 대한 탐지 및 대응을 위하여 침입탐지 시스템들은 시스템에 걸쳐 있는 정보의 공유, 침입이나 공격에 적극 대응하기 위한 자원의 이용 등의 상호협력이 요구되고 있다.

본 논문에서는 시스템에서 기록하는 많은 종류의 로그 정보를 침입 탐지에 이용할 수 있는 정보만을 추출하고 분석하여 저장함으로써, 침입 탐지 시스템이 이용할 수 있게 하고, 다른 시스템에서 정보를 필요로 할 때 제공할 수 있으며, 또한 침입이라 간주되어지는 행위에 대하여 대응하기 위한 추적 및 정보 수집 그리고 접속 거부 등을 행하는 Agent 시스템의 설계 및 개발을 목적으로 한다.

이를 위해 리눅스 시스템 기반 하에서 로그를 기록하는 SYSLOG, 발생한 시스템 콜 정보를 기록하는 LSM, 망에서 시스템으로 들어오는 패킷을 분석하는 Pcap Library를 이용한 로그 등을 통합하는 과정을 설명하고, 침입탐지 시스템에 의해서 침입이라 판단되었을 경우, DART Agent가 그 경로를 역추적하고 여러 시스템에 걸쳐 있는 정보들을 수집하는 과정, 그리고 공격에 대한 대응을 하는 과정을 설명한다.

1. 서 론

침입 탐지와 대응을 위해서 로그의 수집과 분석은 매우 중요한 분야이다. 그러나 기록 목적이 침입 탐지를 위한 것이 아닌 경우, 탐지에 매우 유용한 정보가 있음에도 불구하고 사용되어지지 않아 왔다[1]. 또한 전통적인 침입탐지 시스템들은 시스템내부에서 발생하는 일에 대한 로그와 네트워크에서 발생하는 일에 대한 로그를 따로 분리하여 사용되어져 왔다. 그러나 그들 모두는 침입 탐지 목적으로 기록될 수 있고, 또한 하나의 사건에 대한 로그들로 연결 지어질 수 있다.

일반적으로 공격자가 특정 시스템의 컴퓨터에 침입을 시도하는 경우, 그 공격이 로컬 네트워크의 안쪽에서 시도되거나 바깥쪽에서 시도되는가 등으로 크게 두 가지로 나누어 볼 수 있다. 내부 망에서 시도되는 공격이라면 공격자의 위치가 어느 시스템인지의 여부는 공격에 대한 대응을 위해 매우 중요하다 할 수 있다. 또한 그 공격이 로컬 네트워크의 바깥쪽에서 시작되었다 할 지라도 네트워크의 어느 시스템이 공격의 기점으로 사용되었는지의 여부는 취약점 분석의 측면에서 매우 중요할 것이다. 이는 침입자가 먼저 공격의 목적이 되는 시스템의 네트워크를 조사하고, 보안이 취약한 시스템부터 강한 보안이 적용된 시스템의 순으로 공격을 시도하기 때문이다. 따라서 내부 네트워크에서 공격이 어느 경로로 진행되었는지를 추적하는 시스템은 반드시 필요하다.

이러한 문제는 시스템이 속한 내부망에서 보다 외부망에서 침투하였다면 좀더 어렵다. 분산환경에서 네트워크에 분산된 로그들을 사용하기에는, 시스템의 하드웨어적인 차이와 물리적 네트워크의 차이에서 비롯되는 문제들에 의해, 이진 형식의 로그 전송을 어렵게 한다. 따라서 공통적으로 이용될 수 있는 표준화된 로그 포맷[2]을 이용하여 정보를 전송하도록 하여야 한다.

위와 같은 문제들을 해결하는 방법으로 본 논문에서는 Linux에서 제공하는 대표적인 로그 기록 프로그램인 SYSLOG와, LSM, 그리고 패킷 분석 결과 등의 로그들을 통합함으로써, 여러 가지 로그가 하나의 사건에 대한 기록으로 분류되어질 수 있음을 보인다.

또한 이 논문에서는 시스템에 침입한 침입자가 어떠한 경로를 이용하여 침입하였는지, 그리고 그 침입자가 침입을 위해 경유한 시스템들에서 어떠한 행위를 하였는지, 그리고 공격을 받은 시스템에서 침입의 경로로 이용된 시스템들에게 이 사실을 알렸을 경우, 시스템들에서 가능한 대응들을 설계하고 구현하려 한다.

