DEVS 모델링 및 시뮬레이션을 이용한 침입 탐지 기법의 성능평가

이장세, 이종근, 이종성, 박종서, 지승도
한국항공대학교 컴퓨터공학과
Tel: 02-3158-4866
E-mail : jslee2@mail.hangkong.ac.kr

Performance Evaluation for Intrusion Detection Techniques Using the DEVS Modeling and Simulation

Jang-Se Lee, Jong-Keun Lee, Jong-Sung Lee,
Jong-Seo Park, and Sung-Do Chi
Department of Computer Engineering
Hangkong University, Seoul, Korea

Abstract

본 연구는 DEVS 모델링 및 시뮬레이션을 이용한 침입 탐지 기법의 성능평가를 주목적으로 한다. 최근 컴퓨터망의 확대와 컴퓨터 이용의 급격한 증가에 따른 부작용으로 컴퓨터 보안 문제가 중요하게 대두되고 있으며 이러한 추세에 따라 해커들로부터의 침입을 줄이기 위한 침입 탐지 시스템에 관한 연구가 활발히 진행되고 있다. 한편, 침입 탐지 기법으로 전문가 시스템, 신경망, 유전자 알고리즘 등 인공지능 기법을 이용한 다양한 시도가 이루어지고 있으나 이러한 기법들에 대한 성능평가는 대부분 실제 시스템의 구축을 통해서만 다루어 왔다. 따라서, 이를 극복하기 위하여 시뮬레이션 기법의 도입을 통한 성능평가 방법이 요청된다. 따라서, 본 연구에서는 엔진 베이스 모델링을 통하여 일반적인 침입 탐지 시스템을 설계하고 침입 탐지 기법의 하나인 유전자 알고리즘을 적용하여 시뮬레이션 테스트를 수행함으로써 DEVS 모델링 및 시뮬레이션을 이용한 성능평가의 타당성을 검증한다.

1. 서론

컴퓨터 및 네트워크 기술이 발전함에 따라 컴퓨터 간의 상호 연결성이 증가되고 이로 인해 컴퓨터 보안 문제가 중요하게 대두되고 있으며 네트워크나 시스템 상에서 발생될 수 있는 공격이나 오용을 탐지하기 위한 침입 탐지 시스템에 관한 연구가 활발히 진행되고 있다[1]. 그러나, 침입 탐지 시스템의 평가를 위한 방법이나 표준에 관한 연구는 미흡한 실정으로 DARPA와 MIT Lincoln Laboratory에서 컴퓨터 시스템과 네트워크에 대한 공격을 탐지하기 위한 침입 탐지 시스템의 성능

평가가 시도된 바 있으며[2] IBM의 Zurich Research Laboratory에서는 기존의 침입 탐지 시스템들 간의 비교를 위한 실험장치(Workbench)를 개발[3]하는 등 침입 탐지 시스템의 평가에 대한 연구가 진행된 바 있으나, 대부분이 평가를 위한 테스트 데이터의 획득 또는 테스트를 위한 실제 네트워크 환경 구축에 대한 연구에 그치고 있는 실정이다. 따라서, 본 논문에서는 시스템에 대한 엔진 베이스 모델링을 통하여 침입 탐지 시스템의 시뮬레이션 모델을 설계하고 침입 탐지에 적용되는 여러 가지의 탐지 기법들을 적용함으로써 실제 시스템의 구축과 적용에 따른 문제점을 해결하고 다양한 성능 평가의 기반을 제공할 수 있다.

