인공생명기반의 웜 바이러스 모델링 및 시뮬레이션 방법론

오지연*, 지승도*

Worm Virus Modeling and Simulation Methodology Using Artificial Life.

Ji-yeon Oh, Sung-do Chi

Abstract

Computer virus modeling and simulation research has been conducted with focus on the network vulnerability analysis. However, computer virus generally shows the biological virus characters such as proliferation, reproduction and evolution. Therefore it is necessary to research the computer virus modeling and simulation using Artificial Life. The approach of computer modeling and simulation using the Artificial Life technology provides the efficient analysis method for the effects on the network by computer virus and the behavioral mechanism of the computer virus. Hence this paper proposes the methodology of computer virus modeling and simulation using Artificial Life, which may be contribute the research on the computer virus vaccine.

Key Words: worm virus modeling, 인공생명, SES/MB

* 한국항공대학교 컴퓨터공학과

1. 서론

전 세계적인 정보화와 인터넷의 보급은 컴퓨팅 환경의 변화와 더불어 정보통신에 대한 의존도를 증대시켰다. 그 결과, 정보통신 기반구조에 대한 침해는 개별 서버 및 국가적, 경제적, 사회적 마비를 통하여 막대한 피해를 야기한다. 이는 정보통신 시스템 자체의 버그, 부적절한 구성 설정, 개방형 인터넷 기반구조 등에 따른 취약성을 이용한 해킹 뿐 아니라, 컴퓨터 바이러스에 대한 피해도 포함한다.[1]

컴퓨터 바이러스는 1970년대부터 제작되기 시작하였으며, 컴퓨터 바이러스 자신을 복제하기 위해서 다른 숙주 파일이나 부트 영역을 변경하는 악성 프로그램으로 정의한다.[2] 컴퓨터 바이러스에 감염된 숙주 객체는 컴퓨터 바이러스의 악성 코드를 완전한 복사본으로 포함하고 있기 때문에 숙주 파일이나 부트 영역이 실행되면 컴퓨터시스템에 존재하는 다른 객체는 컴퓨터 바이러스에 감염된다. 따라서 컴퓨터 바이러스는 프로그램을 감염시키고 복제하여 전파하는 특징을 갖는다. 이러한 특징은 성장, 증식, 진화하는 생물학적 바이러스와 유사하기 때문에 컴퓨터 바이러스를 인공생명체로 인식할 수 있다. 따라서 인공생명기반의 컴퓨터 바이러스 모델링을 통하여 컴퓨터 바이러스의 생명체적인 메커니즘을 분석함으로써 컴퓨터 바이러스의 감염 및 확산을 효과적으로 분석 할 수 있다. 이에 따라 외국에서는 Eugene 등을 중심으로 인공생명기반의 컴퓨터 바이러스에 관련된 연구가 진행 중이다.[3,4]

최근 컴퓨터 바이러스는 다른 프로그램을 감염시키지 않고 네트워크를 통해 바이러스 자신을 복제하여 전파하는 웜 바이러스의 형태로 발전하는 추세를 보인다.[5] 지난 2003년에 일어난 '1.25 인터넷 대란'은 전 세계적인 인터넷 마비사태로 웜 바이러스의 위험성을 극명히 보여주었다. 또한 웜 바이러스에 대한 연구의 필요성을 각인시켰다. 그러나 기존의 연구는 웜 바이러스가 네트워크 취약성에 미치는 영향에 대한 분석에 대한 것이었다.[6] 따라서 본 논문에서는 컴퓨터 바이러스 중 최근 그 피해의 심각성이 대두되고 있는 웜 바이러스의 시뮬레이션을 위해서 인공생명기반의 웜 바이러스 모델을 제시한다.

2. 인공생명기반의 웜 바이러스 모델링 및 시뮬레이션 접근방법

인공생명기반의 웜 바이러스의 모델링을 위한 접근 방법은 그림 1과 같다.