논문의 구성은 다음과 같다. 2장에서는 리눅스 로깅 시스템과 침입 탐지를 위해서 시스템간의 정보 교환을 방식들에 대한 연구를 소개하고, 3장에서는 설계된 시스템에서 사용되는 시스템 파일들과 통합하는 방법에 대하여 설명하고, 4장에서는 설계하고 있는 DART 시스템을 소개하고, 5장에서는 아직 결정되지 못한 문제점들을, 6장에서는 이 연구의 결론과 앞으로의 연구 가능분야들을 설명한다.

2. 관련 연구

- LSM(Linux Security Module)

리눅스 시스템의 커널 수준에서 시스템 호출 정보를 기록하고 관리하는 시스템 호출 로깅 모듈로서, 침입탐지 시스템 개발을 위해 기반기술을 위해 전남대학교에서

개발되었다.[8]

- CIDF(Common Intrusion Detection Framework)

침입 탐지 시스템들이 이용하는 정보와 자원들을 물리적으로 다른 시스템의 침입 탐지 시스템들의 요소들이 이용할 수 있도록 하는 연구를 하는 단체로서, 공통된 프로토콜과 응용수준의 프로그램들을 개발을 위한 목적으로 하고 있다.[1][2]

- AAFID(Autonomous Agents for Intrusion Detection)

미국의 COAST라는 단체에 의해서 연구되어지고 있는 분야로서, 하나의 시스템 안에서 다른 프로세스의 동작에 영향을 받지 않고, 독립적으로 보안을 감시하는 프로그램(Autonomous Agent)을 이용하여 네트워크상에서 논리적으로 계층적인 구조를 갖고 IDS들이 정보와 자원을 주고받도록 연구되고 있다.[3][4]

- Mobile Agents

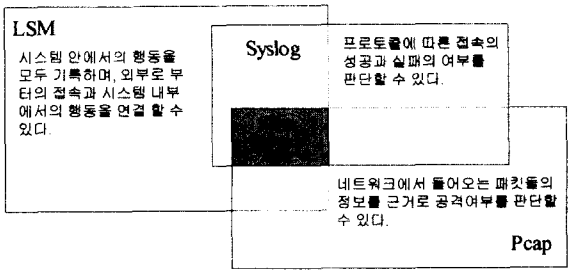
전통적인 침입탐지시스템의 client-server 관계에서 지속되어오던 문제를 극복하기 위하여 연구된 방식으로, 로깅이 이루어지는 시스템에서 연산을 한 후 결과를 전송하는 등의 침입 탐지 시스템이 있는 시스템의 부담을 줄여 더욱 확장된 분산 환경을 구성할 수 있다.[6]

3. 로그 파일의 이용

3.1 시스템 로그 파일

시스템 내부에서 수행되는 각종 로깅 프로그램들은 그 목적에 따라 다른 형식과 내용들을 기록한다. 그러나 이러한 정보들은 시스템의 무결성을 확인하거나 외부로부터의 침입이나 공격의 여부를 판단하는 중요한 근거가 될 수 있다. 이 논문에서 설계하는 DART 시스템은 발생한 시스템콜을 기록하는 LSM, 시스템에 접속하는 로그인 결과를 기록하는 SYSLOG, 그리고 Pcap Library를 이용한 네트워크의 패킷 정보 등을 이용한다.

3.2 DART 로그 파일로의 변환



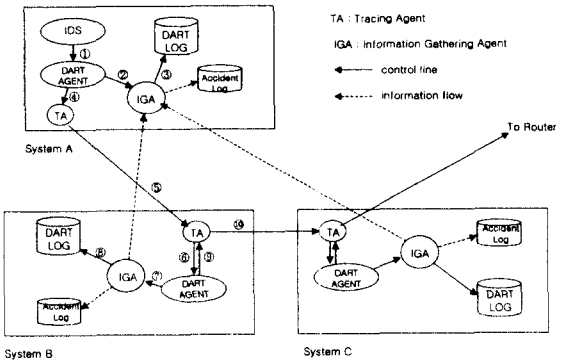
DART Log Generator는 시스템 로그 파일을 기반으로 침입 탐지에 이용될 수 있는 로그만을 골라내어, 이를 시스템에 접속한 시간, Source IP address, Port Number를 기준으로 하여 기록한다. 이 정보에는 침입자의 행위에 대한 정보와 접속한 위치 등의 정보를 저장하고 있다.