2. 모델베이스 설계 방법 개요

2.1. SES/MB

Zeigler에 의해 제안된 SES/MB(System Entity Structure/Model Base)는 시뮬레이션의 동역학적 방법론과 AI의 기호적 방법론을 체계적으로 통합함으로써 시스템 모델링 및 시뮬레이션 환경을 제공한다. SES/MB는 System Entity Structure와 Model Base의 두 구성원으로 이루어진다. SES는 시스템의 구조적 특성을 나타내는 것으로 선언적 성격을 가지며 구성관계, 구성원의 종류, 구성원들의 결합구조, 그리고 제약조건 등의 구조적 지식을 표현할 수 있는 수단을 제공한다[4]. MB는 시스템의 행위적 특성을 나타내는 것으로서 절차적 성격을 가지며 동역학적이고 기호적으로 행위를 표현할 수 있는 수단을 제공하는 모델들로 구성된다[5]. 그림1은 SES/MB개념의 예로서 MB에 저장된 행동적 모델(A, B, C, D, E, EF)들이 SES에 저장된 구조적 관계에 따라 변환 작용을 통해 통합되어 최종적인 시뮬레이션 모델로 구축되는 과정을 보여주고 있다. 여기서 실험장치(EF : Experimental Frame)의 개념은 생성된 시스템에 주어진 입력을 가하고 그 결과(출력)를 받아 볼 수 있는 성능분석용의 모델링 방법으로서, 모듈화된 시뮬레이션 테스트 환경을 제공한다[4,5].

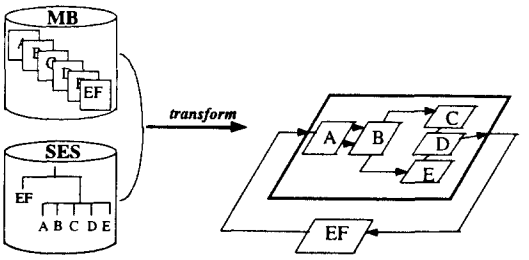


그림1. SES/MB 개념도

2.2. 엔진기반 모델링

Zeigler와 Chi가 제안한 엔진기반 모델링 방법은 전문가 시스템 개념을 도입한 것으로 대상체의 지식과는 독립적으로 설계 운용되는 추론 엔진(inference engine)과 대상체의 지식에 의존하는 지식베이스(knowledge base)로 구성된다. 여기서 추론엔진은 필요시 지식베이스를 참조하여 추론 알고리즘에 의해 최적의 결론을 도출해 내는 알고리즘체로서, 이를 통하여 설계의 일관성 및 용이성을 제공할 수 있다. 즉, 엔진모델과 지식베이스 모델을 분리시킴으로써 시스템의 모듈화와 추상화를 보다 유연하게 할 수 있다. 그림2는 엔진베이스 모델링 개념을 보여준다. 그림에서 원통모양의 모델들은 지식베이스를 나타내며, 실세계에 대한 지식(즉, 대상체에 대한 각종 정보들)들을 담게 된다. 타원형의 모델들은 추론엔진 모델로서, 지식베이스 모델을 알고리즘적으로 처리하여 최적의 결정 사항을 도출해 내는 역할을 담당한다[4,6].

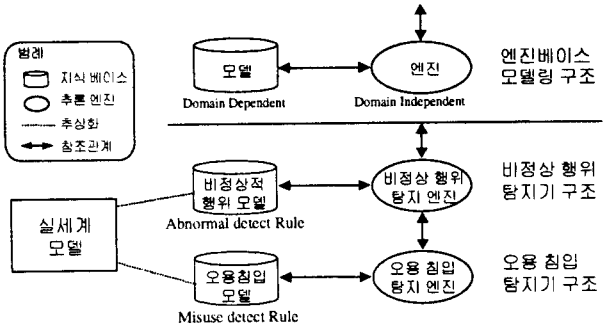


그림2. 엔진기반 모델링 개념

3. 유전자 알고리즘을 적용한 침입탐지 시스템

침입 탐지 기법으로 전문가 시스템, 신경망, 유전자 알고리즘 등 다양한 인공지능 기법이 적용될 수 있으며 본 논문에서는 탐색 및 최적화 기법인 유전자 알고리즘을 적용한 침입 탐지 시스템을 설계하였다.

