2001년 한국정보처리학회 추계 학술발표논문집 제8권 제2호

리눅스 시스템에서의 분산 로그 검색 및 추적

박준형*, 송춘환**, 김민수**, 노봉남**
*전남대학교 멀티미디어협동과정
**전남대학교 전산학과
e-mail:werther@athena.chonnam.ac.kr

Distributed Audit Retrieval and Trail System For Linux

Jun-Hyoung Park*, Choon-Hwan Song**, Min-Soo Kim**
Bong-Nam No**
*Dept of Multimedia, Chon-nam University
**Dept of Computer Science, Chon-nam University

요 약

시스템들은 시스템에서 발생한 일들에 대한 기록을 로그의 형식으로 남긴다. 이러한 로그들은 여러 가지 목적으로 작성되어지며 이용되어진다. 그러나 로그 기록은 침입탐지뿐 아니라 여러 가지 목적을 가지고 이루어진다. 그러한 로그들 역시 침입 탐지에 사용할 수 있는 많은 정보가 있음에도 불구하고 사용되지 않고 버려지고 있음이 현실이다. 본 논문에서는 침입 탐지를 위하여 이러한 로그들을 적극적으로 이용함을 목적으로 한다.

또한 최근의 공격형태의 변화로 인해 하나의 시스템관점이 아닌 네트워크 관점에서의 탐지와 대응이 필요하게 되었다. 이를 위해 공격에 이용된 시스템들을 파악하고 그들의 정보를 이용할 수 있게 함으로써 상호 협력이 가능케 한다.

1. 서론

시스템에 접근한 사용자의 행동은 특정 목적을 가진 로그파일에 기록되고, 그 기록은 차후 여러 목적으로 분석되어진다. 그러나 목적은 서로 달리한다 하더라도 그 로그에 기록된 내용은 시스템에서 일어난 사건들의 기록이므로, 침입 탐지 및 대응의 목적으로 이용될 수 있는 중요한 내용일 수 있다[1]. 또한 전통적인 침입 탐지 시스템들은 시스템 내부에서 발생하는 일에 대한 로그와 네트워크에서 발생하는 일에 대한 로그를 따로 분리하여 사용하여 왔다. 이러한 로그 역시 침입 탐지의 목적으로 기록될 수 있고, 하나의 사건에 대한 로그들로 연결되어질 수 있다.

일반적으로 공격자가 특정 시스템에 공격자가 특정 시스템에 공격 및 침입을 시도하는 경우, 그 공격은 내부 네트워크에서 시도되었거나 외부 네트워크에서 시도되었는가 등으로 크게 두 가지로 나누어 볼 수 있다. 내부 네트워크에서 시도되는 공격이라면 공격자의 위치가 어느 시스템인지의 여부는 공격에 대한 대응을 위해 매우 중요하다. 또한 그 공격이 외부 네트워크에서 시작되었다 할지라도 네트워크의 어느 시스템 공격의 기점으로 사용되었는지 여부는 취약점 분석의 측면에서 매우 중요하다 이는 침입자가 공격을 시도할 때, 보안이 취약한 시스템부터 강한 보안이 적용된 시스템의 순서로 침입을 시도하기 때문이다. 따라서 내부 네트워크에서 공격이 어느 경로로 진행되었는지를 추적하는 시스템의 연구가 필요하다[5]. 또한 이러한 경로를 추적하는 시스템은 내부네트워크에서 발생한 침입자의 모든 행위에 대한 정보를 수집할 수 있도록 하여 효과적인 분석과 대응을 가능하게 한다.

위와 같은 문제들을 해결하기 위해, DART 시스템은 Linux에서 제공하는 로그 기록 프로그램인 SYSLOG와 리눅스 시스템에서 발생한 시스템 콜을 기록하는 LSM, 그리고 패킷 분석 결과 등의 로그를 통합, 분석하여 기록할 수 있음을 보인다.

또한 시스템에 침입한 침입자가 어떠한 경로를 통하여 침입하였는지 추적하고, 경유한 시스템에서 어떠한 행위가 이루어졌는지 정보를 수집하도록 한다.

