중소기업정보기술융합학회 논문지 제2권 제2호 pp.1-6, 2012

인간의 면역체계 시스템을 적용한 침입 탐지자 생성 모델

신미예[1*], 최신형[2], 이상호[3]
[1]충북대학교 소프트웨어학과, [2]강원대학교 제어계측공학과, [3]충북대학교 소프트웨어학과

A Model of Applied to Immune System in Intrusion Detector

Mi-Yea Shin[1*], Shin-Hyeng Choi[2] and Sang-Ho Lee[3]
[1]Department of Software, Chungbuk National University,
[2]Department of Control & Instrumentation Engineering, Kangwon National University,
[3]Department of Software, Chungbuk National University

요 약 본 논문에서는 침입탐지 방법 중 오용탐지모델에서 오용탐지율을 향상시키기 위하여 인간의 면역 시스템을 적용한 탐지자 생성 모델을 제안 한다. DARPA에서 제공된 sendmail 데이터에 대하여 10CV 방법을 이용하였다. 정상적인 시스템 호출을 비정상적인 시스템 호출로 판단하거나 비정상 시스템 호출을 정상적인 시스템 호출로 판단하는 오용 탐지율을 실험하고 분석하였다. 실험에서 임의의 비정상적인 시스템 호출만으로 생성된 탐지자와 임시 탐지자 중에서 정상적인 시스템 호출을 '비정상'이라고 판단하거나 비정상적인 시스템 호출을 '정상'이라고 판단하는 임시 탐지자를 제거한 탐지자를 생성하여 실험하였다. 실험에서 임의의 비정상 시스템 호출만을 탐지자로 선택한 오용 탐지율보다 면역 시스템을 적용한 탐지자의 오용 탐지율이 0.3% 향상되었다.

키워드 : 침입 탐지자, 면역 시스템

Abstract In this paper, we propose a detector generation model which is applied to immune system to improve the misuse detection rates in misuse detection models. 10cv method is used to sendmail data which is provided by the DARPA. We experimented and analyzed the misuse detection rate that is either judgment of the normal system call as abnormal system call or judgment of the abnormal system call as normal system call. In the experiment, between detector which was generated by any abnormal system call and temporary detector. I did experiments with a new detector which was removed temporary detector which made a wrong decision for normal system call as an abnormal system call and abnormal system call as a normal system call. The misuse detection rate of detector which is applied to the immune system is greater than the other detector by 0.3%.

Key Words : Intrusion detector, Immune system

1. 서론

호스트 기반의 침입탐지 시스템은 시스템 호출 순서와 시스템 호출 매개변수를 고려한 모델들이 제안되었다[1-5]. 시스템 호출 순서 기반 침입탐지 방법은 시스템 호출의 순서 상관관계를 이용한다.

시스템 호출 매개변수를 기반으로 하는 침입 탐지 방법은 시스템 호출 순서만을 고려한 경우 시스템 호출 순서는 정상이지만 포맷 스트링 공격[6]과 같이 매개변수의 값만 변하는 공격을 탐지할 수 없으므로 시스템 호출 시에 전달되는 매개변수의 길이에 대한 평균 및 표준편

이 논문은 2012년 중소기업정보기술융합학회 추계학술발표대회의 우수논문을 확장한 것임.
*교신저자(e-mail: shinmiyea@nate.com)
접수일(2012년 11월 05일), 심사완료일(2012년 11월 30일)

차 등 통계적인 방법이 제안되었다.

기존 침입에 대한 정보를 기반으로 시스템의 비정상적인 행위를 탐지하는 오용탐지와 다르게 비정상 침입 탐지 시스템은 기존에 알려지지 않은 침입을 탐지할 수 있다는 장점을 가지고 있다. 멕시코 대학의 Forrest 교수에 의해 제안된 면역시스템은 인간이 항원으로부터 몸을 보호하는 면역체계가 마치 컴퓨터의 보안 체계와 유사한 점을 이용하였다.