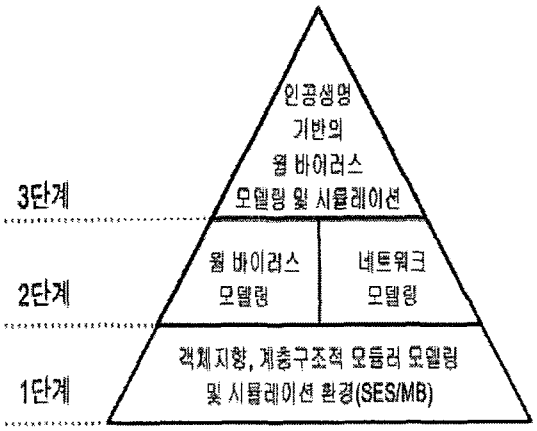

그림 1 계층적 접근 방법

인공생명 기반의 웜 바이러스를 모델링하기 위한 1단계는 객체 지향적이며 계층 구조적 모듈러 모델링 및 시뮬레이션 환경을 사용한다. 2단계에서는 1단계에서 선택한 환경을 기반으로 웜 바이러스를 모델링하고 웜 바이러스가 전파될 네트워크를 모델링한다. 2단계에서 만들어진 모델을

사용하여 3단계에서는 인공생명 기반의 웜 바이러스의 모델링 및 시뮬레이션을 수행한다.

1단계: 객체지향, 계층 구조적 모듈러 모델링 및 시뮬레이션 환경

본 논문에서는 인공생명 기반의 웜 바이러스의 모델링을 위해서 객체지향, 계층 구조적 모듈러 모델링 및 시뮬레이션 환경을 구축하기 위하여 객체지향 모듈러 모델의 개념을 반영한 SES/MB를 사용하였다.

객체지향 모듈러 모델: 일반적인 시스템 모델의 명세화와 객체지향 프로그래밍 개념은 내부적 상태 개념을 갖는다는 점에서 유사성을 갖는다. 그러나 복잡한 동역학에 의한 표현을 필요로 하는 네트워크 구성원 및 바이러스 모델링을 위해서는 객체지향 기법에 동역학 표현을 추가한 모듈화 모델을 필요로 한다. 모듈화 모델이란 시간 축을 기반으로 운영되며 입출력 포트를 통하여 외부환경과 상호작용을 하는 모델링의 개념을 의미한다. 이러한 모듈화 모델은 몇 개가 모여서 보다 큰 모델로 결합될 수 있는 계층 구조적 특성을 갖게 된다. 그러나 전형적인 객체지향 개념에서의 객체들은 동역학 표현을 위한 시간 축을 기반으로 하지 않고 계층 구조적 특성이 없다. 따라서 모듈화 모델링을 지원할 수 없다. 이러한 문제점을 해결하기 위해서 일반적 객체모델과는 달리 명시적인 입출력포트를 포함하는 모듈화 모델은 독립적 개발 및 테스트를 제공한다. 또한, 모델의 확장성 및 재사용성을 제공한다.

SES/MB(System Entity Structure): SES/MB [7,8,9]는 Zeigler에 의해 제안된 개념으로 기존의 동역학적 방법론과 AI의 기호적 방법론을 체계적으로 결합시킨 환경을 제공한다. SES/MB는 System Entity Structure와 Model Base의 두 구성원으로 이루어진다. SES는 시스템의 구조적 특성을 표현한다. 따라서 선언적 성격을 나타내며 구성관계, 구성원의 종류, 구성원들의 결합구조, 그리고 제약조건에 대한 구조적 지식을 표현할 수 있는 수단을 제공한다. MB는 시스템의 행위적 특성을 표현하고 절차적 성격을 나타내며 동역학적이고 상징적으로 행위를 표현할 수 있는 수단을 제공하는 모델들로 구성된다.