논문의 구성은 다음과 같다. 2장에서는 리눅스 로깅 시스템과 침입 탐지를 위한 시스템간의 정보 교환 방식들에 대한 연구를 소개하고, 3장에서는 DART 시스템에서 사용하는 로그 파일들과 통합에 대하여 설명하고, 4장에서는 DART 시스템의 동작 과정과 문제점을 소개하고, 5장에서는 이 연구의 결론과 앞으로의 연구 필요분야들을 설명한다.

2. 관련 연구

- LSM(Linux Security Module)

리눅스 시스템의 커널 수준에서 시스템 호출 정보를 기록하고 관리하는 시스템 호출 로깅 모듈로서, 침입 탐지 시스템 개발을 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 응용 수준에서의 프로그램 개발을 목적으로 연구하고 있다[1][2].

- AAFID(Autonomous Agents for Intrusion Detection)

미국의 COAST라는 단체에서 연구하고 있는 구조로서, 시스템에서 다른 프로세스의 동작에 영향을 받지 않고, 독립적으로 보안을 감시하는 프로그램(Autonomous Agent)을 이용하여 네트워크 상에서 논리적으로 계층적인 구조를 갖고 IDS들이 정보와 자원을 공유할 수 있도록 하는 시스템이다[3][4].

- Mobile Agents

전통적인 침입 탐지 시스템들이 구성하는 client-server 관계에서 오는 문제점들을 극복하기 위하여 연구된 방식으로, 침입이 발생한 시스템에서 직접 보고할 수 있고, 대응이 즉시 이루어지게 한다. 일본의 IPA에서 연구되어지고 있다[6].

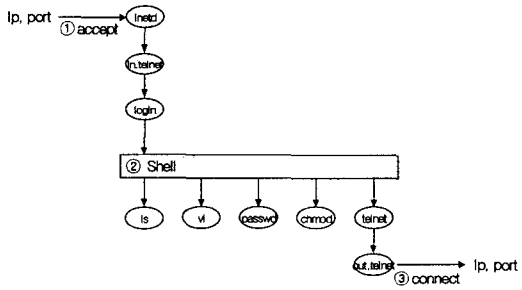
3. 시스템 로그

시스템에 접근한 사용자의 행동은 동작하고 있는 프로세스의 상태를 검사함으로써 알 수 있다. 따라서 이러한 프로세스들이 발생시킨 시스템 콜들의 내역을 기록한 로그들은 시스템이 수행한 모든 일을 파악할 수 있게 해주는 중요한 기록이다.

외부에서 접속한 사용자가 어느 시스템에서 접근해 들어왔으며, 또한 어느 시스템으로 접속해 나갔는지 역시 프로세스들의 상태 변화를 검사함으로써 어렵지 않게 추적할 수 있다. 이를 위해 DART 시스템은 LSM 로그를 기본으로 하고, SYSLOG 기록 내용, 네트워크 패킷 분석 로그 등을 통합하도록 한다.

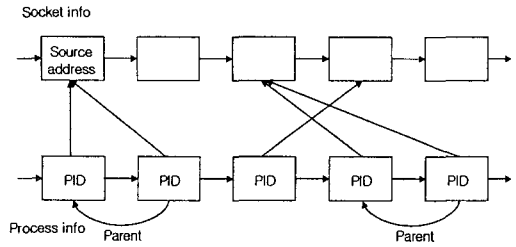
3.1 LSM

접속은 프로세스의 accept 시스템 콜의 발생이후 이루어지게 되고, 또한 외부로 나가는 접속은 프로세스의 connect 요청에 의해서 이루지게 된다. 또한 접속자의 모든 작업은 accept한 프로세스에 의해 fork된 프로세스들에 의하여 작업을 하게 된다. 따라서 시스템에 접근한 사용자가 다른 시스템으로 접근을 시도할 때 발생하는 connect 요청을 하는 프로세스 역시 처음 연결을 맺은 프로세스에 의하여 fork된 프로세스이다. 이러한 사실은 네트워크에서 시스템을 경유하여 공격과 침입을 시도하는 사용자의 경로를 추적하는데 매우 중요한 근거가 된다.



