기능과 품질의 관점에서의 침입탐지시스템의 평가를 위한 실험 데이터 패턴 생성

조범래[0], 김종, 홍원기
포항공대 정보통신대학원 정보통신학과
{ brcho, jkim, jwkhong }@postech.ac.kr

Test Data Pattern Generation To Evaluate Intrusion Detection Systems In Aspect Of Functionality And Quality

Bum-Rae Cho[0] , Jong Kim, Won-Ki Hong
Graduate School of Information Technology, Pohang University of Science and Technology

요 약

컴퓨터의 역사와 함께 하여왔던 보안 문제를 해결하기 위하여 여러 연구소 및 기업에서 침입을 능동적으로 탐지하기 위한 침입탐지시스템의 개발을 수행하여 왔다. 기존 보안 평가가 보안 시스템의 안정성 인증에 국한되는 평가에 국한되는 한계를 극복할 수 있는 방법들을 UC Davis와 MIT에서 발표하였지만 그들의 평가가 공격에 대한 식별을 근간으로 하고 있으므로 침입 관련 정보의 갱신 수준에 대한 평가에 그치고 말았다. 본 논문에서는 기존의 보안 시스템 평가의 문제점을 해결할 수 있는 기능과 품질의 관점에서의 침입탐지시스템의 평가를 위한 실험 데이터의 패턴 생성 방법을 말하고, 공개된 침입탐지시스템인 snort를 대상으로 기능과 품질 관점의 평가에 대한 구체적인 적용 사례를 소개한다.

I. 서론

인터넷의 범용적인 사용으로 인하여 각종 기관들의 컴퓨터 시스템들이 네트워크로 연결되고 있다. 네트워크의 강력한 연결성으로 인하여 각종 정보 업무를 수행하는 일이 쉬워지고 그 정보의 교류 속도가 매우 빨라져 업무의 효율성에 중대한 장점을 가지게 된 것이 사실이다. 그러나 이에 상반되는 결점인 보안문제의 심각성은 날로 그 수위를 더하고 있다. 방화벽(Firewall)[1]과 같은 수동적인 보안 장비의 한계를 극복하기 위하여 침입탐지시스템(Intrusion Detection System)[1][2]이 개발되었으나 평가를 위한 보안 지침서들은 시스템 자체의 안정성에 대한 인증 항목들만을 제시하고 있다[3]. 침입탐지시스템은 능동적으로 침입을 탐지하고 이에 대응하는 기능들을 갖고 있다. 그러나 침입탐지시스템의 침입탐지 능력에 대한 평가 기준이 없으므로 연구기관, 기업 그리고 소비자들은 침입탐지시스템의 개발과 도입에 있어서 객관적인 기준을 갖지 못하게 된다.

국내에서는 이미 2000년 7월에 정보통신망 침입탐지시스템 평가기준[5]을 발표하였으나 이 또한 안정성 인증 평가를 벗어나지 않고 있다[3]. 이

는 침입탐지시스템의 침입탐지 능력 혹은 기능을 평가하는 데에는 적절하지 않은 항목들로 구성되어 있다. 침입탐지시스템의 기능을 평가하고자 UC Davis[6]와 MIT[7]가 처음으로 방법론을 제시하였다. UC Davis에서는 침입식별, 자원사용량, 스트레스의 세 항목에 걸쳐 실험하였고 MIT는 최신 침입에 대한 식별에 비중을 두고 평가를 수행하였다. 그러나 이들 실험들의 공통점은 침입탐지시스템이 갖고 있는 분석을 위한 침입에 대한 최신 정보의 보유 수준에 대한 평가라는 것이다. MIT에서는 침입 정보량과 갱신 수준에 대한 평가가 갖는 평가의 불공정성을 해소하고자 실험에 사용될 침입탐지시스템을 미리 제공한 데이터로 훈련을 수행하도록 하였다. 또한 일정 시간 동안 제공된 데이터로 시스템을 훈련한 것도 각 시스템의 분석 엔진의 특성에 따라 훈련 효과에 대한 신뢰성을 떨어지게 한다.

