情報保護學會論文誌
第18卷 第1號, 2008. 2

안전한 홈네트워크 서비스를 위한 계층적 분산 침입탐지에 관한 연구*

유 재 학[1†], 정 용 화[1‡], 최 성 백[2], 양 성 현[3], 박 대 희[1]
[1]고려대학교, [2]조은시큐리티, [3]광운대학교

A Study on Hierarchical Distributed Intrusion Detection for Secure Home Networks Service*

Jaehak Yu[1†], Yongwha Chung[1‡], Sungback Choi[2], Sunghyun Yang[3], Daihee Park[1]
[1]Korea University, [2]Joeunsecurity, [3]Kwangwoon University

요 약

본 논문에서는 구조적으로 분산 침입탐지시스템의 구조를 계승하면서 동시에 홈네트워크의 환경을 최대한 고려하여 *HNHDIDS*(Home Network Hierarchical Distributed Intrusion Detection System)로 명명된 새로운 계층적 분산 침입탐지 시스템을 제안한다. 제안된 시스템은 단일 클래스 support vector machine(support vector data description)과 지역적 에이전트(agent)들을 계층적으로 결합한 구조로써, 홈네트워크의 환경을 위하여 최적화되었다. 만족스러운 침입 탐지율과 안전한 FNR(false negative rate) 수치 등을 실험을 통하여 확인함으로써 제안된 시스템이 홈네트워크 환경에 적합함을 검증하였다.

ABSTRACT

In this paper, we propose a novel hierarchical distributed intrusion detection system, named *HNHDIDS*(Home Network Hierarchical Distributed Intrusion Detection System), which is not only based on the structure of distributed intrusion detection system, but also fully consider the environment of secure home networks service. The proposed system is hierarchically composed of the one-class support vector machine(support vector data description) and local agents, in which it is designed for optimizing for the environment of secure home networks service. We support our findings with computer experiments and analysis.

Keywords : *Intrusion detection, Home network security, Support vector machine*

Ⅰ. 서 론

홈네트워크는 유선(wired)과 무선(wireless)의 통합 네트워크를 기반으로 정보 인프라, 홈오토메이션, 그리고 고품질 인터넷 서비스 등의 다양하고 편리한 서비스를 제공하여 사용자의 편의를 극대화시키기 위한 기술 집약적인 산업으로 성장하고 있다.

그러나 홈네트워크의 댁내망은 PC뿐만 아니라 홈오토메이션, 백색 가전 등 다양한 기기들로 구축되어 있기에 외부의 공격 형태 및 피해 양상은 다양하게 나타날 수 있다. 안전한 홈네트워크을 위한 요구 사항으로는 사생활 보호를 위한 보안 및 홈기기의 안전성 확보 등을

접수일: 2007년 6월 29일; 채택일: 2007년 11월 22일
* 본 연구는 산업자원부 및 한국산업기술평가원의 성장동력 기술개발사업의 연구결과로 수행되었습니다.
† 주저자, dbzzang@korea.ac.kr
‡ 교신저자, ychungy@korea.ac.kr

들 수 있다. 특히 많은 갱신 정보들이 디지털화되어 인터넷에 노출됨으로써, 홈네트워크의 가정 내 시설 및 통신 등을 위한 침입탐지시스템을 포함한 보안기술의 필요성이 학계의 중요한 이슈이다[1-2].

침입탐지 방법론은 침입에 대한 탐지 전략에 따라 크게 오용 탐지(misuse detection) 모델과 비정상 탐지(anomaly detection) 모델로 나누어진다[3-4]. 오용 탐지모델은 이미 발견된 공격 유형에 대한 면밀한 분석을 통하여 규칙 베이스(rule-base)화 하고 이를 기반으로 탐지를 수행하는 방법으로, 새로운 공격 유형이 발견될 시에는 수동으로 규칙 베이스를 갱신해야만 새로운 공격에 대처할 수 있다는 문제점을 가지고 있다. 비정상 탐지 모델은 미리 정의된 정상 행동에 대한 프로파일로부터 크게 벗어나는 데이터를 비정상 행동으로 판단하여 공격을 탐지하는 방법으로, 새로운 공격유형을 탐지할 수 있다는 점에서는 실용적이나 탐지된 공격 유형에 대한 추가적인 세부 정보를 알 수 없기에 침입에 따른 적절한 대처를 할 수 없다는 한계점을 피할 수 없다. 최근의 연구문헌 조사에 의하면, 보다 지능적인 침입탐지 모델의 설계를 위하여 데이터마이닝 및 기계학습 기법을 침입탐지시스템에 적용하려는 시도가 활발히 진행중이다. 이러한 연구 동향 중 특히 패턴 분류(pattern classification) 및 함수 근사(function approximation) 등의 문제에서 매우 우수한 성능을 보이는 SVM(support vector machine)을 침입탐지에 적용하려는 연구가 주목을 받고 있다[3-5].