2단계: 웜 바이러스 모델링

웜 바이러스는 컴퓨터 바이러스가 발전한 종류로 자체 프로그램 코딩을 이용하여 전파되는 자기 복제가 가능한 악성코드로 정의한다.[2] 이러한 정의로 볼 때 웜 바이러스는 인공생명체로서의 특성을 가진다. 이러한 특성은 아래의 8가지로 구분된다.[3, 4]

- 생명의 시공성
- 자기복제
- 정보 저장소
- 신진대사
- 환경에 반응
- 안정성/복원력
- 진화하는 능력
- 성장과 확장

웜 바이러스는 컴퓨터에서 하나의 프로세스로 실행되는 동안 수행됨으로 생명의 시공성을 가진다. 또한 웜 바이러스는 자신의 코드를 복사해 전파함으로 자기 복제를 수행한다. 웜 바이러스가 수행되는 동안 컴퓨터에 웜 바이러스에 대한 정보를 남김으로 정보 저장소를 사용한

다. 웜 바이러스의 신진 대사는 웜 바이러스가 하나의 프로세스로서 실행되는 것으로 이해해야 한다. 이는 생명에 있어서 신진대사가 의미하는 것이 에너지를 사용하여 어떤 행위를 수행하는 것이기 때문이다. 웜 바이러스는 컴퓨터의 운영체제, 혹은 SQL 서버 프로그램과 같은 특정 조건이 만족하는 환경에서 실행됨으로 웜 바이러스는 환경에 반응한다. 웜 바이러스가 인공생명으로서 보이는 안정성과 복원력은 웜 바이러스가 이 기종의 컴퓨터 시스템 간에서도 전파되는 것으로 이해할 수 있다. 모든 웜 바이러스는 생명체가 진화하는 것처럼 능동적으로 진화하지는 않지만 특정한 환경아래서 진화하도록 프로그램 되어질 수 있다. 일반적으로 웜 바이러스는 자체코드를 통해서 E-메일을 이용해 전파되는데 이러한 과정은 인공생명으로서 웜 바이러스가 성장하고 확장되어지는 예이다. 또한 이는 웜 바이러스가 기존의 컴퓨터 바이러스와 다른 점이다. 이러한 웜 바이러스의 특징은 컴퓨터 바이러스가 숙주 객체를 필요로 하기 때문에 생기는 다른 프로그램에 대한 의존도를 낮춘다. 그리고 웜 바이러스의 자체코드는 E-메일에 첨부된 파일로 전파되기 때문에 컴퓨터시스템은 웜 바이러스를 정상적인 프로그램으로 인식하게 된다. 이러한 특성 때문에 웜 바이러스는 네트워크에 속한 노드 모델 상에서 동작하는 하나의 프로세스 모델로 표현 가능하다.

2단계: 네트워크 모델링

네트워크 모델링은 웜 바이러스가 출현하는 컴퓨터 네트워크에 대한 모델링으로 웜 바이러스가 컴퓨터 네트워크에 미치는 영향을 분석하는 역할을 한다. 네트워크 모델은 정보보호 관점에서 바이러스에 대한 시뮬레이션 접근을 위한 보안대책 및 취약성 분석을 위한 필수요소로 인식되고 있다.[10] 최근까지 네트워크 보안 모델링에 있어서 Cohen[11], Amoroso[12], NongYe[13] 등의 연구는 나름대로의 의미 있는 연구 결과들을 제시하고 있지만 대부분 개념적 단계의 모델링으로서 구체적인 분석이 어려운 실정이다. 따라서 웜 바이러스 파급에 따른 구체적인 변화를 분석하기 위하여, 컴퓨터 네트워크 상에 존재하는 호스트, 라우터, 방화벽, 침임 탐지 시스템 등과 같은 다양한 구성원들에 대한 모델링이 요구된다.