본 논문은 2장에서 관련 연구를 통하여 기존의 보안 시스템 평가 요소들을 통하여 국제 인증 기준들의 기능과 품질을 고려한 인증 항목들과 UC Davis와 MIT의 평가 시도의 실험 방법에 대해 소개한다. 3장에서는 기존의 평가가 갖는 문제점들을 해결하기 위한 실험 데이터 패턴의 요구사항을, 4장에서는 실제적인 평가 데이터의 패턴의 생성 방안에 관해 논하고 5장에서는 그 구현을, 6장에서는 snort[11]에 대한 적용에 대해 기술한다.

II. 관련연구

1. 기존의 보안시스템 평가기준

세계 각 국들은 보안 문제의 심각성이 더해지면서 이를 해결하고자 컴퓨터 시스템이 갖추어야 보안 평가 요소들을 제정하였다. 아래의 기준들을 통해서 모든 평가 항목들이 보안 시스템의 안정성 인증을 위한 항목들임을 알 수 있다. 그러나 UC Davis와 MIT의 실험에서 구체적으로 논의되고 고려되지 않은 측면인 기능과 품질에 관한 보안 수준별 평가 항목들은 앞으로 추가적인 국제적인 침입탐지시스템의 평가 기준 또한 이를 고려하여 제정될 것을 시사해 주고 있다[3][4][5].

1.1. 국내의 침입탐지시스템 평가기준[3][4]

한국 정부는 2000년 7월에 정보통신망 침입탐지시스템 평가기준을 발표하여 침입탐지시스템의 안정성 인증을 위한 발판을 마련하였다. 이 기준은 K1~K7 의 등급을 가지며 각 등급별 보안기능요구사항과 보증요구사항을 가지고 있다.

표 1. 보안기능과 보안인증 요구사항

요구사항	내 용
보안기능	축약감사데이터 생성, 보안위반분석, 보안감사대응, 신분확인, 데이터 보호, 보안감사 , 보안관리, 보안기능의 보호
보안인증	개발과정, 시험과정, 형상관리, 운영환경, 설명서, 취약성

1.2. 미국과 유럽의 보안 시스템 평가 기준

1.2.1. 미국의 TCSEC[9]

1983년에 미국은 국방부와 NBS(National Bureau of Standards) 등을 중심으로 안전한 컴퓨터 시스템의 구축 및 평가등에 지속적인 연구 결과인 TCSEC(Trusted Computer System Evaluation Criteria)을 발표했다. A,B,C,D의 4개의 기본적인 영역을 가지며 어떤 요구사항도 없는 D를 제외하고 C1, C2, B1, B2, B3, A1의 6개 영역으로 나뉘어 있다.

표 2. TCSEC의 보안 요구사항

	내 용
보안요구사항	보안정책, 표시, 신분확인, 감사기록, 보증, 지속적인 보호

1.2.2. 유럽의 ITSEC[10]

ITSEC(Information Technology Security Evaluation

Criteria)은 1990년에 영국, 독일, 프랑스, 네덜란드가 협력하여 출간한 유럽 공통 평가 기준이다. ITSEC은 보안기능 요구사항과 보증 요구사항을 가지는데 E1~E6의 6등급으로 나누어지며 E0 등급은 부적합 판정을 의미한다.

표 3. ITSEC의 보안 요구사항

	내 용
개발과정	요구사항, 구조설계, 상세설계, 구현
개발 환경	형상관리, 프로그래밍 언어 및 컴파일러, 개발 보안과 운영환경에 관한 문서에 대한 평가

1.3. 미국과 독일의 평가 기준 비교

독일의 그린 북을 작성한 GISA(German Information Security Agency)는 8개 항목에 걸쳐 품질 평가 항목을 제시하고 10개의 기능 요구사항을 작성하여 계산적으로 볼 때 80가지의 다른 평가 결과를 도출할 수 있도록 하였다. 이는 미국의 TCSEC의 평가 클래스들과도 대응되며 이는 국제적인 평가의 흐름이 기능과 품질을 모두 고려하는 것이 바람직함을 말해주는 단적인 예이다[5].