이상에서 살펴본 기존의 침입탐지 모델들은 다양한 도메인에서 적절히 적용되어 왔으나 홈네트워크 환경으로의 적용 시에는 여러 가지 문제점들이 발생한다. 특히, 기존의 중앙 집중 형태의 처리방식은 침입에 대한 정확한 탐지를 못했을 경우 전체 시스템의 마비를 초래할 수 있을 뿐만 아니라 시스템의 갱신 및 새로운 기능의 추가 시 상당한 비용을 초래해야만 한다.

따라서 본 논문에서는 홈네트워크의 환경을 최대한 고려하여 *HNHDIDS*(Home Network Hierarchical Distributed Intrusion Detection System)로 명명된 새로운 계층적 분산 침입탐지 시스템을 제안한다. 본 시스템은 단일 클래스 SVM과 지역적 에이전트(agent)들을 계층적으로 결합한 구조로써, 다음의 평가 기준들을 모두 만족하는 차원에서 설계되었다: 1) 실시간 침입탐지 및 서비스의 안전성 보장; 2) 지역화된 에이전트의 역할로 침입에 의한 피해 발생 시 시스템 피해의 최소화; 3) 시스템에서 학습되지 않은 새로운 공격 유형의 탐지; 4) 빠르고 효율적인 학습 및 점증적(incremental) 갱신으로 인한 경제적인 시스템의 유지/보수 및 확장성; 5) 저가의 시스템 구축 및 경량적 시스템.

본 논문의 구성은 다음과 같다. 2장에서는 본 논문의 연구배경 및 홈네트워크의 보안 요구사항에 대해 설명하고, 3장에서는 본 논문에서 제안하는 계층적 분산 침입탐지 모델에 대해 기술한다. 4장에서는 실험결과 및 성능 분석을 기술하며, 마지막으로 5장에서는 결론 및 향후 연구과제에 대해 논한다.

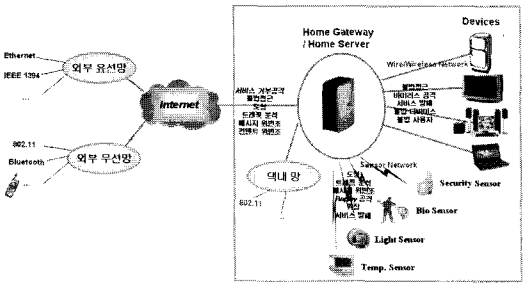
Ⅱ. 연구배경

2.1. 기존 침입탐지방법론의 문제점

기존의 침입탐지방법론들을 홈네트워크 환경에 그대로 적용할 경우 발생하는 문제점들을 나열하면 다음과 같다[2]: 1) 네트워크 및 호스트 기반 침입탐지 시스템은 모놀리식 구조(monolithic architecture)를 이용한 중앙 집중 형태의 데이터 수집과 분석을 수행함으로써, 침입을 정확히 탐지 못했을 경우 전체 시스템의 마비 또는 홈기기들의 오작동을 유발하는 등 심각한 문제점을 야기할 수 있다. 2) 시스템을 변경하거나 새로운 기능의 추가 시 시스템의 인스톨 및 재부팅이 요구된다. 따라서 실시간 침입탐지를 보장할 수 없다. 3) 홈게이트웨이에서 모든 정보의 모니터링 및 프로세싱이 진행됨으로, 처리절차와 시간이 오래 걸린다. 결과적으로 중앙 집중식 침입탐지 시스템의 한계를 극복할 수 있는 새로운 대안이 요구된다.

2.2. 홈네트워크 환경에서의 보안 요구사항

홈네트워크 환경을 고려한 시스템 요구사항은 다음과 같이 요약된다[2]: 1) 대부분의 홈네트워크 사용자는 IT비전문가임으로 시스템 갱신을 포함한 많은 부분이 자동화 되어야한다. 2) 훈련된 보안 관리자가 운영하는 것이 아니므로 사용이 간편하고 정책설정이 용이한 사용자 인터페이스를 제공하여야한다. 3) 홈네트워크의 구축비용을 감안할 때 저가로 시스템을 구축할 수 있어야 하며 시스템이 가볍게 동작해야한다. 4) 홈네트워크의 침입이 홈기기의 동작에 영향을 미치기 전에 처리할 수 있도록 시스템의 실시간성이 보장되어야한다. 5) 기



〔그림 1〕 홈네트워크에서의 접근경로 및 보안 취약점

존의 보안 제품과 달리 홈네트워크 기기는 각 가정에 보급되는 것을 목표로 하기 때문에 대량공급 및 운영의 편의성이 보장되어야한다.

