생체 면역시스템을 이용한 하이브리드 침입 탐지 시스템 설계

양은목*, 이상용**
*공주대학교 컴퓨터 공학과
**공주대학교 정보통신공학부
e-mail : {*emyang,**sylee}@kongju.ac.kr

The Design of a Hybrid Intrusion Detection System using Immune Systems

Eun-mok Yang*, Sang-yong Lee**
*Dept of Computer Science & Eng, Kongju National University,
**Division of Information & Communication, Engineering

요약

최근 컴퓨터와 인터넷의 급속한 발전과 더불어 컴퓨터의 데이터를 파괴하거나 바이러스를 이용해 정보를 빼내기 위한 해킹 등이 만연하고 있다. 이에 컴퓨터의 데이터를 외부 침입물질에 대해 자체적인 보호와 제거 기능을 가진 생체 면역시스템을 이용한 연구가 활발히 진행되고 있다. 생체 면역 시스템은 바이러스나 병원균 등의 낯선 외부의 침입자로부터 자신을 보호하기 위해 크게 선천성 면역과 후천성 면역을 제공한다.

본 논문은 선천성 면역에는 오용탐지기법과 후천성 면역에는 비정상행위 탐지 기법을 이용한 하이브리드 침입탐지 시스템을 제안한다. 감사 자료 수집은 멀티레벨 파라미터 모니터링을 통해 감사 자료를 수집한다. 선천성 면역에서는 피부와 여러 가지 감각 기관의 분비물을 이용하듯이 방화벽과 같은 비슷한 기능을 하는 서비스 제한 에이전트와 기존에 알려진 버그와 해킹 기법을 시나리오 지식베이스를 이용하는 오용탐지 기법을 사용한다. 그리고, 후천성 면역에서는 유전자 알고리즘을 이용해 침입을 탐지하고 대응한다.

1. 서론

컴퓨터 바이러스는 생물학의 바이러스와 같은 자기 복제와 파괴능력을 갖고 컴퓨터에서 실행되는 프로그램의 일종으로서 감염 대상인 컴퓨터 프로그램이나 데이터 파일을 파괴한다. 최근 컴퓨터의 사용이 보편화되면서 악의적 사용자에 의해 이러한 컴퓨터에 해킹과 바이러스에 의한 피해가 급속히 증가하고 있으며, 파일의 손상에 그치지 않고 시스템을 파괴하거나 전체 네트워크를 마비시키는 바이러스들도 자주 등장한다. 해킹은 주로 다른 사람의 컴퓨터에 침입해 정보를 가져오거나 그 컴퓨터가 가지고 있는 정보를 없애는 작용으로, 인터넷의 지속적인 발전에 따라 하나의 네트워크로 연결된 많은 컴퓨터가 피해를 입고 있다. 컴퓨터에 침입하는 해킹이나 데이터를 파괴하는 컴퓨터 바이러스에 의한 피해를 막기 위해 최근에 생명체 면역시스템의 특징을 이용한 시스템 침입탐지와 바이러스 탐지 및 치료에 대한 연구가 활발히 진행되고 있다[1][2].

침입탐지시스템은 네트워크, 호스트, 다중 호스트 기반의 침입탐지시스템으로 분류할 수 있다. 네트워크 기반의 침입탐지는 전체적인 네트웍에 과부하를 줄 수 있고 호스트 기반의 오용탐지는 새로운 해킹기법에 신속한 대응이 미비하고 비정상행위 탐지 기법은 기존에 알려진 버그나 해킹기법에 능동적인 대응이 부족하다.

본 논문에서는 침입탐지 시스템에서 방화벽과 유사한 기능을 하는 서비스 제한 에이전트를 도입함으로 인해 시스템의 외부물질을 탐지하는데 드는 비용과 시간을 절약할 수 있다. 또 선천성 면역에는 기존에 잘 알려진 버그나 해킹방법을 지식베이스에 저장하였다가 침입이 발생하면 신속하게 대응할 수 있도록 하였고, 후천성 면역에서는 선천성 면역에서 탐지 할 수 없었던 침입에 대해 학습하고 적응할 수 있는 시스템을 제안한다.