3.1 침입 탐지 시스템 구조

그림3은 침입 탐지 시스템의 시뮬레이션 모델을 나타낸다. 여기서 EF(Experimental Frame)의 Generator와 Transducer는 각각 사용자 모델과 시스템 모니터링 모델로서 사용자의 입력을 발생시키고 입력에 대한 침입 탐지 시스템의 결과를 보고 받는다. EF와 결합된 IDS는 크게 감사 이벤트 변환기(Audit Event Translator), 침입 탐지기(Intrusion Detector), 경고기(Alarmer)로 구성되며 감사 이벤트 변환기는 사용자 입력을 이벤트로 변환시키고 변환된 이벤트를 침입 탐지기로 보내는 역할을 담당한다. 침입 탐지기는 엔진 베이스 모델링에 의하여 이미 알려진 공격 이벤트를 가지고 있는 공격-이벤트 지식베이스와 추론 엔진으로 구성되며 추론 엔진은 감사 이벤트 변환기로부터 입력된 이벤트들을 가지고 공격-이벤트 지식베이스에 대한 질의를 하고 최적화 탐색 기법인 유전자 알고리즘을 적용하여 입력된 감사 이벤트에 포함된 공격을 찾는다. 이때, 감사 이벤트로부터 공격을 탐지하면 탐지된 결과를 경고기로 보낸다. 끝으로, 경고기는 침입에 대한 적절한 조치를 취하고 탐지된 결과를 시스템 모니터로 보내는 역할을 담당한다.

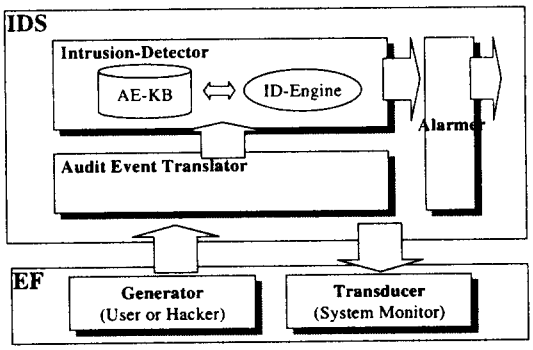


그림3. 침입 탐지 시스템의 시뮬레이션 모델

3.2 유전자 알고리즘을 적용한 침입 탐지

침입 탐지는 입력되는 사용자의 감사 이벤트에 포함된 이미 알려진 공격 패턴을 찾는 것으로 알려진 공격의 개수 Na가 많아지면 정확한 H 벡터 즉, Na차원의 가설벡터로서 가능한 특정 공격의 집합(만약 Hi=1이면 공격 i가 있음을 의미하고 Hi=0이면 공격 i가 없음을 의미한다)을 찾는 것은 NP-complete 문제로서 해결이 불가능하므로 탐색 및 최적화 기법인 유전자 알고리즘[7]을 적용하여 빠른 시간 안에 H 벡터를 찾을 수 있다. 세대를 구성하는 개체는 Na의 길이를 가지는 비트열로서 표현되며 개체들에 대하여 적용된 적합도 함수는 다음과 같다.

$$Fitness = \sum_{i=1}^{Na} Ri \cdot Ii \quad \cdots\cdots\cdots\cdots\cdots \text{식 (1)}$$

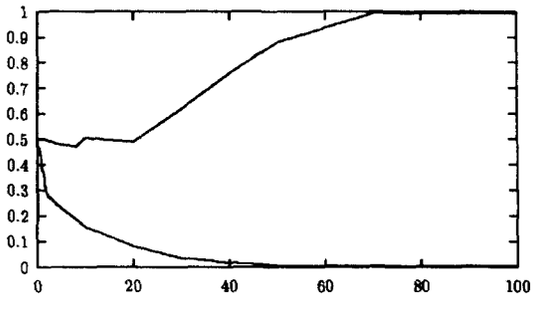
$$= \alpha + (\sum_{i=1}^{Na} Ri \cdot Ii - \beta \cdot Te^2) \quad \cdots \quad 식 (2)$$