[그림 1]은 홈네트워크 환경에서의 사용자 접근경로 및 그에 따른 보안 취약성을 보여주고 있다[2]. 홈네트워크에서는 다양한 홈기기를 댁외나 댁내에서 접근하여 서비스를 이용하기 때문에 중간 매개체인 홈게이트웨이에 대한 보안 취약성 분석이 우선적으로 요구되며, 또한 홈네트워크 환경에서의 접근경로인 유·무선 네트워크 기술에 대한 보안 취약성 분석도 필요하다[1-2].

2.2.1. 홈게이트웨이의 보안 취약점 분석

홈게이트웨이는 외부망과 댁내망의 연결, 댁내망 내의 홈기기들 사이의 인터페이스를 담당한다. 특히 외부망으로부터 요구되는 다양한 서비스를 안전하게 제공해야 하며 해킹, 악성코드, 웜 및 바이러스, DoS, 도·감청 등의 공격들을 고려해야 한다[1]. 아래의 [표 1]은 홈게이트웨이에서 고려해야할 보안 취약점을 정리한 내용이다.

〔표 1〕 홈게이트웨이의 보안 취약점 분석

항목	보안 취약성
인증	-인가 사용자로 위장하여 접근
접근제어	-비인가 사용자의 불법 접근 -정상 사용자에 대한 접근 차단
기밀성/무결성	-사용자 정보의 유출 -홈기기 제어 데이터의 유출 및 위·변조
DoS 공격	장비사용 불능

〔표 2〕 유·무선 홈네트워크 기술의 보안 취약점 분석

기술		보안 취약성
유선	HomePNA	-데이터 변조 및 DoS공격
	PLC	-데이터 유출 및 위변조
	IEEE1394	-송·수신 데이터 유출
무선	Bluetooth	-블루스니핑, 블루버깅, 블루재킹
	무선랜	-비인가 사용자의 불법 접근
	Zigbee	-비인가 사용자의 불법 접근

2.2.2. 유·무선 네트워크의 보안 취약점 분석

홈네트워크는 유·무선 네트워크 기술을 이용하기 때문에 이에 대한 보안 취약성 분석이 필요하다[1]. [표 2]는 홈네트워크 환경에 적용 가능한 유·무선 네트워크 기술들의 대표적 보안 취약점을 정리한 내용이다.

Ⅲ. *HNHDIDS*

3.1. 분산 침입탐지 시스템

중앙 집중식 침입탐지 시스템의 한계를 극복할 수 있는 새로운 대안으로써 분산 침입탐지에 관한 연구가 최근 학계에서 관심의 대상이다[6-8]. 분산 침입탐지시스템의 특징들을 요약하면 다음과 같다: 1) *지역적 침입탐지*: 각각의 침입탐지 에이전트는 클러스터 헤드(cluster head)에 연결된 클러스터 멤버(cluster member)들의 현재 상태 및 정보들을 수집하여 지역적이고 독립적으로 침입탐지를 수행한다. 2) *자율성 보장*: 에이전트는 미리 정의된 임계값으로 정상작동과 오작동 그리고 새로운 침입에 대한 자율적 탐지를 수행한다. 3) *분산된 작업과 처리 효율성*: 클러스터 멤버에서 침입이 탐지되면 해당 에이전트에 포함되어 있는 지역만을 차단함으로써 피해를 최소화 한다. 4) *경제적인 시스템의 유지보수*: 새로운 클러스터 멤버의 추가 그리고 시스템의 확장 및 변경은 해당 에이전트만을 갱신하기 때문에 경제적인 시스템의 유지보수가 가능하다.

따라서 본 논문에서는 기본적으로 분산 침입탐지시스템의 구조를 계승하면서 동시에 홈네트워크 환경을 최대한 고려하여 시스템을 설계하였다.

3.2. 단일 클래스 SVM 및 *HNHDIDS*

본 장에서는 우선 침입탐지를 위한 단일 클래스 SVM을 소개하고, 이를 주요 구성 요소로 하는 새로운 계층적 분산 침입탐지 모델인 *HNHDIDS*(Home Network Hierarchical Distributed Intrusion Detection System)을 자세히 설명한다.