2. 침입탐지시스템의 침입탐지 성능 평가

2.1. UC Davis

이 실험에서 사용된 평가항목은 표4와 같다. 식별(Intrusion Identification)은 잘 알려진 공격을 선정하여 식별하는지의 여부에 관한 것이었고, 자원 사용량(Resource Usage)은 생성된 가상의 사용자들의 수가 증가함에 따른 디스크 공간의 소모율을 조사한 것이다. 그리고 스트레스 실험(Stress Test)은 침입탐지를 수행하는 네트워크에 침입탐지를 방해할 수 있는 트래픽의 양을 증가시키면서 침입에 대한 식별의 유무를 실험한 것이다.

표4. UC Davis의 침입탐지시스템 평가 영역

	내 용
평가항목	식별, 자원 사용량, 스트레스 실험

2.2. MIT

MIT의 실험은 UC Davis에 비해서 더욱 공정한 실험을 수행하기 위하여 미리 훈련 데이터를 제공하여 2주간의 시스템 훈련 기간을 주었다. 그리고 훈련 데이터와 제공하지 않은 새로운 침입들을 개발하여 침입탐지시스템을 평가하였다.

표 5. MIT의 침입탐지시스템 평가 영역

	내 용
평가항목	식별(Intrusion Identification)

2.3. UC Davis와 MIT의 실험의 한계

UC Davis와 MIT의 실험은 공통적으로 침입 식별을 기반으로 한 평가였다.

침입 식별을 통한 평가는 평가 대상 침입탐지시스템의 침입관련 데이터의 갱신 수준에 따라 그 결과가 달라질 수 있기 때문에 불공정하다[8][9]. 그리고 침입탐지시스템을 훈련시키고 다시 같은 데이터와 유사한 침입 데이터에 대한 탐지를 수행하는 것도 여전히 침입탐지시스템의 침입탐지를 위한 고유 기능에 대한 평가가 아닌 침입 데이터의 보유 수준에 대한 평가에 그친다[8].

III. 침입탐지시스템의 기능과 품질을 평가하기 위한 데이터 패턴의 요구사항

1. 기능과 품질을 고려한 데이터 패턴의 생성

이 평가 항목의 분류에서는 국제 평가 기준들의 고려 관점인 기능과 품질을 평가하기 위한 데이터 패턴의 생성을 목적으로 한다.

2. 침입탐지시스템의 기능 평가

침입탐지의 보안계층은 표6과 같이 3개의계층으로 나누어 고려한다[1]. 각 계층에서 침입탐지시스템이 가져야 할 기능 요소들에 대한 데이터 패턴들은 [2]에서 소개된 침입탐지시스템의 다양한 침입탐지 방법들의 특성들을 침입탐지의 관점이 아닌 공

격과 침입의 관점을 적용하여 표 6과 같이 분류된다.

표 6 침입탐지시스템의 기능 분류

보안계층	데 이 터 패 턴
응용계층	응용의 exploit을 이용한 침입
운영체제 계 층	식별과 인증을 침해하는 행위
	접근제어를 침해하는 행위
	시스템 자원의 비정상적인 소모
	정상 명령어에 의한 침입
네트워크 계 층	단일 패킷에 의한 침입
	다중 패킷에 의한 침입
	조각화된 패킷에 의한 침입
	보안상 중요한 파일의 전송
	Promiscuous 모드 프로그램

3. 침입탐지시스템의 품질 평가

침입탐지시스템의 품질을 평가하기 위해서 먼저 품질 요소들을 표7에서 기술한다. 표7의 항목들은 현존하는 침입탐지시스템들의 특성에서 발췌한 것이다.