3단계: 인공생명을 이용한 웜 바이러스 모델링 및 시뮬레이션

웜 바이러스 모델링 및 시뮬레이션 방법론에 대한 개념도는 그림 2와 같다.

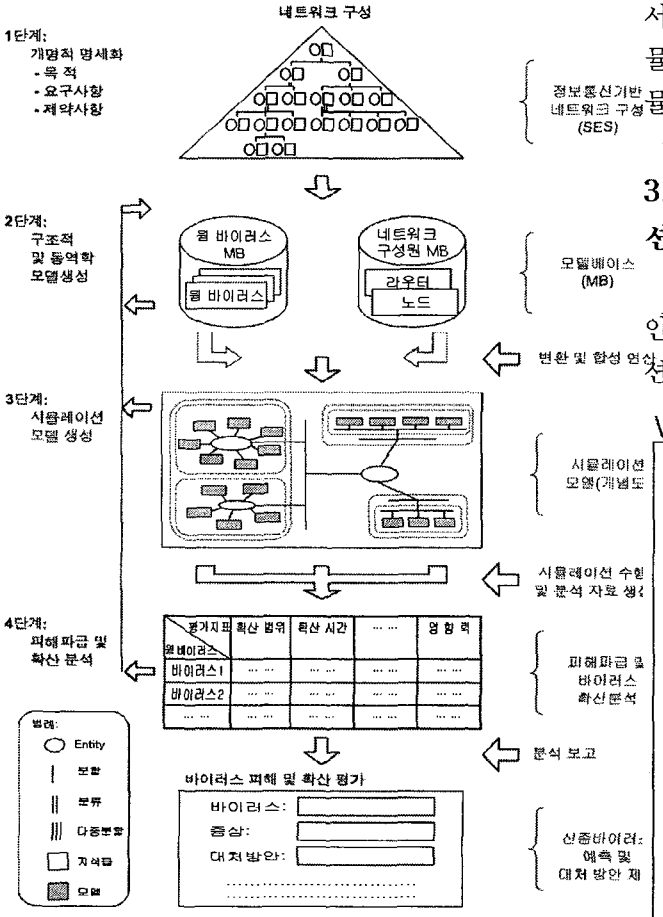

그림 2 인공생명기반의 웜 바이러스 모델링 및 시뮬레이션을 위한 접근방법

그림 2에서 1단계는 개념 명세화 단계로 정보통신 기반 네트워크의 전반적인 구조를 도식화하는 단계이다. 이 단계에서는 시스템의 구성관계, 구성원의 종류, 구성원들의 결합구조, 그리고 제약조건에 대한 구조적 지식의 표현수단을 제공한다. 2단계는 이미 라이브러리화되어 있는 네트워크 구성원 모델들과 인공생명을 이용한 각종 웜 바이러스 모델들을 1단계의 구조와 통합시키는 단계로서, 웜 바이러스 모델, 네트워크 구성원 모델을 포함한다. 이러한 구조적 및 동역학 모델들을 통합시킴으로 3단계에서는 최종적인 시뮬레이션 모델이 생성되어 시뮬레이션을 수행한다. 마지막 4단계에서는 시뮬레이션 수행 결과에 대해 분석한다.

3. 인공생명기반의 웜 바이러스 시뮬레이션 구조

인공생명기반의 웜 바이러스 모델링의 시뮬레이션을 위한 전체적인 구조는 그림 3과 같다.