3.2.1. 단일 클래스 SVM

SVM(support vector machine)은 주어진 문제의 전역적 최적해(global optimum solution)를 보장함으로 패턴 분류 및 함수 근사 등에 적용되어 매우 우수한 성능을 보이고 있다. 특히, 침입탐지의 경우 대부분의 데이터는 정상 그리고 일부의 데이터만이 공격데이터로써 학습에 사용가능한 데이터의 크기는 차이가 크다. 따라서 이진 분류기인 SVM은 관측되지 않은 영역을 포함하여 결정 경계면을 생성함으로 새로운 학습 데이터에 대해서 오분류(misclassification)할 가능성이 크다. 그러므로 해당 클래스만을 독립적으로 표현하는 단일 클래스 분류기(one-class SVM)로서 결정 경계면을 선택하는 것이 보다 유리하다. 따라서 본 논문에서는 단일 클래스 SVM의 대표적인 알고리즘인 SVDD(support vector data description)[3]를 기반으로 한다.

d-차원 입력공간상에 존재하는 K-개의 데이터의 집합 $D_k = \{x_i^k \in R^d | i = 1,\ldots,N_k;\ k = 1,\ldots,K\}$ 이 주어졌을 경우, 각각의 클래스를 분류하기 위한 분류기는 각 클래스의 학습 데이터를 포함하면서 체적을 최소화하는 구체(sphere)를 구하는 문제로 정의되며, 다음의 최적화 문제를 통하여 수식화 된다.

$$\begin{aligned} &\min L_0(R_k^2, a_k, \xi_k) = R_k^2 + C\sum_{i=1}^{N_k}\xi_i^k \\ &s.t \quad \|x_i^k - a_k\|^2 \le R_k^2 + \xi_i^k,\ \xi_i^k \ge 0,\ \forall i. \end{aligned} \tag{1}$$

여기에서, a_k는 k-번째 클래스를 표현하는 구체의 중심이며, R_k^2은 구체의 반경의 제곱, ξ_i^k는 k-번째의 클래스에 속한 i-번째 학습 데이터 x_i^k가 구체에서 벗어나는 정도를 나타내는 벌점 항이며, C는 상대적 중요성을 조정하는 상수(trade-off constant)이다.

식 (1)에 관한 쌍대 문제(dual problem)를 구하기 위하여 라그랑제함수(Lagrange function) L을 도입한다.

$$\begin{aligned} &L(R_k^2, a_k, \boldsymbol{\xi_k}, \boldsymbol{\alpha_k}, \boldsymbol{\eta_k}) \\ &= R_k^2 + C\sum_{i=1}^{N_k}\xi_i^k \\ &\quad + \sum_{i=1}^{N_k}\alpha_i^k[(x_i^k - a_k)^T(x_i^k - a_k) - R_k^2 - \xi_i^k] \\ &\quad - \sum_{i=1}^{N_k}\eta_i^k\xi_i^k \end{aligned} \tag{2}$$

단, $\alpha_i^k \ge 0$, $\eta_i^k \ge 0$, $\forall i$.

식 (2)는 변수 $R_k^2, a_k, \boldsymbol{\xi_k}$에 대해서는 최소값을 변수 $\boldsymbol{\alpha_k}, \boldsymbol{\eta_k}$에 대해서는 최대값을 가져야 함으로, 아래의 조건식을 만족해야 한다.

$$\begin{aligned} \frac{\partial L}{\partial R_k^2} &= 0: \sum_{i=1}^{N_k} a_i^k = 1. \\ \frac{\partial L}{\partial \xi_i^k} &= 0: C - \alpha_i^k - \eta_i^k = 0 \quad \therefore \alpha_i^k \in [0, C],\ \forall i. \\ \frac{\partial L}{\partial a_k} &= 0: a_k = \sum_{i=1}^{N_k}\alpha_i^k x_i^k \end{aligned} \tag{3}$$

조건 식 (3)을 라그랑제 함수 L에 대입하면, 다음의 쌍대 문제를 얻는다.

$$\begin{aligned} &\min_{\alpha_k} \sum_{i=1}^{N_k}\sum_{j=1}^{N_k}\alpha_i^k\alpha_j^k < x_i^k, x_j^k > - \sum_{i=1}^{N_k}\alpha_i^k < x_i^k, x_i^k > \\ &s.t. \quad \sum_{i=1}^{N_k}\alpha_i^k = 1,\ \alpha_i^k \in [0, C],\ \forall_i \end{aligned} \tag{4}$$

입력 공간위에서 정의되는 구체는 매우 간단한 형태의 영역만을 나타낼 수 있다. 이러한 한계를 극복하기 위하여 커널 함수(kernel fuction) k를 통하여 정의되는 고차원의 특징 공간(feature space) F 위에서 정의되는 구체를 사용하는 방향으로 확장될 수 있다. 각각의 클래스는 각자의 특징공간에서 자신의 경계를 보다 정확하게 표현할 수 있으므로, 시스템의 학습은 각각의 클래스들이 매핑되는 특징공간의 독립성을 고려하여 아래의 convex QP(quadratic problem) 문제의 해답을 얻음으로써 이루어진다.

$$\begin{aligned} &\min_{\alpha_k} \sum_{i=1}^{N_k}\sum_{j=1}^{N_k}\alpha_i^k\alpha_j^k k_k(x_i^k, x_j^k) - \sum_{i=1}^{N_k}\alpha_i^k k_k(x_i^k, x_i^k) \\ &s.t. \quad \sum_{i=1}^{N_k}\alpha_i^k = 1,\ \alpha_i^k \in [0, C],\ \forall_i \end{aligned} \tag{5}$$

특히 가우시안 커널(Gaussian kernel)을 사용할 경우,

$k(x,x)=1$이 성립함으로 식 (5)는 아래와 같이 단순화 된다.