사용자가 시스템에 접속하여 다른 시스템으로 접속하는 동안의 프로세스 변화는 위의 그림과 같다. 따라서 DART 시스템에서는 사용자 위치의 역추적을 위해 아래와 같이 기억장소에 연결정보와 프로세스의 관계를 저장하도록 한다. 이는 실시간 역추적이나 로그를 기반으로 한 사후 역추적 모두에서 사용될 수 있다.

① 시스템에 사용자가 접속을 요청하면 accept 시스템 콜이 발생하게 되고, 이때 접속한 시스템의 source ip, port, destination ip, port 등을 알 수 있다. 이 정보들을 Socket_info 구조체 리스트에 저장하고, Process_info 구조체 리스트에 해당 프로세스 번호를 갖는 기억장소를 추가시킨다. 이때 각각의 process 구조체는 접속정보를 저장하고 있는 socket_info 구조체의 주소를 기억하도록 한다.



외부 접근에 의한 Process의 주소 기억

② 프로세스 구조체 리스트에 있는 Process들은 fork를 할 때마다 그에 대한 자식 프로세스를 구조체 리스트에 추가하고 연결 정보를 상속시킨다.

③ 프로세스 구조체 리스트에 있는 PID중에서 connect 시스템 콜이 발생했을 때의 소켓 정보는 socket_info 구조체에 기록되어있는 소켓 정보의 연결에 대한 기록은 DART Agent에 기록되며, TA의 사용자 위치 추적에 사용된다.

3.2 이 기종의 로그와의 결합.

DART Log Generator는 시스템 로그 파일 중 침입 탐지에 이용될 수 있는 로그만을 골라내어, 시스템에 접속한 시간, Source IP, Port를 기준으로 기록한다. 이러한 정보는 사용자의 행위에 대하여 이 기종의 로그들을 연결하는 중요한 기준이 된다.

구현중인 DART 시스템에서는 Linux 시스템에서 발생한 시스템 콜 정보를 기록하는 LSM, 시스템에 시도되는 접근정보를 기록하는 SYSLOG, 그리고 Pcap Library를 이용하여 네트워크 패킷 정보를 기록하는 TCP Dump 등의 기록을 사용한다.

4. DART 시스템

4.1 목적

- 이상행위 발견 시, 접속 시스템 추적
- 공격 및 침입시 경유한 시스템에서 이루어진 사전 작업 및 취약성 분석
- 공격 및 침입에 대한 실시간 대응

4.2 정의

- DART Agent

IDS에서의 비정상행위라는 판단에 따른 요청이나, 다른 시스템의 TA의 추적 요청에 응답할 수 있다. Agent의 동작을 관리하는 Agent이다. 또한 다른 시스템에 설치된 Agent들에게 자신의 동작 여부를 알림으로써, 동적으로 시스템의 영역을 확장할 수 있도록 한다.

- Tracing Agent(TA)[7]

침입탐지 시스템이 추적하기를 지시한 process에 대한 외부로부터의 연결을 찾아내고, 연결을 요청한 시스템의 TA에게 다음 추적을 지시하며 해당 DART Agent에게 IGA로 하여금 해당 접속에 대한 정보를 수집, 분석하도록 한다.

- Information Gathering Agent(IGA)

TA의 경로 추적 정보와 추적한 연결에 대한 시스템에서의 활동 내역에 대한 로그를 탐지 및 분석하여 Accident Log에 저장한다. 또한 이 정보를 요청한 시스템으로의 전송도 담당한다.

- DART Log

DART Log Generator에 의해서 제공된 로그들의 모음이다. 침입탐지 시스템에 의하여 사용되며, TA의 역추적과 IGA의 정보 수집을 위해 이용된다.