표7. 침입탐지시스템의 품질 요소

	내 용
품질 요소	분석시간, 감사기능을 이용한 탐지결과 조정 및 로그화 , 경고, 경보, 침입탐지시 대응, 로그 검색 및 재현, 다양한 통계 자료 제공, 데이터에 대한 암호화, 탐지 공격의 종류, 네트워크 관리, 분석 서비스의 종류, 트러블 슈팅, 자원의 사용량, 침입 데이터 갱신 , 분석 엔진 버그 패치

표7의 14가지의 요소들 가운데 질의 측면과 양적 측면의 요소들을 살펴보면 다음의 표 8와 같다. 품질을 고려함에 있어서 항상 수량과 질적인 항목의 분류가 검토되어야 하므로 표8과 같이 분류하였다.

표8. 침입탐지의 품질 평가 요소

	수량적 관점	질적 관점
침 입 탐 지	• 시스템 엔진의 분석 시간 •침입에 사용된 서비스의 종류 • 침입 프로그램의 명칭	• 침입 데이터 갱신 • 분석엔진 버그 패치
침입탐지 시스템	자원의 사용량	표7의 모든 항목들

표8의 2가지 관점에 속하는 평가 항목들 중에서 실제적인 평가를 위한 실험 데이터 패턴 생성을 위한 항목은 시스템 자체의 특성을 평가하기 위한 침입탐지시스템의 평가를 위한 내용을 배제한 침입탐지의 평가를 위한 내용들이다. 그리고 실질적인 데이터 생성과 관계 없는 침입데이터 갱신과 분석 엔진 버그 패치 요소들을 제외하면 표9와 같은 2가지의 실험 데이터 패턴을 얻게 된다.

표9. 품질 평가를 위한 실험 데이터 패턴

	데 이 터 패 턴
침입탐지	컴퓨터 서비스를 이용한 침입, 공개된 잘 알려진 공격에 의한 침입

IV. 침입탐지시스템의 평가를 위한 실험데이터 패턴 생성 방안

1. 기능 관점의 평가

표10. 기능을 평가하기 위한 데이터 패턴 생성 방법

기능적 요소	데이터 패턴 생성 방안
응용의 exploit을 이용한 침입	More를 이용한 버퍼 오버플로우 공격
식별과 인증을 침해하는 행위	CLI에서 이용한 무한 로그인 시도
접근제어를 침해하는 행위	CLI에서 TCPWrapper로 접근이 금지된 호스트로부터의 로긴 시도
시스템 자원의 비정상적인 소모	메모리 할당 프로그램을 이용한 시스템 자원 소모 공격
정상명령어에 의한 침입	패스워드 등의 시스템 파일의 일반 사용자 디렉터리로의 복사
단일 패킷에 의한 침입	한 개 호스트로부터의 DoS 공격
다중 패킷에 의한 침입	DDoS 공격
조각화된 패킷에 의한 침입	Nmap의 스텔스 기능을 이용한 포트 스캔 침입
보안상 중요한 파일의 네트워크 전송	FTP를 통한 임의의 호스트로부터의 passwd 파일 전송
Promiscuous 모드의 프로그램	Sniff 프로그램 수행

(CLI: Command Line Interface)

표10에서는 침입탐지시스템의 탐지 기능을 평가하기 위한 데이터 패턴 생성 방안을 보여준다. 이곳의 공격 방법들에 대한 설명은 [8]에 나타나 있다.