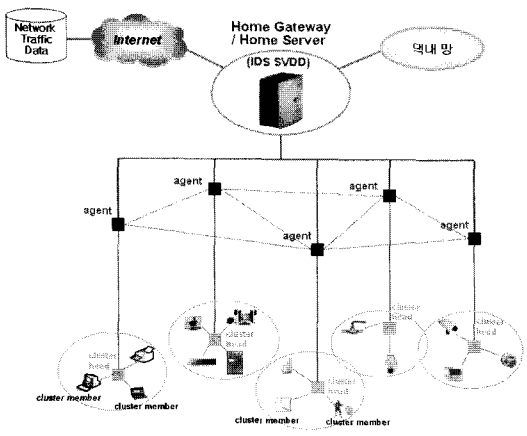
$$\min_{a_k} \sum_{i=1}^{N_k}\sum_{j=1}^{N_k} \alpha_i^k \alpha_j^k k_k(x_i^k, x_j^k) \quad s.t. \quad \sum_{i=1}^{N_k} \alpha_i^k = 1,\ \alpha_i^k \in [0, C],\ \forall_i. \tag{6}$$

학습 종료 후 적용 과정에서, 각각 클래스의 결정함수는 다음과 같이 정의 된다.

$$f_k(x) = R_k^2 - \left[1 - 2\sum_{i=1}^{N_k} \alpha_i^k k_k(x_i^k, x) + \sum_{i=1}^{N_k}\sum_{j=1}^{N_k} \alpha_i^k \alpha_j^k k_k(x_i^k, x_j^k)\right] \geq 0 \tag{7}$$

3.2.2. *HNHDIDS*(Home Network Hierarchical Distributed Intrusion Detection System)

본 장에서는 유·무선 네트워크 환경에서의 분산 침입탐지 모델을 기반으로 홈네트워크 환경에 적합한 *HNHDIDS*(Home Network Hierarchical Distributed Intrusion Detection System)로 명명된 새로운 침입탐지 모델을 제안한다([그림 2] 참조). Peng[9] 등이 제안한 계층적 분산 침입탐지 모델은 MANET(Mobile Adhoc Network)에 근간을 이루고 있는 바, 이는 홈네트워크 환경에 그대로 적용하기에 어려움이 있다. 이 방법은 지역적 침입탐지와 실시간적 탐지 보장 그리고 계층적이며 분산처리에 의한 효율성 등의 장점을 가지고는 있지만, MANET의 기본적 특성인 망 이동성 보장으로 인하여 시스템의 구조 변경이 가능하다. 또한 동일 계층 간의 접근 및 정보교환 등의 특성들은 홈네트워크 환경에는 적절치 않다.



[그림 2] *HNHDIDS*(Home Network Hierarchical Distributed Intrusion Detection System)의 구조도

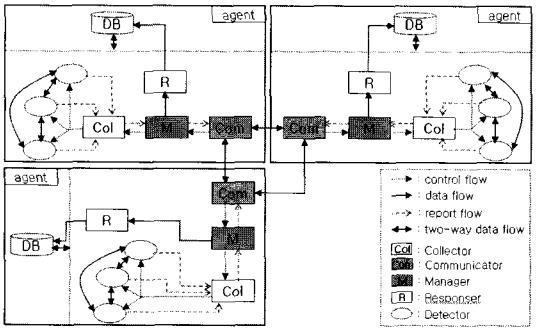
따라서 본 논문에서는 Peng[9] 등이 제안한 모델의 장점은 유지하되 문제점인 망의 이동 및 자율적 그룹화, 동일 계층 간 불필요한 접근 및 정보교환 등의 단점은 보완하는 견지에서 홈네트워크 환경에 적합한 새로운 계층적 분산 침입탐지 모델을 설계하였다. 특히 Peng[9] 등이 제안한 모델에서 사용되는 각 에이전트의 침입탐지 구조는 홈네트워크 환경에서는 과도한 하드웨어를 요구함으로 본 모델에서는 경제적 시스템의 구축을 고려하여 삭제하였다. 새롭게 제안된 모델(*HNHDIDS*)의 각 계층별 기능은 다음과 같다:

첫 번째 계층의 홈게이트웨이에서는 외부망과 댁내 망의 트래픽 정보 중 정상 데이터만으로 학습된 경량적 실시간 탐지를 보장하고 점증적 갱신이 가능한[4] SVDD의 사용으로 침입을 실시간으로 탐지 및 차단한다. 즉, 침입 데이터로 판단된 경우는 접근을 원천 봉쇄한다. 또한 홈게이트웨이는 홈기기들을 위한 인터페이스를 제공함과 동시에 시스템의 전체 상태 모니터링 및 제어를 담당한다.