감사 자료 수집에서는 사용자 레벨, 시스템 레벨, 프로세스 레벨, 패킷 레벨 파라미터 모니터링을 수행하였다. 이로 인해 호스트 기반에서 탐지 할 수 없었던 네트워크의 감사자료를 통한 침입탐지도 가능하게 되었다[3][4].

선천성 면역에 해당하는 오용탐지 모델은 응용프

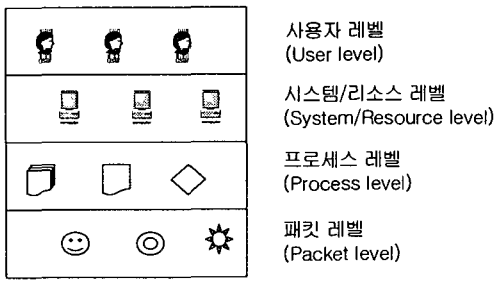
로그램의 약점을 통하여 시스템에 침입하는 이미 알려진 공격행위를 면역 메모리(memory cell)에 저장한 후 동일한 공격행위를 면역 메모리와 비교를 통해서 탐지하는 기법이다 이때 면역 메모리에 저장되어 있는 공격행위가 정확한 것만 구별할 수 있어야 하며, 그렇지 못한 경우 부정적 결함(false negative error)이 발생할 확률이 높다.

후천성 면역에 해당하는 비정상행위 탐지 모델은 시스템 또는 사용자의 행위가 정상 행위로부터 벗어난 경우를 탐지하는 것으로 시스템 또는 사용자의 정상적인 행위를 면역 메모리에 저장한 후 이 감사 데이터로부터 유전자 알고리즘을 사용하여 정상 행위를 추출한 후, 지금 수행되는 행위가 정상 행위로부터 벗어나면 불법 침입으로 간주한다. 이 방법은 예측하지 못한 시스템 취약점을 이용하려는 시도를 탐지할 수 있어 새로운 침입을 자동으로 탐지 할 수 있다. 그러나 면역 메모리에 저장된 정상행위 정보가 모든 정상행위를 포함하지 못하기 때문에 긍정적 결함(false positive error)이 발생할 확률이 높다.

2. 관련연구

2.1 감사자료 수집과 필터링

연결유형과 사용자의 유형의 해석에 대하여 수치적 표현이 가능할 만한 파리미터를 모니터링 한다. 여러 가지 Unix 명령어를 사용하고, 선택된 값들을 얻기 위해 결과를 필터링 한다(그림 1).



(그림 1) 다른 레벨의 파라미터 모니터링

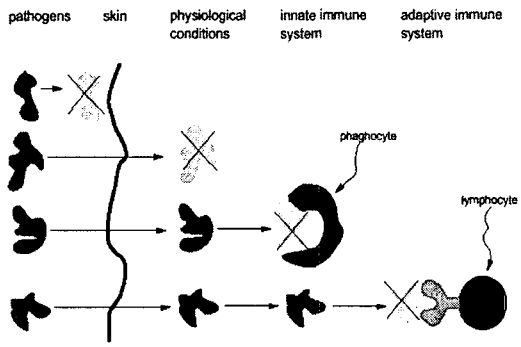
사용자 레벨의 파라미터들은 주로 시스템 로그에 의한 것들로서 사용자의 유형과 권한, 로그인과 로그아웃의 주기와 위치, 엑세스 디렉토리와 리소스, 사용 소프트웨어 또는 프로그램의 종류, 사용 명령어의 종류 등을 수집하고 필터링 한다. 사용자의 파라미터들을 수집하는 것은 세션을 시작한 ID에 기초를 두고, 매 세션은 ID에 따라서 구분 가능하다.