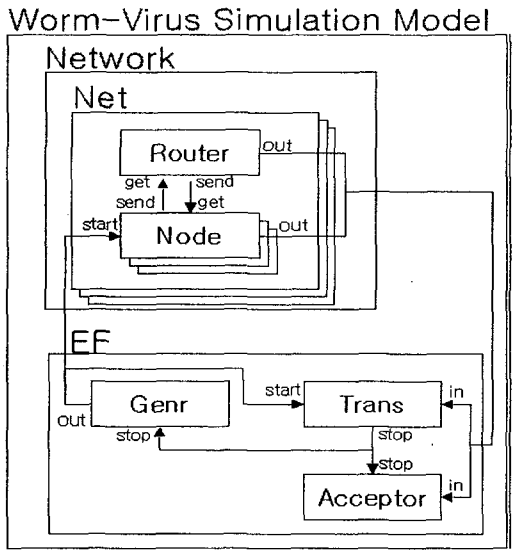

그림 3 인공생명기반의 웜 바이러스 모델링의 시뮬레이션을 우한 전체구조

Network 모델: Network 모델은 웜 바이러스가 발생하는 컴퓨터 네트워크를 표현하는 모델이다.
Net 모델: Net 모델은 네트워크에서 하나의 Router와 연결된 단말 노드들을 포함하는 단위망을 나타낸다.
Router 모델: Router 모델은 단위망에 속하는 Router에 대한 모델로 단위망에서 속하는 모든 단말 노드와 연결되어 있다. 그리고 하나의 단위망이 다른 단위망과 연결되어 있다는 것은 각 단위망에 속하는 Router 모델 객체간의 연결이 존

재하는 것을 의미한다. Router 모델은 네트워크의 단말노드에서 생성한 메시지를 전달한다. 이때, Router 모델은 자신과 연결된 어떤 노드나 라우터 동시에 여러 메시지를 받을 수 있기 때문에 버퍼를 사용한다.

Node 모델: Node 모델은 웜 바이러스가 실행되는 컴퓨터시스템을 나타내고 그림 4와 같이 Network Manager 모델, Scheduler 모델, Processor 모델로 구성된다.

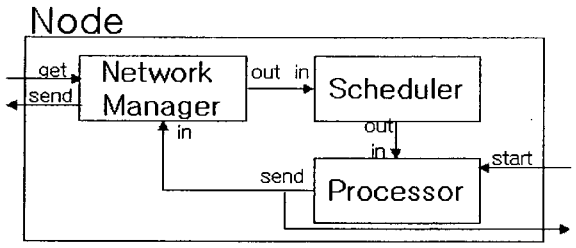

그림 4 Node 모델의 구조

Scheduler 모델: Scheduler 모델은 Network Manager 모델에서 받은 메시지를 확인하여 웜 바이러스를 실행하는 Processor 모델을 실행시킨다.

Network Manager 모델: 웜 바이러스는 E-메일을 통해서 전파되기 때문에 Network Manager 모델은 Node 모델에 전달된 메시지를 관리하는 모델이다. 표 1은 Network Manager 모델에 대한 수도코드이다. Node 모델은 외부로부터 웜 바이러스가 포함된 메시지를 받으면 메시지를 Scheduler 모델에 전달한다. 그리고 웜 바이러스를 포함하는 메시지를 다른 Node에 전파하기 위해서 Processor 모델에서 메시지를 Router 모델에 전달한다. Network Manager 모델에 전달된 E-메일 메시지는 Network Manager 모델이 다른 메시지를 처리 중이면 버퍼에 싸인다. 이때, 웜 바이러스의 공격으로 인해서 버퍼오버플로우가 발생하면 해당 Node의 상태는 'impossible'이 된다. 네트워크에 존재하는 모든

```
Network Manager
External Transition
    If Phase is 'passive'
        If receive value on port 'get'
            Hold-in 'getting'
        If receive value on port 'in'
            Hold-in 'sending'
    else
        If Buffer is not overflow
            Store message into the Buffer
        else
            Hold-in 'impossible'
Internal Transition
    If phase is 'getting'
        if the Buffer is Empty
            Hold-in 'passive'
        else
            Check the message
            Hold-in 'getting' or 'sending'
    If phase is 'sending'
        if the Buffer is Empty
            Hold-in 'passive'
        else
            Check the message
            Hold-in 'getting' or 'sending'
Output
    If phase is 'getting'
    Output: send the message to port 'out'
    If phase is 'sending'
    Output: send the message to port 'send'
```

표 1 Network Manager 모델의 수도코드

Node의 상태가 impossible'이 되면 네트워크가 완전히 마비된 것을 나타내고 시뮬레이션을 종료한다.