하위계층의 각 에이전트들은 추가적인 침입탐지기능은 수행치 않고 클러스터 헤드별로 그룹화 된 홈기기들의 상태 모니터링 및 보고 그리고 관리를 수행하는 역할을 담당한다. 단 침입에 의한 이상 현상이 감지되면 주변 에이전트에게 경보를 신속히 전달하고 모든 연결고리를 끊음으로써 피해를 최소화한다, 또한 감지된 이상 현상을 홈게이트웨이에 보고하여 재인증 및 접근제어 등의 추후 대응방법을 결정할 수 있게 도와준다.

마지막 계층으로 클러스터 멤버들은 실제 홈네트워크를 구성하는 홈기기들을 말하며 각 장치들의 상태 및 정보는 자신을 포함하는 클러스터 헤드에게 보고한다.

MANET의 특성인 망 이동성과 동일 계층 간 접근에 관한 문제를 극복하기 위하여 클러스터 헤드에 묶인 홈기기들의 연결을 임의로 고정하였으며, 하나의 클러스터 헤드에 연결된 홈기기들의 자율적 그룹화 과정은 차단하였다. 또한, 하위 계층에 구성된 홈기기들은 서로의 연결 경로를 통해 다른 장치들과 통신할 수 없으며 독립적으로 작동하도록 정의하였다.



〔그림 3〕. 에이전트 기반 분산 시스템의 구조도

[그림 3]은 본 시스템의 에이전트에 관한 기능설명도로써 [6]에서 소개된 침입탐지 에이전트를 홈네트워크 환경에 맞게 일부 변형하였다. 각 요소 중 가장 기본이 되는 detector는 클러스터 멤버에 해당되며 독립적으로 수행하나 클러스터 헤드에 포함된 detector들은 상호 협력하여 침입 및 시스템의 이상여부를 감지한다. 이러한 detector들은 침입 또는 홈기기의 이상여부를 탐지하면 즉시 collector에게 현재 상태를 보고한다. collector는 클러스터 헤드로써 소속된 detector들의 제어 및 관리를 담당하며 detector들의 현재 상태를 manager에게 보고하는 역할을 한다. manager는 collector로부터 침입 또는 오작동이 확인되면 실시간으로 홈게이트웨이와 responser 모듈로 보고하여 재인증 및 접근제어 등의 대응방법을 결정하게 도와주며 침입에 대한 피해를 최소화 하는 역할을 담당한다. communicator는 주변 에이전트간의 상태 및 정보 전달을 담당한다.

Ⅳ. 실험 결과 및 성능 분석

본 실험에서 사용한 KDD CUP 99는 1998년 DARPA Intrusion Detection Evaluation Program에서 침입탐지 분야의 표준 데이터 집합을 얻기 위하여 미국의 군사 네트워크상에서 시뮬레이션을 통해 만든 데이터로써[10], 본 실험에서는 실험 결과의 정확한 분석을 위하여 KDD CUP 99 데이터 중 corrected labeled된 데이터 집합만을 이용하였다. 사용된 데이터의 수는 전체 311,029개 중 정상데이터 50,000개, 공격데이터는 DoS와 R2L 그리고 probing은 각각 100개, U2R은 labeled된 88개 모두를 포함하여 388개 데이터를 랜덤하게 추출하였다. KDD CUP 99 데이터의 속성은 9개의 기호형(symbolic) 속성과 32개의 숫자형(numeric) 속성으로 구성되어있다. 데이터의 모든 속성을 사용하기 위하여 기호형 속성에 대하여 다음과 같은 변환을 수행하였다. Σ는 한 속성에 나타날 수 있는 기호형 데이터 값들의 집합이며, $|\Sigma|$는 집합의 구성요소의 수를 의미한다. 한 기호형 데이터의 값이 i 번째 구성요소라고 가정하면 해당 기호형 데이터의 값은 아래의 방법과 같이 변환된다[3].

$$\underbrace{0\ 0\ 0\ 0\ 0\ \ 1/|\Sigma_i|\ldots..\ 0\ 0\ 0}_{|\Sigma|}$$

또한 DoS 공격을 분산화, 자동화시켜 더욱 발전시킨 DDoS(Distributed Denial of Service) 공격의 탐지실험을 위하여 DDoS의 대표적 툴인 TFN2K[11]을 이용하여 정상 데이터 50,000개와 공격 데이터 400개를 생성하여 실험하였다.