시스템/리소스 레벨의 파라미터들은 시스템 사용양의 측정값들을 수집하는 것들로서 사용자당 CPU 사용량, 실제 메모리와 가상 메모리 사용량, 현재 사용 가능한 스왑(swap)의 양, 사용 가능 메모리의 양, I/O와 디스크 사용량 등의 현재 탐지되고 있는 시스템의 주요 리소스들을 모니터링을 하므로 인해서 감사자료를 수집하고 필터링 한다. 이 모든 파라미터들은 시스템 자원의 사용 정보를 제공하고 처음 몇 번은 네트워크와 호스트 기반상의 침입 탐지를 하기 위해 정상적 시스템 사용을 비교적 정확하게 모니터링 한다. 시스템 레벨의 정보를 얻기 위한 명령어로 vmstat(가상 메모리 정보 제공)와 iostat(I/O 통계를 제공)를 사용한다.

프로세스 레벨의 파라미터들은 어떤 작업들이 이루어지는 것을 실시간으로 모니터링 하는 것으로 프로세스 수와 유형, 프로세스 사이의 관계, 프로세스 시작 후 경과 시간, 프로세스 현재 상태(실행, 중지, 준비)와 급증 프로세스, 여러 가지 시간(사용자 프로세스 시간, 시스템 프로세스 시간과 유휴시간)의 백분율 등을 감사자료로 수집하고 필터링 하여 침입 탐지에 사용한다. 프로세스 레벨의 프로세스 ID, 터미널 식별자, 누적실행 시간, 명령어 이름 등을 얻기 위한 명령어로 ps - ef(활동 중인 프로세스의 정보를 출력)와 ps(제어 터미널과 프로세스를 조합한 정보를 출력)를 사용한다.

패킷 레벨의 파라미터들은 현재 호스트에서 사용되어지고 있는 네트워크의 모든 상황을 모니터링 하는 것으로 연결의 수와 연결상태(예. close_wait, time_wait, established), 보내고 받는 패킷의 평균 크기, 연결 지속 시간, 연결의 유형(remote/local)과 프로토콜과 포트 등의 정보 등을 감사 자료로 수집하고 필터링 한다. 패킷 레벨의 정보를 얻기 위한 명령어로 netstat(네트워크의 여러 가지 데이터를 다양한 형태로 제공)를 사용한다.

2.2 자연 면역



(그림 2) 생체 면역시스템

생체 면역 시스템은 병원균이 침입으로부터 1차적은 방어는 피부가 담당하고 2차 적인 대응은 선천적 면역 시스템인 자연 면역이 담당한다(그림 2). 2차 방어선인 자연 면역을 통과하게 되면 3차 방어인 후천적 면역 시스템이 담당하게 되는데 이때 병원균에 대한 학습과 적응이 이루어지게 된다[4][5].

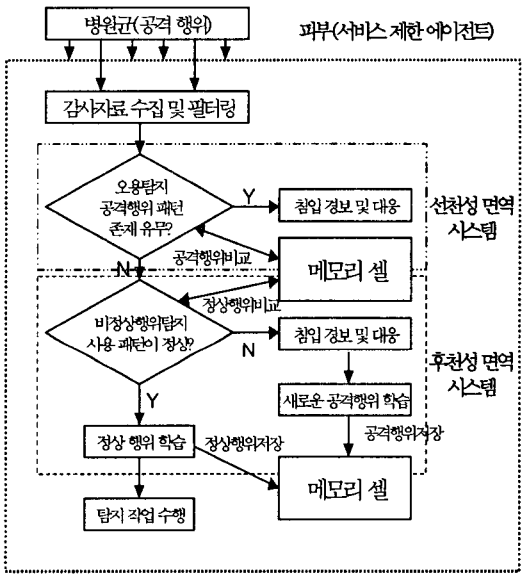
유전자 알고리즘은 자연적인 생물학 전개의 흉내낸 다목적이고 복합적인 알고리즘이다[6].