비정상 침입 탐지 시스템 또는 오용 탐지 시스템에서 긍정적 결함율 및 부정적 결함율을 향상시키기 위하여 여러 단계의 과정을 수행하는 인간의 면역 체계 시스템을 침입 탐지 시스템의 탐지자 생성시에 적용할 수 있다.

따라서 본 논문은 인간의 면역시스템을 컴퓨터 보안 시스템에 적용하기 위해 시스템 호출이 정상적으로 이루어지는지 판단하기 위해서 10CV 방법을 적용하여 정상적인 시스템 호출을 비정상적인 시스템 호출로 판단하는 오용 탐지율을 향상시키기 위한 모델을 제안한다.

이 논문의 구성은 다음과 같다. 2장에서는 인간의 면역체계 시스템을 적용한 방법을 기술하고 3장에서는 DARPA에서 제공된 정상적인 시스템 로그 파일과 비정상적인 로그 파일을 이용하여 탐지자를 10CV 방법으로 추출하는 모델을 제안한다. 4장에서는 제안 모델을 10CV 방법에 따라서 오용 탐지율을 비교 분석한다. 마지막으로 5장에서는 결론을 맺는다.

2. 관련연구

인간의 면역 체계는 자연면역과 적응면역으로 구분되고 자연면역(innate immunity)은 감염에 대한 일차적인 방어선을 의미한다. 자연면역에 중요 역할을 담당하는 것은 대식세포 및 중성구와 같은 식세포와 피부 그리고 체내의 항생 물질 등이 있다.

적응면역(adaptive immunity)은 항원에 노출되고 나서 5 ~ 6일 후에 반응하는 것을 의미한다. 또한 같은 항원에 노출되면 기억된 면역 작용이 일어나므로 항원의 두 번째 침입에 대한 면역반응은 훨씬 신속하고 효과적으로 이루어진다. 림프구와 항체는 적응면역의 가장 중요한 요소이다[7, 8]. 이와 같이 인간의 면역체계 시스템이나 사회적 곤충인 개미와 꿀벌이 동료와 상호작용하는 모델을 컴퓨터 보안 시스템에 적용한 모델이 제안되었다.

2.1 면역의 단계적 방어 시스템

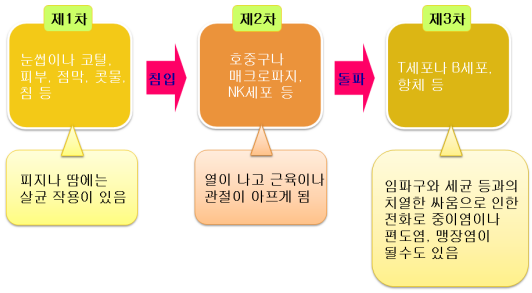

그림 1. 항체 생성 과정
Fig 1. Process of antibody creation

그림 1과 같이 인간의 단계적 면역 시스템은 피부나 눈썹 등에서 일차적으로 방어를 하고 세균들과 싸우는 과정에서 열과 염증 등이 발생되고 T세포 등에서 항체를 생성하게 된다.

2.2 negative selection

흉선에서 T-림프구가 negative selection 방법에 의해 생성되는 원리를 적용하여 컴퓨터에서 자기세포(self)를 인식하지 못하는 탐지자를 생성한다[9]. 자기세포(self) 문자열 집합으로부터 r-contiguous 방법에 의해 탐지자를 생성한다.

2.3 positive selection

침입을 받은 정보들을 이용하므로 자기세포(self) 집합은 공격을 받은 패턴들을 모아 놓은 집합을 의미한다, 요구되는 탐지자의 수 n이 충족될 때까지 임의의 수 d를 생성하여 자기세포(self)의 집합과 비교한 적합도가 임계치 보다 크면 탐지자에 추가하는 모델이다[10].