R_i는 이미 알려진 공격i의 시스템에 대한 심각한 정도를 나타내고 I_i는 H_i로서 가설 벡터의 i번째 공격의 유무를 나타낸다. 따라서 적합도는 식(1)과 같이 가설에 포함된 공격과 그 공격의 심각한 정도와의 곱에 대한 합으로 나타낼 수 있다. 또한, 감사 이벤트 타입의 개수를 Ne, Ne × Na로서 각 공격에 의해서 생성되어지는 이벤트의 집합으로 구성되는 공격-이벤트 행렬을 AE, Ne차원의 입력되는 감사 벡터를 O(Oi는 감사 데이터에 포함된 i 타입의 이벤트 개수를 나타낸다)라고 할 때, $(AE \cdot H)_i > O_i$일 경우는 당연히 올바른 H일 수 없으므로 식(1)에 페널티를 적용하여 적합도를 조절한다. 이때 적용된 페널티 함수는 $P = Te^p$으로서 T_e는 $(AE \cdot H)_i > O_i$ 에 해당되는 이벤트 타입의 개수이며 p는 실험적으로 2정도가 적당하다. 따라서, 페널티 함수를 적용한 적합도 함수는 식(2)와 같다. 여기서 I는 개체를 의미하며 α는 적합도를 양수로 만들기 위한 값이고 β는 페널티 함수의 기울기를 조절한다. 또한 α와 β는 실험적으로 각각 200과 2를 사용하였다[8,9].

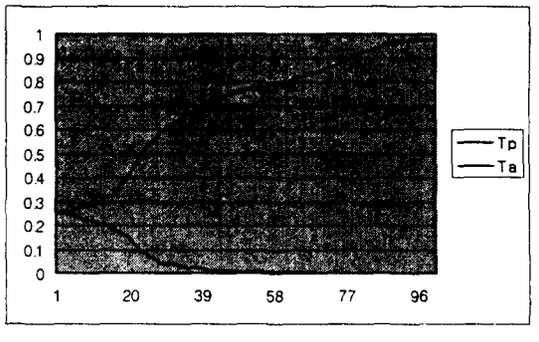
4. 시뮬레이션 테스트

본 논문에서는 침입 탐지 기법으로서 탐색 및 최적화 기법인 유전자 알고리즘을 적용하고 엔진 기반 모델링에 의하여 침입 탐지 시스템을 설계하였으며 RS6000의 AIX 시스템의 보안 감사 부시스템에서 제공하는 감사 이벤트를 가지고 시뮬레이션 하였다. 유전자 알고리즘은 [8,9]에서와 동일한 실험 조건으로서 500개체에 대하여 교차 연산 0.6, 돌연변이 연산 0.002의 확률을 적용하고 10회에 대한 평균을 취하여 결과를 비교함으로써 시뮬레이션 기법을 통한 성능평가의 타당성을 검증하였다.

그림4(a)와 (b)는 Tp와 Ta의 세대별 변화에 대한 실제 시스템에서의 테스트 결과[8,9]와 시뮬레이션 모델에서의 결과를 나타낸 것으로 Tp는 개체들 중에서 존재하는 공격에 대한 비트가 1로 설정된 개체의 수를 나타내고 Ta는 개체들 중에서 존재하지 않는 공격에 대한 비트가 1로 설정된 개체의 수를 나타낸다. 이때 Tp와 Ta가 각각 1과 0일 경우 모든 존재하는 공격과 존재하지 않는 공격을 탐지한 것으로서 100세대 후 그림4(a)의 경우 Tp와 Ta는 각각 0.996과 0.0044, 그림4(b)의 경우 Tp와 Ta는 각각 0.977과 0으로서 그림4(a)와 (b)는 유사한 분포를 나타내며 공격의 유무를 정확하게 탐지하는 것을 보이고 있다.