유전자 알고리즘에서 염색체 신텍스는 문제를 공간에 표현하는 것이고, 염색의 해석은 실행가능을 해석하고 체크한다. 또 염색체의 적정 평가는 문제

해결의 질을 결정한다. 마지막으로 유전학적인 연산은 교배와 돌연변이를 이용하여 후보자를 결정하고 조정하는 알고리즘을 수행하게 된다.

3. 침입 탐지 시스템 설계

(그림 3)은 전체적인 침입탐지 시스템의 동작 알고리즘을 나타내며, 3.1에서 선천성 면역시스템과 3.2에서 후천성 면역시스템에 대하여 자세하게 기술한다.



(그림 3) 제안한 침입탐지 시스템 동작 알고리즘

3.1 선천성 면역

1차 방어인 서비스 제한 에이전트를 통과한 외부 침입에 대하여 오용탐지로서 2차적 대응을 하게 된다. 여기에서는 필터링된 감사자료와 메모리 셀과 패턴 비교로서 공격을 탐지하게 된다. 여기서 침입탐지 시스템의 오류로 치명적인 결함이 될 수 있는 부정적 결함을 비정상행위 탐지 기법으로 다시 한번 비교하게 된다. 이로 인해, 오판으로 시스템에 악영향을 줄 수 있는 요인을 최대한 줄였다.

여기서, 메모리 셀에 있는 정보는 잘 알려진 버그나 공격 방법을 미리 학습시키고, 새로운 공격 방법의 학습은 비정상행위 탐지에서 공격행위로 판정된 것을 학습하게 된다.

3.2 후천성 면역

오용탐지 기법에서 탐지 못한 것들과 정상행위 들이 비정상행위 탐지 기법의 입력으로 들어오게 된다. 부정적 결함을 유전자 알고리즘 이용하여 탐지하게 된다.

(표 1) 파라미터 값의 바이너리 인코딩

0	00	Normal
1	01	Minimal
2	10	Significant
3	11	Dangerous

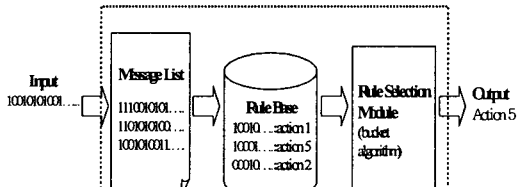
(표 1)은 침입 정도를 0~3으로 표시하고 2비트 문자열로 변환한다.

학습의 과정은 조건과 선택으로 된 고정길이 문자열로 유전자 알고리즘의 선택의 방법을 진화시켜 환경에 대처하기 위한 적응형 학습시스템이다. 보안정책을 바탕으로 정상행위를 정의해서 내장시킨다. 이것은 조건(0~3)과 공격 유형에 따라서 명확한 반응(0~7)으로 표현한다. 전문가들에 의해 디자인된 일반적인 지식베이스로부터 일반적인 규칙을 만들고, 그러한 지식베이스는 (표 2)와 같이 반응의 높은 값은 조금 더 강한 대응을 한다.

(표 2) 서로 다른 레벨들에서 침입한 활동들에 대한 평가들과 반응작용의 예

가설	유저 레벨	시스템 레벨	프로세스 레벨	패킷 레벨	반응
1	0.2	0.0	0.0	0.1	1
2	0.4	0.0	0.0	0.4	3
3	0.0	0.1	0.2	0.8	6

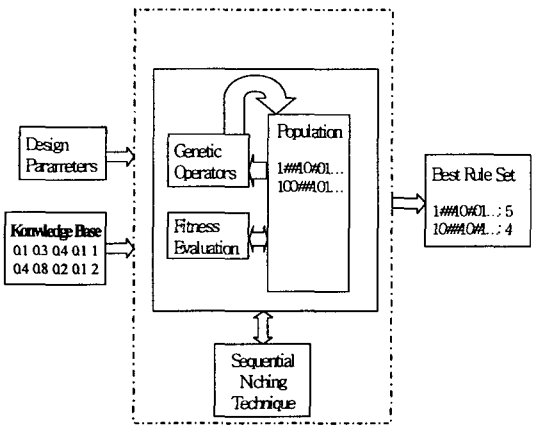
(그림 4)는 바이너리 스트링을 입력으로 하는 분류에 근거한 결정 지원 시스템으로, 입력에서 일정 프레임의 범위를 잡고, 유전자 알고리즘으로 추출한 최적의 규칙 값과 비교하는 규칙 선택 모듈(Rule Selection Module)을 통해서 최적의 반응 값을 찾아낸다.