3.3 DART Message

침입탐지 시스템들이 서로 정보를 교환하기 위하여 사용된다. 이는 외부 시스템으로 정보를 전송하기 위하여, IGA가 수집한 정보를 축약 및 추상화하여 보다 전송하기 쉽고 다른 시스템에서 이해할 수 있는 형식으로 바꾸게 된다. 이러한 목적으로 네트워크상으로 전송되는 시스템 정보를 이 기종의 시스템에서의 이해할 수 있도록, CIDF에서 제안한 S-expression을 이용하여 정보를 전송한다.

4. DART 시스템 설계

4.1 기능



DART Agent : IDS가 침입/공격으로 의심되는 행위를 판단한 후, 그에 대한 대응을 담당하는 Agent로서 Manager Agent에게 자신의 정상 동작 여부를 전송한다.

Tracing Agent : 침입/공격자의 위치를 추적[7]하는 Agent

Information Gathering Agent : 침입/공격으로 의심되는 행위에 대한 정보를 추출하며, TA가 추적하는 경로 정보를 제공하는 Agent

Manager Agent : 내부 네트워크에서 동작하는 DART Agent의 동작을 관리하고, 외부 네트워크의 Manager Agent에게 정보를 요청한다.

DART Log : 침입 탐지 시스템이 이용할 수 있는 모든 정보의 집합으로서 IDS 뿐 아니라, DART Agent의 활동에 필요한 정보를 보관한다.

DART Log Generator : 시스템 로그를 침입 탐지에 관련된 정보만을 골라내어 DART Log에 기록하는 Agent

Accident Log : IDS에 의해 대응 지시와 관련된 정보를 DART Agent에 의한 기록한다.

4.2 DART Log

침입 탐지에 이용될 수 있는 모든 로그를 기록한다. 구현 중인 시스템은 아래와 같은 정보를 기록한다.

① LSM : 발생하는 시스템 콜 정보 중 accept, connect, fork, exec, login, logout 시스템 콜의 정보,
- 접속하여 들어온 정보, 시스템 내부에서 활동의 내역, 외부로 접속한 정보등을 연결시켜 주는 고리 역할을 한다.

② Syslog : 각 접속 요청에 대한 Time, Daemon, PID, User name, Source IP, Port,
- 로그 인 성공/실패 여부, 실패 이유 등을 기록.

③ Pcap : 내부 시스템에 들어오는 모든 패킷들의 Source IP, Port, Destination IP, Port, Service Type, Sequence number, Acknowledgement

number, Code Bits, Time, Device, Source MAC address, Destination MAC address, Ether type

이들 정보는 시스템에 접속을 시도한 원격지의 IP 주소와 Port 정보, 그리고 시간을 이용하여 하나의 사건과 관련된 기록으로 구분되어질 수 있다.

4.3 DART 시스템 동작 설명

① IDS는 침입으로 간주되어지는 Process 활동을 발견하고, DART Agent를 동작시킨다.

② DART Agent는 침입으로 보이는 Process와 관련된 정보를 수집하도록 Information Gathering Agent(IGA)에게 요청한다.

③ IGA는 주어진 PID와 관련된 모든 정보(시스템 내부에서의 행동, 시스템에 접속한 원격지의 위치, 외부 시스템으로 시도한 접속 등)를 수집한다.

④ DART Agent는 IGA에 의해 수집된 정보를 기반으로 시스템으로 접속한 원격지의 위치를 알아내고 Tracing Agent(TA)에게 추적을 지시한다.

⑤ 추적된 시스템의 TA는 자신의 시스템의 DART Agent로 하여금 침입으로 간주되어지는 접속 정보를 제공하고, 관련 정보를 수집하도록 요청한다.

⑥ 위의 ②번부터 반복...

⑦ 더 이상 내부 망에서 추적할 수 없다면 침입에 이용되는 원격지에서의 내부 망으로 접근하는 접속을 종료시키도록, 내부 네트워크를 관리하는 Manager Agent에게 정보를 요청한다.