성능 측정을 위하여 침입 탐지율(detection rate), FPR(false positive rate) 및 FNR(false negative rate)을 성능지표로 사용했으며 실험결과는 [표 3]에 정리하였다. 여기서 조정상수 C는 0.1, 커널 함수인 가우시안 함수의 변수인 σ 값은 0.7로 고정하였다.

$$\text{침입탐지율} = \frac{\sum_{i=1}^{n} T_i}{\sum_{i=1}^{n} I_i} \tag{8}$$

$$FPR = \frac{\sum_{i=1}^{n} P_i}{\sum_{i=1}^{n} N_i} \tag{9}$$

$$FNR = \frac{\sum_{i=1}^{n} F_i}{\sum_{i=1}^{n} I_i} \tag{10}$$

위 식에서 I는 공격데이터, T는 공격데이터를 정확히 공격으로 분류한 데이터, N은 정상데이터, P는 정상데이터를 공격으로 분류한 데이터, F는 공격데이터를 정상으로 판단한 데이터를 의미한다.

KDD CUP 실험의 경우, 정상 데이터 45,000개로 학습하여 98.1%의 학습율을 얻었으며, 테스트 시에는 공격데이터 388개와 학습에 참여하지 않은 정상데이터 5,000개를 사용하였다. 또한 DDoS 실험의 경우, 45,000개의 정상데이터로 학습하여 97.89%의 학습율

〔표 3〕 성능 평가를 위한 측정 표

항목 / data set	침입 탐지율	FPR	FNR
KDD CUP	96.65	14.95	0.52
DDoS	96.25	11.75	0.50

을 얻었으며, 테스트 시에는 400개의 공격데이터와 학습에 참여하지 않은 5,000개의 정상데이터로 테스트하였다.

[표 3]의 성능 평가를 위한 항목 중 FPR은 정상 데이터를 공격데이터로 오 판정한 비율을 나타내며 이는 시스템에 큰 영향을 미치지 않지만, FNR은 공격 데이터를 정상 데이터로 판단하는 비율로써 보안상 커다란 문제점을 야기하는 매우 중요한 지표이다. 본 실험의 결과에 의하면, KDD(탐지율: 96.65%, FNR: 0.52%)와 DDoS(탐지율: 96.25%, FNR: 0.50%)에서 모두 만족스러운 침입 탐지율과 안전한 FNR을 보여줌을 확인할 수 있었다.

[표 4]에서는 DDoS에서의 공격유형별 실험 결과이며 대체로 만족스러운 침입 탐지율과 FNR을 확인할 수 있었다.

V. 결 론

본 논문에서는 홈네트워크 환경에서 발생 가능한 보안 취약성을 분석하고, 이를 바탕으로 보안 요구사항들을 제시하였으며, 안전한 홈네트워크 서비스를 위한 새로운 계층적 분산 침입탐지 모델을 제안하였다. 제안된 모델은 홈게이트웨이에서 SVDD에 기반한 실시간 침입탐지를 수행할 뿐만 아니라 새로운 공격유형의 탐지 및 경제적 시스템의 갱신 등을 만족하는 차원에서 설계되었다. 또한, 에이전트에서는 클러스터 헤드별로 그룹화 된 홈기기들의 상태 모니터링 및 보고 그리고 관리를 수행하도록 설계되었다. 단 침입에 의한 이상 현상이 감지되면 주변 에이전트에게 경보를 신속히 전달하고 모든 연결고리를 끊음으로써 피해를 최소화한다, 만족스러운 침입 탐지율과 안전한 FNR 수치 등을 실험을 통하여 확인함으로써 제안된 시스템이 홈네트워크 환경에 적합함을 검증하였다.

〔표 4〕 DDoS에서의 공격유형별 성능 평가를 위한 측정 표

항목 / 공격 유형	침입 탐지율	FPR	FNR
TCP syn flood	98.00	10.00	0.00
UDP flood	97.00	11.00	0.00
ICMP echo request	96.00	13.00	1.00
Smurf	94.00	13.00	1.00
계	96.25	11.75	0.50

향후 연구 과제로는 이종의 유·무선 네트워크와 프로토콜의 혼재로 인한 보안 취약성을 극복하기 위한 종합적인 보안 인프라 구축과 분산 처리 및 홈네트워크 환경에 적합한 데이터베이스의 설계가 요구된다. 또한, 침입탐지 시스템에 의해 탐지된 공격들의 세부 정보를 침입방지 시스템의 정책 수립에 연동하는 방안과 지능적이고 적응적인 침입대응 방법론에 관한 연구가 요구된다.