(그림 4) 분류에 근거한 결정지원 시스템

(그림 5)는 (그림 4)에서 제안된 분류에 근거한 결정 지원 시스템에 유전자 알고리즘을 적용한 일반적

인 모듈이다. 이로서 비정상 행위를 탐지기법인 후천성 면역은 감사자료 수집과 필터링이 이루어진 후 바이너리 스트링으로 변환, 유전자 알고리즘을 이용한 학습과 적응의 단계를 거치면서 탐지율을 높이게 된다. 또 적합도 평가를 통해 학습의 강도를 조정할 수 있다. (그림 4)에서 보이는 규칙 선택 부분은 (그림 5)에서 생성한 베스트 규칙 중에서 대응할 규칙을 선택하는 부분이다. 서열비교평가를 통해 새로운 규칙을 찾아내고 진화시켜 새로운 규칙을 생성한다.



(그림 5) 규칙에 근거한 일반적인 유전자 알고리즘 모듈

4. 결론 및 향후 연구 과제

본 논문에서는 생체 면역 시스템에 가장 가깝게 접근할 수 있도록 1차적 방어인 피부와 2차적 대응인 선천성 면역 그리고 면역의 마지막 단계인 후천성 면역을 바탕으로 서비스 제한 에이전트와 오용탐지기법, 그리고 비정상행위 탐지기법을 이용한 침입탐지 시스템을 제안하였다. 그리고 침입 탐지 시스템의 치명적인 결함인 부정적 결함을 오용탐지와 비정상행위 탐지 기법을 동시에 사용함으로써 원천적으로 봉쇄하고, 선천성 면역에서는 잘 알려진 버그와 해킹 유형을 메모리 셀에 규칙기반으로 저장하여 침입이 이루어지면 곧 바로 대응할 수 있도록 하였다. 또 후천성 면역에서는 정상행동을 학습해 메모리 셀에 규칙기반으로 저장하고, 또 침입으로 탐지된 것은 다음에 유사한 침입이 이루어졌을 경우 선천성 면역에서 탐지할 수 있도록 설계하였다.

향후 연구 과제로서는 유전자 알고리즘의 적합도 함수를 구하는 부분과, 제안한 시스템을 구현하고 테스트 작업을 통해 본 시스템의 우수함과 실효성을 검증하는 작업이 요구된다.

참고문헌

[1] Dipanker Dasgupta, "An Immune Agent Architecture for Intrusion Detection", Genetic and Evolutionary Computation Conference Workshop Program pp.42 - 44, 2000.

[2] Paul K. Harmer, Gary B. Lamont, "An Agent Based Architecture for Computer Virus Immune System", Genetic and Evolutionary Computation Conference Workshop Program pp.45 - 46, 2000.

[3] Dipankar Dasgupta, "Immunity-Based Intrusion Detection Systems: A General Framework", In the proceedings of the 22nd National Information Systems Security Conference (NISSC), October 18-21, 1999.

[4] S. A. Hofmeyr. "An Interpretative Introduction to the Immune System" Oxford University Press, Eds, I. Cohen and L. Segel. 2000.

[5] Dopankar Dasgupta., "An Artificial Immune System as A Multi-agent Decision Support System", In proceedings of the IEEE International Conference on System, Man, and Cybernetics, San Diego, pages 3816 - 3820, 1998.

[6] Crosbie M., Spafford G., "Applying Genetic Programming to Intrusion Detection", COAST Laboratory, Purdue University, 1997.