4.4 문제점과 해결 방안

- 데이터 전송 : DART Agent는 AAFID의 계층적 구조를 따르고 있다. 이는 탐지하는 시스템에 따른 로그 전송이 필요하다는 사실이 가장 큰 이유일 것이다. 그러나 이는 전통적인 Master, Slave 관계에서 그러하듯이 여러 시스템에서 전송되는 메시지를 받아 탐지하는 시스템은 네트워크 대역폭의 한계와 병목현상을 일으킬 수 있는 지점이 될 수 있다는 문제를 초래할 수 있다. 그러나 이러한 문제는 DART Log Generator에 의해 로그를 축약하고 IGA에 의해 전송될 때 로그 자체가 아닌 분석 결과를 전송함으로써 극복될 수 있다.

- 로그 전송 포맷 : 모든 시스템들이 이해할 수 있는 형태의 전송이라 함은 매우 중요하지만 어려운 문제이다. 또한 많이 연구되어진 분야라고도 할 수 있다. 본 논문의 구현에서는 새로운 포맷을 제안하기 보다, CIDF에서 제안하고 있는 S-expression 형식의 정보 전송 형태를 이용하도록 한다. 또한 문서 표준으로 자리잡고 있는 XML 형식의 로그 표현 형태도 연구하고 있다.

5. 결론 및 향후 연구 방향

이 논문에서 설계하고, 현재 개발 중인 DART 시스템 구현의 기본 방향은 세 가지로 나누어 질 수 있다. 침입 탐지에 이용될 수 있는 로그 수집, 이상 행위 발견 시 이루어지는 증거수집, 그리고 더 이상의 공격이 이루어지지 않도록 대응하는 부분이다.

먼저 현재 시스템 내부에서 기록되고 있는 모든 로그들을 침입 탐지에 이용될 수 있는 정보를 골라내고, 이들을 하나의 사건에 대한 로그들로 분류함으로써 로그를 더욱 효과적으로 이용할 수 있다. 그러나 침입 탐지 시스템에서 더욱 더 효과적인 탐지를 할 수 있도록 하는 정보의 분류와 축약은 앞으로도 연구되어져야 한다.

또한 침입이나 공격으로 판단된 행위에 대하여 가해질 대응을 위한, 침투 경로의 추적, 침투 방법을 찾기 위한 정보 수집, 그리고 내부 망으로의 접근 통제는 적극적인 대응을 위하여 다른 시스템을 이용할 수 있도록 하는 시스템이다.

앞으로의 연구방향은 침입 탐지 시스템이 더욱 더 효과적으로 로그를 이용할 수 있도록 하는 로그 분석과 분류 방법이며, 인터넷을 경유하여 이루어지는 침입이나 공격에 대해 적극적인 대응을 할 수 있도록 하는 연구가 진행되어져야 한다.

6. 참고 문헌

[1] M. Bishop, "A Standard Audit Trail Format." In Proceeding of the 18th National Information Systems Security Conference, Baltimore, pages 136-145, 1995.

[2] Clifford Kahn, Phillip A. Porras, Stuart Staniford-Chen, Brian Tung, "A Common Intrusion Detection Framework."1998.

[3] Eugene H. Spafford, Diego Zamboni, "Intrusion detection using autonomous agents." 2000.

[4] Jai Sundar Balasubramaniyan, Jose Omar Garcia Fernandez, David Isacoff, Eugene Spafford, and Diego Zamboni. "An architecture for intrusion detection using autonomous agents." In Proceedings of the Fourteenth Annual Computer Security Applications Conference. IEEE Computer Society, December 1998.

[5] Thomas H. Ptacek, Timothy N. Newsham, "Insertion, Evasion, and Denial of Service: Eluding Network Intrusion Detection." January, 1998.

[6] Midori Asaka, Shunji Okazawa, Atusushi Taguchi, Shigeki Goto, "A Method of Tracing Intruders by Use of Mobile Agents." INET'99, June 1999.

[7] Choonhwan H. Song, "A Design of Distributed Audit Trail System For Solaris and Linux Systems." 2001.

[8] 박남열, 송춘환, 김정일, 노봉남, "호스트 기반 침입 탐지 시스템을 위한 리눅스 보안모듈." 제 11회 통신정보 합동학술대회 논문집, pp81-84, 4. 2001.