2.4 클론 선택

임의의 난수를 이용하여 탐지자 D를 생성하고, 인식되어질 항원 집합 S가 있을 때, 각 항원 s에 대하여 탐지자 집합 d와의 적합도를 구한다. 적합도가 가장 높은 n에 대해 돌연변이를 발생시켜서 이를 탐지자 D에 추가한다. 그리고 탐지자 D에서 적합도가 가장 높은 것을 memory에 저장한다[10].

2.5 danger theory

새롭게 침입한 항원에 대한 정보가 없어서 탐지할 수

도 없지만, 항원과 높은 적합도로 결합된 림프구들은 큰 수로 재생성 되도록 자극을 받는다. 이와 같이 “non-self”는 아니어도 위험한 손상을 입은 cell이 있는 일정 부분은 "danger zone"으로 처리하게 된다[10, 11].

2.6 swarm intelligence

개미는 페르몬을 분비하여 먹이의 위치를 알리는 것은 물론 다른 동료 개미들과 상호 협력하여 먹이를 옮긴다. 개미나 꿀벌과 같은 사회적 곤충의 상호작용을 컴퓨터 시스템 보안에 적용한 모델이 제안되었다[12].

3. 제안 모델

이장에서는 비정상적인 시스템 호출 순서와 정상적인 시스템 호출 순서를 이용하여 10CV 방법에 따라 탐지자를 생성하는 모델에 대하여 기술한다.

3.1 개요

임의의 정상적인 시스템 호출 70%를 탐지자로 선택하여 정상적인 시스템 호출 30%와 비정상적인 시스템 호출 30%를 탐지할 때 각각 오용 탐지하는 경우 탐지자에서 제거된다. 또한 임의의 비정상적인 시스템 호출 70%를 탐지자로 선택하여 정상적인 시스템 호출 30%와 비정상적인 시스템 30%를 탐지할 때 오용 탐지하는 경우 탐지자에서 제거된다.

3.2 정상적인 시스템 호출 순서에 의한 탐지자

그림 2는 정상적인 시스템 호출의 70%를 임의로 선택하여 임시 탐지자를 생성한 후 정상적인 시스템 호출 30%를 임시 탐지자가 비정상으로 판단하는 경우는 실제 탐지자로 선택되지 않고 정상이라고 판단된 시스템 호출만을 실제 탐지자로 채택하는 모델이다.

그림 3은 정상 시스템 호출 70%를 선택하여 임시 탐지자를 생성한 후 비정상적인 시스템 호출 순서를 정상적인 시스템 호출로 판단하는 탐지자는 제거하고 비정상적인 시스템 호출로 판단하는 탐지자만 선택하는 모델이다.