2. 품질 관점의 평가

표11. 품질 평가를 위한 데이터 패턴 생성 방법

품질적 요소	데이터 패턴 생성 방안
시스템 서비스를 이용한 침입	PING의 ICMP Flooding 공격
잘 알려진 프로그램에 의한 침입	NMAP을 이용한 스캔

표11의 시스템 서비스를 이용한 침입의 테스트 결과는 구체적인 서비스의 명칭을 찾아낼 수 있어야 한다. 잘 알려진 프로그램에 의한 침입은 침입 프로그램의 특성을 분석하여 구체적인 사용된 프로그램의 명칭을 보고할 수 있는가를 실험하는 것이다.

V. 데이터 패턴의 구현

1. 사용자 인터페이스

반복적인 평가가 필요하므로 사용자 인터페이스를 가진 시뮬레이터 프로그램을 구현하여야 한다. 이를 위하여 Qt 라이브러리[10]를 사용한다.

2. 시스템 내외부 사용자에 의한 침입

이미 UC Davis와 MIT의 실험에서 그 유용성이 검증된 expect를 사용하여 시스템 내외부로부터의 침입을 구현한다. Expect[9]는 tcl/tk의 확장 패키지로 개발된 컴퓨팅 환경의 사용자 시뮬레이션용 언어이다.

3. 응용 프로그램에 의한 공격

네트워크 프로그램을 위해서는 C언어를 expect와 함께 사용하고 웹을 통한 공격은 expect 만으로 구현한다.

4. 실험 데이터 패턴 생성기 제작

이 실험을 위하여 구현한 생성기는 실험에 쓰일 데이터들을 각각의 실험용 컴퓨터로 전송하고 설치하며 수행하는 역할을 하게 된다. 그리고 반복적인 실험을 위하여 전술된 행위들을 사용자의 조작에 따라 반복 수행하는 역할을 한다.

생성기의 구조는 그림 1과 같다. GUI는 시뮬레이터의 조작을 위한 사용자 인터페이스이며 GUI는 Control Module에서 발생한 명령 이벤트들을 입력으로 받아들여 환경 설정에 따른 Data Module의 전송, 설치, 동작을 수행한다. Data Module은 실험 데이터 패턴을 따라 개발되고 모아진 공격 패키지들과 수신 대상 컴퓨터들의 IP 주소를 가진 환경설정 데이터이다. Communication Module은 실험 컴퓨터들 상의 결과들을 보여주기 위해 실험자의 콘솔에 telnet을 통한 표준 I/O를 담당하는 expect의 고유 기능 부분이다.

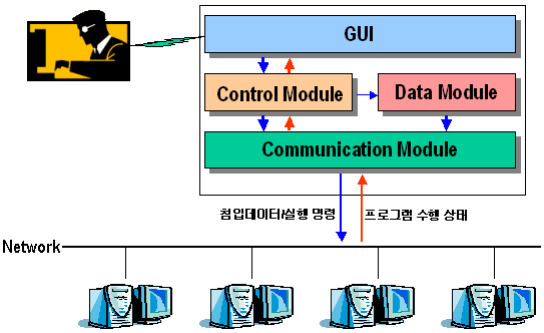

그림 1. 실험 데이터 생성기의 아키텍처

VI. 실험 데이터 패턴의 적용

이 실험에는 4대의 리눅스 머쉰을 사용하였고 그림 2와 같은 실험 환경 구조로 되어 있다.