Processor 모델: 컴퓨터시스템이 웜 바이러스에 감염되었다는 것은 웜 바이러스 프로그램이 실행되었다는 것을 의미한다. 이를 표현하기 위해서 인공생명 기반의 웜 바이러스의 모델링에서는 Processor 모델을 사용한다. Processor 모델은 그림 5와 같이 구성된다.

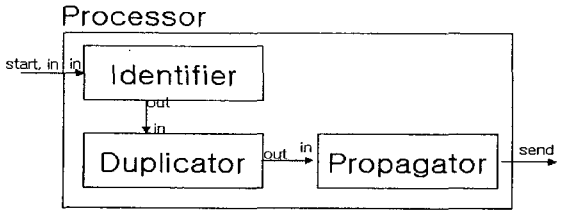

그림 5 Processor 모델의 구조

Identifier 모델: 웜 바이러스가 실행되기 위한 컴퓨터 시스템의 환경을 확인하는 모델이다. 웜 바이러스를 포함하여 컴퓨터 시스템 상에서 실행되는 모든 프로그램은 컴퓨팅 환경에 따라서 실행여부가 결정되기 때문이다. 표2는 Identifier 모델의 의사결정 코드이다.

Duplicator 모델: Duplicator 모델은 웜 바이러스를 포함하는 E-메일 메시지를 반복적으로 생성하는 모델이다. 새로 생성된 E-메일은 컴퓨터 시스템에 미리 저장되어 있는 E-메일 리스트에서 임의의 값을 읽어 웜 바이러스를 포함한 E-메일을 전달할 노드를 설정한다. 그리고 웜 바이러스 자신의 코드를 복사하여 첨부파일로 E-메일에 추가한다. 표3은 Duplicator 모델에 대한 의사결정 코드이다.

Propagator 모델: Duplicator 모델에서 생성된 E-메일 메시지에 대한 정보를 저장한 뒤 메시지를 Node 모델로 전달한다. 표4는 Propagator 모델의 수도코드이다.

```
Identifier
External Transition
    If Phase is 'passive'
        If receive value on port 'start'
            Hold-in 'checking'
        If receive value on port 'in'
            Hold-in 'checking'
    else
        Store message into the Buffer
Internal Transition
    If phase is 'checking'
        Hold-in 'passive'
Output
    If phase is 'checking'
        If the processing condition is fired
            Output: send the message to port 'out'
```

표 2 Identifier 모델의 수도코드

EF 모델: 인공생명기반의 웜 바이러스 모델을 이용하여 시뮬레이션을 하기 위해 필요한 모델로 시뮬레이션의 시작과 종료, 결과 분석을 위한 Gern 모델, Trans 모델, Acceptor 모델로 구성된다.

Gern 모델: 웜 바이러스를 포함하는 E-메일을 네트워크의 한 Node에 전달하는 역할을 수행한다.

Trans 모델: Trans 모델은 시뮬레이션이 종료조건을 검사하는 모델이다. Trans 모델은 웜 바이러스에 의해서 네트워크가 마비되거나 정의된 시간이 되면 시뮬레이션의 종료조건이 만족한 것으로 판단하여 인공생명기반의 웜 바이러스에 대한 시뮬레이션을 종료 시킨다.

```
Duplicator
External Transition
    If Phase is 'passive'
        If receive value on port 'in'
            Hold-in 'creating'
Internal Transition
    If phase is 'creating'
        Hold-in 'copying'
    If phase is 'copying'
        Hold-in 'ip setting'
    If phase is 'setting'
        Hold-in 'passive'
Output
    If phase is 'ip setting'
    Output: send the message to port 'out'
```

표 3 Duplicator 모델의 수도코드

```
Propagator
External Transition
    If Phase is 'passive'
        If receive value on port 'in'
            Hold-in 'writing'
Internal Transition
    If phase is 'writing'
        Hold-in 'passive'
Output
    If phase is 'writing'
    Output: send the message to port 'out'
```

표 4 Propagator 모델의 수도코드

Accrptor 모델: 시뮬레이션을 종료하는 메시지를 받고 시뮬레이션의 결과를 분석한다.