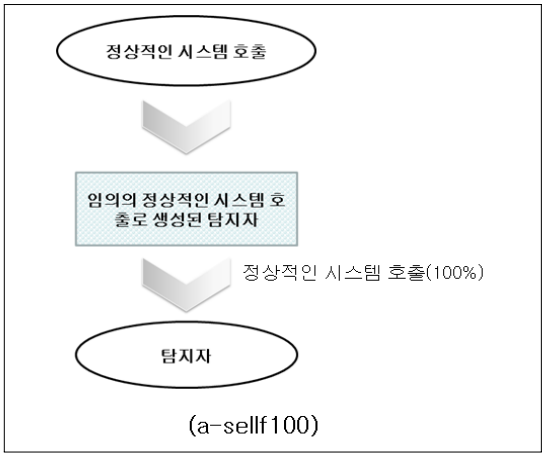

그림 2. 정상적인 시스템 호출를 이용한 정상적인 시스템 호출 탐지자 선택

Fig 2. Self system call detector selection using self system call

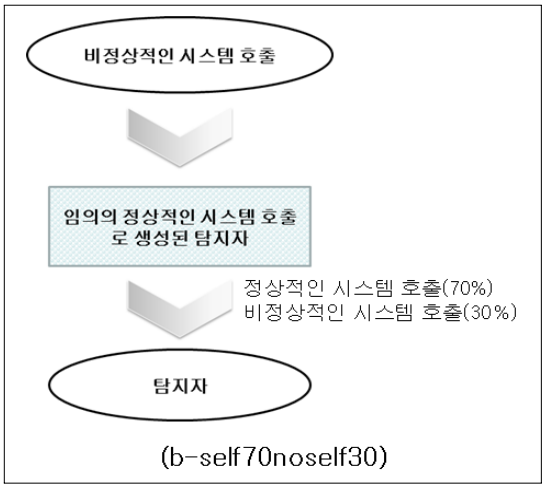

그림 3. 비정상적인 시스템 호출을 이용한 정상적인 시스템 호출 탐지자

Fig 3. Self system call detector selection using non-self system call

3.3 비정상적인 시스템 호출 순서에 의한 탐지자

그림 4는 비정상적인 시스템 호출의 70%를 임의로 선택하여 임시 탐지자를 생성한 후 정상적인 시스템 호출 30%를 임시 탐지자가 비정상적인 시스템 호출로 판단하는 경우는 실제 탐지자로 선택되지 않고 정상이라고 판단된 시스템 호출만을 실제 탐지자로 채택하는 모델이다.

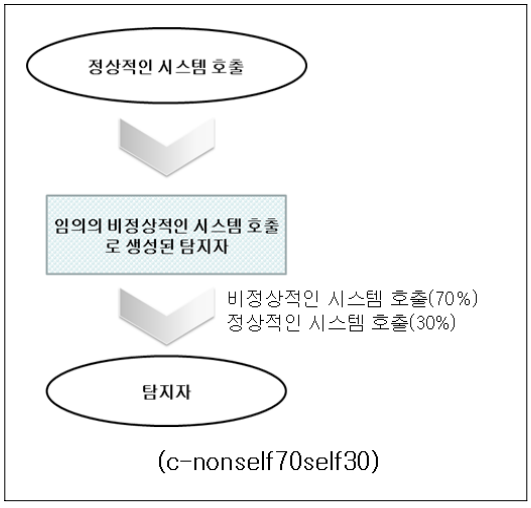

그림 4. 정상적인 시스템 호출을 이용한 비정상적인 시스템 호출 탐지자
Fig 4. Non-self system call detector selection using self system call

그림 5는 비정상 시스템 호출 70%를 선택하여 임시 탐지자를 생성한 후 비정상적인 시스템 호출 순서를 정상적인 시스템 호출로 판단하는 탐지자는 제거하고 비정상적인 시스템 호출로 판단하는 탐지자만 선택하는 모델이다.

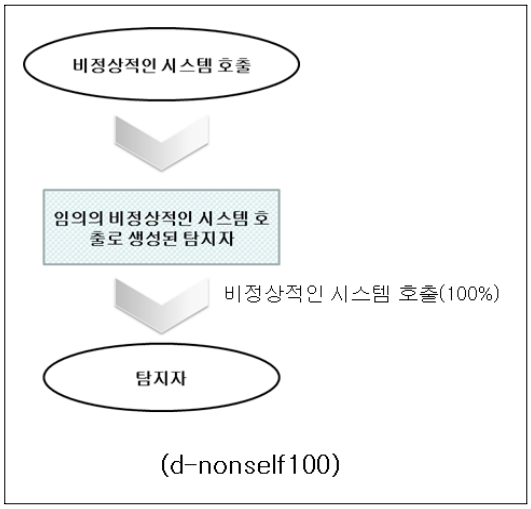

그림 5. 비정상적인 시스템 호출을 이용한 비정상적인 시스템 호출 탐지자
Fig 5. Non-self system call detector selection using non-self system call

4. 평가

4.1 개요

제안 모델에서의 실험평가는 DARPA 데이터로 실험 비교 평가하였다. 시스템 콜의 윈도우 크기는 30으로 하였다. 탐지자는 143개의 정상적인 시스템 호출 파일들 중에서 임의로 10%, 20%, 30%씩 그룹화한 후, 한 그룹 내에서 10CV방법으로 탐지자를 생성하고 정상적인 시스템 호출 파일 30%와 비정상적인 시스템 호출 파일 30%에 대한 탐지율을 실험하였다. 제안 모델의 실험 환경은 OS는 Windows 7이고 컴파일러는 Visual C++ 2010을 이용해 감사 데이터를 추출한다.