(a) 실제 시스템의 경우

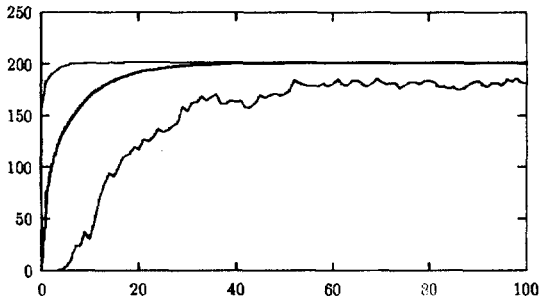


(b) 시뮬레이션 모델의 경우

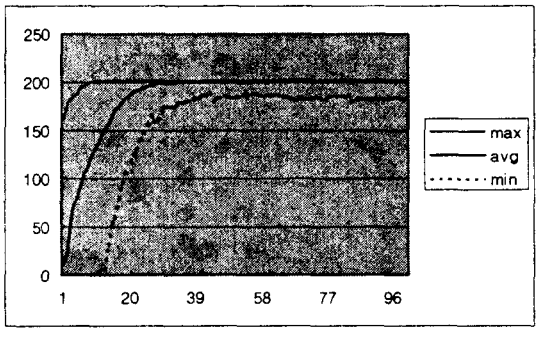
그림4. 세대별 Tp, Ta

또, 그림5(a)와 (b)는 세대별 적합도의 최대, 최소, 평균에 대한 실제 시스템에서의 테스트 결과[8,9]

와 시뮬레이션 모델에서의 결과를 나타낸 것으로 그림5(a)와 (b) 모두 동일한 분포로 적합도가 수렴되는 것을 보이고 있다.



(a) 실제 시스템의 경우



(b) 시뮬레이션 모델의 경우

그림 5. 세대별 적합도의 최대, 최소, 평균

5. 결론 및 향후연구

본 논문에서는 이산 사건 모델링 방법론에 의하여 침입 탐지 시스템의 모델을 설계하고 시뮬레이션 기법을 통하여 유전자 알고리즘을 적용한 침입 탐지 기법에 대한 성능평가를 수행하였다. 또, 시뮬레이션 기법을 통한 성능평가 결과와 실제 시스템 상에서의 성능평가 결과를 비교함으로써 시뮬레이션 기법을 통한 성능평가의 타당성을 검증하였다. 이를 통하여 실제 시스템의 구축과 적용에 따른 문제점을 해결하고 다양한 성능 평가의 기반을 제공할 수 있을 것으로 기대된다. 향후 연구로는 실질적인 침입 탐지 기법에 대한 연구와 다양한 침입 탐지 기법에 대한 시뮬레이션 기법을 통한 성능평가가 진행되어야 할 것이다.

참고문헌

[1] 이종성, 채수환, "분산 침입 탐지 에이전트를 기반으로 한 지능형 침입탐지시스템 설계," *한국정보처리학회 논문지*, 제6권 제5호, pp.1332-pp.1341, 1999.5

[2] H. Debar, M. Dacier, A. Wespi and S. Lampart, "An Experimentation Workbench for Intrusion Detection Systems", *Research Report*, IBM Research Division, March, 9. 1998

[3] *DARPA Intrusion Detection Evaluation*, http : //www.ll.mit.edu/SST/ideval/index.html.

[4] S.D. Chi, *Modelling and Simulation for High Autonomy Systems*, Ph.D.Dissertation, Univ. of Arizona, 1991.

[5] B.P. Zeigler, *Object-oriented Simulation with Hierarchical, Modular Models : intelligent Agents and Endomorphic Systems*, Academic press, 1990.

[6] 이종근, 지승도 "자치적 방어 시스템을 위한 모델 베이스기반 설계", *한국시뮬레이션 학회 논문지*, 제8권 제1호, 1999, 4.

[7] D.E. Goldberg, *Genetic Algorithm in Search, Optimization, and Machine Learning*, Addison-Wesley, 1989.

[8] Ludovic ME, "Genetic Algorithms, a Biologically Inspired Approach for Security Audit Trails Analysis", *the 1996 IEEE Symposium on Security and Privacy*, Oakland, May 1996.

[9] Ludovic ME, "GASSATA, a Genetic Algorithms, as an Alternative Tool for Security

Audit Trails Analysis", *First international workshop on the Recent advances in Intrusion Detection(RAID98).* September, 14-16, 1998., Louvain-la-Neuve, Belgium.