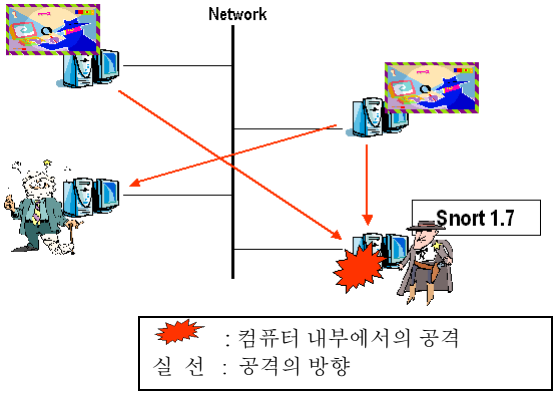

그림 2. 실험환경

표12. Snort에 대한 기능 평가 실험 결과

보안계층	기능적 데이터 패턴	결과
응용 계층	응용의 Exploit을 이용한 침입	무
운영체제 계층	식별과 인증을 침해하는 행위	유
	접근제어를 침해하는 행위	유
	시스템 자원의 비정상 소모	무
	정상 명령어에 의한 침입	무
네트워크 계층	단일 패킷에 의한 침입	유
	다중 패킷에 의한 침입	무
	조각화된 패킷에 의한 침입	무
	보안상 중요한 파일의 전송	유
	Promiscuous 모드 프로그램	무

표12와 표13는 snort[11]에 대한 각 평가 요소들의 소유 유무에 대한 결과를 나타낸다.

표13. Snort에 대한 품질 평가 실험 결과

품질적 데이터 패턴	결과
컴퓨터 서비스에 의한 침입	유
잘 알려진 공격 프로그램에 의한 침입	유

VII.결론 및 향후 과제

기존의 보안 시스템의 인증을 위한 국제 보안 시스템 평가 기준들은 단지 안정성에 관한 보증 결과를 산출한다. 이러한 안정성 인증은 침입탐지시스템의 탐지 능력에 대한 평가 결과를 나타내어 주지 못한다. UC Davis와 MIT는 침입탐지시스템의 침입 식별을 이용한 침입탐지시스템의 성능 평가를 수행하였으나 그들의 평가 방법은 단지 침입탐지시스템의 데이터의 보유수준에 따라 결과가 변동될 수 있다.

본 논문에서는 국제 보안 시스템 평가 기준의 흐름인 기능과 품질의 관점에서의 평가를 수행하는 방안을 제안한다. 침입탐지시스템의 개발자와 사용자들을 위한 침입탐지시스템 평가 요소들을 제시하고 공개된 침입탐지시스템인 snort를 평가 대상으로 그 구체적인 방법을 보여준다.

Snort에 대한 평가 결과를 통해서 snort가 가지는 기능들은 일반적으로 호스트 기반의 침입탐지시스템에서 고려되어야 할 많은 항목들이 결여되어 있다는 것을 알 수 있다. 그러나 품질의 관점에서는 두 가지 항목을 모두 만족하고 있다.

이 실험에서는 자체 제작된 실험용 데이터들을 사용하여 수동적으로 실험을 수행하였으나 향후에는 제작중인 생성기를 완성하여 실험 데이터의 조작이 쉽도록 할 것이며 그래픽 사용자 인터페이스를 통해서 실험을 위한 환경 설정이 가능하게 할 것이다.

참고문헌

[1] Terry Escamilla, ***Intrusion Detection : Network Security Beyond the Firewall***, WILEY, 1998.

[2] Sandeep Kumar, “Classification and Detection Of Computer Intrusions”, Ph.D thesis, Purdue University, 1995.

[3] 한국정보보호센터, *정보보호개론*, 교우사, 2000.

[4] Charles P. Pfleeger, ***Security In Computing 2nd Edition***, Prentice Hall Inc., 1997.

[5] 정보통신부, *정보통신망 침입탐지시스템 평가기준*, 정보통신부, July 2000.

[6] Nicholas Puketza, “A Software Platform for Testing Intrusion Detection Systems”, IEEE Software, 1997.

[7] Richard P. Lippmann, “Evaluating Intrusion Detection Systems : The DARPA Off-line Intrusion Detection Evaluation”, IEEE Computer Society Press, 2000.

[8] Stephen Northcutt, ***Network Intrusion Detection : An Analyst’s Handbook***, New Riders, June 1999.

[9] Don Libes, ***Exploring Expect***, O’REILLY, 1996.

[10] Trolltech, “Qt Library”, URL http://www.trolltech.com.

[11] Martin Roesch, “Snort”, URL http://www.snort.org.