표 1. 하드웨어 환경
Table 1. Hardware specification

구분	내용
컴파일러	Visual C++ 2010
메모리	2GB
CPU	AMD AthlonTM 64 * 2 Dual Core Processor 4000+
OS	Window 7
데이터	1999 DARPA

4.2 실험방법

실험 데이터는 DARPA에서 제공된 데이터를 이 모델에서 필요한 요소로 추출하여 임시 탐지자에 저장하고, 새로운 이벤트가 발생되었을 경우 비정상으로 탐지된 경우 보고 및 조치가 이루어진다.

4.2.1 실험 방법 1

그림 2와 그림 3의 모델에서는 임의의 정상적인 시스템 호출로 생성된 임시 탐지자를 이용하여 정상적인 시스템 호출을 '비정상'이라고 판단하거나 비정상적인 시스템 호출을 '정상'이라고 판단하는 임시 탐지자들을 제거한 후 최종 탐지자로 선택한다. 이와 같이 생성된 최종 탐지자와 임의의 정상적인 시스템 호출로 생성된 임시 탐지자에 대하여 정상적인 시스템 호출을 '비정상' 이라고 판단하는 긍정적 결함율과 비정상적인 시스템 호출을 '정상'이라고 판단하는 부정적 결함율을 실험했다.

4.2.2 실험 방법 2

그림 4와 그림 5의 모델에서는 임의의 비정상적인 시

스템 호출로 생성된 임시 탐지자를 이용하여 정상적인 시스템 호출을 '비정상'이라고 판단하거나 비정상적인 시스템 호출을 '정상'이라고 판단하는 임시 탐지자들을 제거한 후 최종 탐지자로 선택한다. 이와 같이 생성된 최종 탐지자와 임의의 비정상적인 시스템 호출로 생성된 임시 탐지자에 대하여 정상적인 시스템 호출을 '비정상' 이라고 판단하는 긍정적 결함율과 비정상적인 시스템 호출을 '정상'이라고 판단하는 부정적 결함율을 실험했다.

4.3 실험 결과

그림 2와 그림 3의 모델에 대한 실험에서는 긍정적 결함율과 부정적 결함율은 각각 0.7%와 0.8%로 나타났으며, 그림 4와 그림 5의 모델에 대한 실험에서는 표 2와 같이 오용 탐지율이 각각 0.6%와 0.9%였다. 따라서 이 실험에서 면역 시스템을 적용한 탐지자의 성능은 윈도우 크기를 30으로 하고 10CV 방법으로 비정상적인 시스템 호출을 이용한 탐지자가 오용 탐지율에서는 더 효과적으로 나타났다.

표 2. 10CV긍정적 결함율 및 부정적 결함율
Table 2. 10CV false positive rate and false negative rate

제인모델	긍정적 결함율(%)	부정적 결함율(%)
A-Self100	0.7	
B-self70nonself30		0.8
C-nonself70self30	0.6	
D-nonself100		0.9

5. 결론 및 향후 연구

이 논문에서는 효과적으로 오용 탐지율을 향상시키기 위하여 인간의 면역 시스템을 적용한 탐지자 생성 모델을 제안하였다. 제안 모델은 DARPA 자료를 이용하였다. 실험에서 임의의 비정상 시스템 호출을 그대로 이용한 탐지자의 오용 탐지율은 0.6%, 면역 시스템을 적용한 탐지자의 오용 탐지율은 0.9%로 향상되었다.