4. 향후연구

본 논문은 인공생명기반의 웜 바이러스에 대한 모델을 제시하였다. 따라서 향후 연구로는 본 논문에서 제시한 인공생명기반의 모델을 사용한 시뮬레이션을 수행함으로써 인공생명으로서 웜 바이러스에 대한 생명적인 특성에 대한 연구가 이루어져야 한다.

5. 결론

기존에 컴퓨터 바이러스 모델링에 관한 연구는 컴퓨터 바이러스자체에 대한 연구 보다는 네트워크에 관한 연구가 주를 이루었다. 하지만 본 논문에서는 인공생명의 관점에서 웜 바이러스에 대해 모델링함으로써 웜 바이러스가 나타내는 인공생명체적인 특성까지 모델링의 범주에 포함할 수 있다. 아직까지 다양한 시뮬레이션 분석이 진행되지는 않았지만, 향후 웜 바이러스가 네트워크에 미치는 영향뿐 아니라 웜 바이러스의 탐지 방법에 대한 연구에도 활용할 수 있을 것으로 기대된다. 또한 웜 바이러스의 동작 메커니즘에 대한 분석도 가능하게 되어 컴퓨터 바이러스 백신 연구에도 기여할 것으로 기대된다.

참고문헌

[1] T.A Longstaff, C.Chittister, R.Pethia, Y. Y.Haimes, "Are We Forgetting the Risks of I nformaion Technology", IEEE Computer, De c.2000.

[5] Grimes, Roger A., "Malicious Mobile Cod

e", O'REILLY, 2001.
[3] Eugene H, Spafford, "Computer Viruses as Artificial Lifr", Journal of Artificial Life, MIT Press, 1994
[4] Eugene H, Spafford, "Computer Viruses, A Form of Artificial Life", Technical Report CSD-TR-985, Purdue University, Sept. 1991
[5] http://www.kisa.or.kr : 한국정보보호 진흥원
[6] 유용준, 이장세, 지승도, "SIMVA를 이요한 시뮬레이션 기반의 네트워크 취약성 분석", 한국시뮬레이션학회 논문지, 13권 3호, pp21~29, 9월, 2004.
[7] Zeigler, B.P. Object-oriented Simulation with Hierarchical, Modular Models: Intelligent Agents and Endomorphic systems, Academic Press, San Diego, CA.[11], 1990.
[8] Zeigler, B.P. Kim, T.G. and Praehofer, H. Theory of Modeling and Simulation. 2 ed. Academic Press, New York, NY. 1990.
[9] S.D. Chi, Modeling and Simulation for High Autonomy Systems, Ph.D. Dissertation, Dept. of Electrical and Computer Engineering, Univ. of Arizona, 1991.
[10]이철원, 김홍근, "정보보증:컴퓨터보안의 새로운 패러다임", 정보과학회지, 제18권 제1호, pp53~61, 1월, 2000.
[11] Fred Cohen, "simulating Cyber Attacks Defenses, and Consequences". 1999 IEEE Symposium on Security and Privacy Special 20th Anniversary Program, The Claremont Resort Berkeley, California, May 9-12. 1999
[12] Amoroso, E., Intrusion Detection, AT&T Laboratory, Intrusion Net Books, January, 1999
[13] Nong Ye, Joseph Giordano, CACS - A Process Control Approach to Cyber Attack Detection, Communications of the ACM.