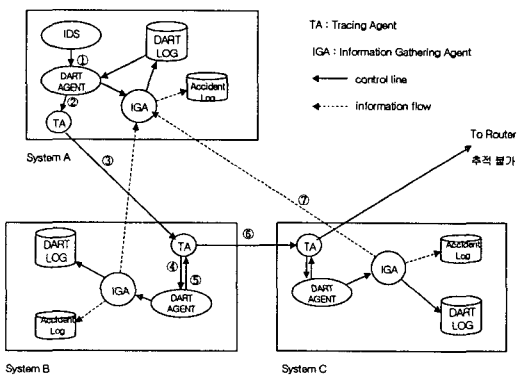
- DART Log Generator

이 기종의 시스템 로그들 중, 침입 탐지에 이용될 수 있는 모든 정보를 찾아내고, 그 내용을 분석하여 하나의 사건에 대한 기록들로 연결되어질 수 있도록 기록하는 프로그램. 그러나 단순히 시스템 로그들의 복사가 아닌 그 정보들을 분석하여 기록한다.

- Accident Log

TA들의 경로 추적 결과, IGA에 의해 수집된 정보의 저장 및 전송 내용 등을 저장하여 차후 침입에 대한 증거로서 이용될 수 있다.

4.3 DART 시스템 동작 설명



① IDS는 침입으로 간주되어진 Process 활동을 발견하고, DART Agent에 추적을 요청한다.

② DART Agent는 요청 받은 Process와 관련된 accept 요청을 찾아내고 TA에게 추적을 명령한다.

③ TA는 추적할 시스템의 ip 주소를 받아 해당 시스템 TA에게 추적을 요청한다.

④ 추적된 시스템의 TA는 자신의 DART Agent로 하여금 추적할 접속이 어느 시스템에서 연결되었는

지 정보를 요청한다.

⑤ IGA는 추적할 접속에 대한 프로세스관계를 조사하여 증거를 수집하고, 시스템에 connect한 시스템을 찾아내어 TA에게 정보를 제공한다.

⑥ ③번부터 계속...

⑦ TA는 추적할 연결이 이루어진 시스템에 DART Agent가 설치되지 않았거나, 어떠한 이유로 인해 추적을 계속할 수 없다면, 처음 추적을 요청한 시스템에게 정보를 전송한다.

4.4 문제점과 해결 방안

- Manager Agent : DART Agent의 기본 목적은 하나의 시스템 관점에서는 볼 수 없는 네트워크 관점에서의 판단 및 대응을 가능케 한다. 이를 위하여 AAFID의 계층적 구조를 따르고 있다. 그러나 이는 전통적인 Master, Slave 관계에서 오는 네트워크 대역폭의 문제점과 시스템에서 처리해야 하는 자료의 폭주로 인한 병목현상을 초래할 수 있다. 그러나 분석에서 오는 시스템의 부하를 분산함으로써 극복할 수 있다.

- 로그 전송 포맷 : DART Agent는 시스템과 네트워크에 제약받지 않고 모든 시스템에서 이해할 수 있는 형태의 전송을 목적으로 한다. 이는 CIDF를 비롯한 여러 단체에서 연구하고 있는 분야이다. 따라서 CIDF에서 제안하고 있는 S-expression을 이용한 정보 전송형태를 이용하도록 한다. 또한 문서 표준으로 자리잡고 있는 XML 표현 형태도 연구 중에 있다.

5. 결론 및 향후 연구 방향

현재 구현중인 DART 시스템이 갖는 기능은 침입 탐지에 이용될 수 있는 로그의 수집, 이상행위 발견 시 이루어지는 침입자 위치 추적 및 증거 수집 부분이다.

먼저 시스템에서 기록하는 모든 로그에서 침입 탐지에 이용할 수 있는 정보를 골라내고, 이들을 탐지 시스템이 이용할 수 있도록 제공하여 로그 정보를 더욱 효과적으로 이용할 수 있다. 그러나 침입 탐지 시스템에서 더욱 효과적인 판단을 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 통제 등은 적극적인 방어를 위하여 다른 시스템을 이용할 수 있도록 하는 시스템이다.

앞으로의 연구 방향은 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. 참고문헌

[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 침입탐지 시스템을 위한 리눅스 보안모듈." 제11회 통신정보 합동학술대회 논문집, pp81-84, 4. 2001.