향후 연구에서는 실험에서 사용된 DARPA에서 제공된 자료 외에 다양한 종류의 시스템 호출이 필요하며, 오용 탐지 모델뿐만 아니라 비정상 탐지 모델에 면역 시스템을 적용한 탐지자를 고려하는 연구가 필요하다.

참 고 문 헌

[1] S. Forrest, Steven A. Hofmeyr, Anil Somayaji, Thomas A. Longstaff, "A Sense of Self for Unix Process", In Proc. of the 1996 IEEE Symposium on Research in Security and Privacy, Los Alamos, CA, pp. 120-128. IEEE Computer Society Press.

[2] J. B. D. Cabrera, L. Lewis, and R.K. Mehara. "Detection and classification of intrusion and faults using sequences of system calls". ACM SIGMOD Record, Vol.30 No.4, 2001.

[3] G. Tandon and P. Chan. "Learning rules from system call arguments and sequences for anomaly detection". In ICDM Workshop on Data Mining for Computer Security (DMSEC), pp 20 - 29, 2003.

[4] G. Casas-Garriga, P. Diaz, and J.L. Balcazar. "ISSA : An integreated system for sequence analysis". Technical Report DELIS-TR-0103, Universitat Paderborn, 2005.

[5] 황현욱, 김민수, 노봉남, " 감사로그 상관관계를 통한 호스트 기반의 침입탐지 시스템", 한국정보보호학회 논문지, 제13권 제3호, pp. 81-90, 2003.

[6] D. Wagner and P. Soto. "Mimicry attacks on host based intrusion detection systems". In ACM conference on Computer and Communications Security (CCS), 2002.

[7] KUBY 면역학 5판, 강호영 외12인 역, 월드사이언스, 2006

[8] 자연면역(innate immunity)
(류호룡의 한방과 건강 http://www.herbu.co.kr)

[9] S.Forrest, A. S Perelson, L. Allen and R. Cherukuri, Self-nonself discrimination in a computer, In Proceedings of the IEEE Symposium on Research in Security and Privacy, IEEE Computer Society Press, Los Alamotos, CA, pp. 202-212, 1994

[10] http://www-course.cs.york.ac.uk

[11] Uwe Aickelin1, Steve Cayzer, The Danger Theory and Its Application to Artificial Immune Systems,International Conference on Artificial Immune Systems, Canterbury, UK. 2002

[12] Diego Alejandro Ingaramo, Guillermo Leguizamon, Marcelo Errecalde, "Adaptive clustering with artificial ants", Journal of Computer Science &Technology, December, 2005

저 자 소 개

신 미 예(Mi-Yea Shin) [정회원]

- 1990년 8월: 한밭대학교 전자계산학과 학사
- 1998년 8월 : 충북대학교 전자계산학과 석사
- 2010년 2월 : 충북대학교 전자계산학과 박사

<관심분야> : 유·무선 네트워크 보안, 정보보호, 인공지능, 정보검색

최 신 형(Shin-Hyeong Choi) [정회원]

- 1993년 2월: 울산대학교 전자계산학과 공학사
- 1995년 2월 : 경남대학교 전자계산학과 공학석사
- 2002년 8월 : 경남대학교 컴퓨터공학과 공학박사
- 2003년 9월 ~ 현재 : 강원대학교 제어계측공학과 부교수

<관심분야> : 임베디스 시스템, USN, 유·무선 보안

이 상 호 (Sang-ho Lee) [정회원]

- 1976년 2월: 숭실대학교 전자계산학과 학사
- 1981년 2월: 숭실대학교 전자계산학과 석사
- 1989년 2월: 숭실대학교 전자계산학과 박사
- 1981년 3월 ~ 현재 : 충북대학교 전자정보대학 교수

<관심분야> : 컴퓨터 네트워크, 정보보호, 데이터 통신