참고문헌

[1] 고 훈, "홈네트워크 취약점 분석 및 인증 분석", 정보보호학회지, Vol. 16, No. 6, pp. 42~47, 2006.

[2] 유재학, 이한성, 정용화, 최성백, 양성현, 박대희, "유비쿼터스 홈네트워크를 위한 경량화된 침입탐지", 한국정보보호학회 하계학술대회, Vol. 16, No. 1, pp. 269~272, 2006.

[3] 이한성, 송지영, 김은영, 이철호, 박대희, "다중 클래스 SVM기반의 침입탐지 시스템", 퍼지 및 지능시스템학회 논문지, Vol. 15, No. 3, pp. 277~281, 2005.

[4] 박주영, 임채환, "비정상 상태 탐지 문제를 위한 서포트벡터 학습", 퍼지 및 지능시스템학회 논문지, Vol. 13, No. 3, pp. 266-274, 2003.

[5] Tarun Ambwani, "Multi class support vector machine implementation to intrusion detection", Proceedings of the International Joint Conference on Neural Networks, Vol. 3, pp. 2300~2305, 2003.

[6] Z. Wang, H. Wang, Q. Zhao, and R. Zhang, "Research on distributed intrusion detection system", Machine Learning and Cybernetics, Vol. 5, pp. 181~184, 2006.

[7] Z. Xie, Q. Shyu, S. Chen, and L. Chang, "A distributed agent-based approach to intrusion detection using the lightweight PCC anomaly detection classifier", IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing, SUTC 06, Vol. 1, pp. 446~453, 2006.

[8] K. Juszczyszyn, N. Nguyen, G. Kolaczek, A. Grzech, A. Pieczynsk, and R. Katarzyniak, "Agent-based approach for distributed intrusion detection system design", ICCS 2006, LNCS, Vol. 3993, pp. 224~231, 2006.

[9] P. Fu, D. Zhang, L. Wang, and Z. Duan, "Intelligent hierarchical intrusion detection system for secure wireless adhoc network", ISNN 2005, LNCS, Vol. 3498, pp. 482~487, 2005.

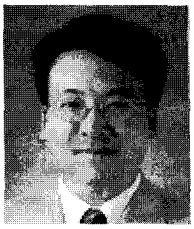
[10] KDD CUP 1999 DATA, Available in http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html

[11] Paul J. Criscuolo, "Distributed denial of service-Trinoo, Tribe Flood Network, Tribe Flood Network2000", CIAC-2319, 2000.

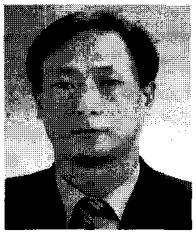
[12] Support Vector Machine, Available in http://svmlight.joachims.org/

〈著者紹介〉

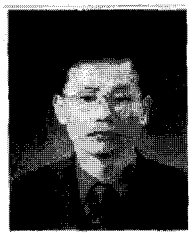
유 재 학 (Jaehak Yu) 학생회원
2001년 : 건국대학교 전산과학과 학사
2003년 : 고려대학교 전산학과 석사
2003년~현재 : 고려대학교 전산학과 박사과정
<관심분야> 기계학습, 데이터마이닝, 홈네트워크, 침입탐지



정 용 화 (Yongwha Chung) 종신회원
1984년 : 한양대학교 전자통신공학과 학사
1986년 : 한양대학교 전자통신공학과 석사
1997년 : Univ. of Southern California 전기공학과(컴퓨터공학 전공) 박사
1986년~2003년 : 한국전자통신연구원 생체인식기술연구팀장
2003년~현재 : 고려대학교 컴퓨터정보학과 부교수
<관심분야> 생체인식, 정보보호, 생체정보 보호



최 성 백 (Sungback Choi)
1982년 : 성균관대학교 전자공학과 학사
2004년 : 성균관대학교 정보보호학과 석사
1984년~ 1993년 : LG전자 과장 구 LG정보통신
1993년~ 2000년 : 마니네트웍 대표이사
2000년 ~ 현재 : 조은시큐리티 대표이사



양 성 현 (Sunghyun Yang)
1983년 : 광운대학교 전기과(공학사)
1987년 : 광운대학교 대학원 전기과(공학석사)
1992년 : 광운대학교 대학원 전기과(공학박사)
1991년 ~ 현재 : 광운대학교 전자공학부 교수

박 대 희 (Daihee Park)
1982년 : 고려대학교 수학과 학사
1984년 : 고려대학교 수학과 석사
1989년 : 플로리다 주립대학 전산학과 석사
1992년 : 플로리다 주립대학 전산학과 박사
1993년~현재 : 고려대학교 컴퓨터정보학과 교수
<관심분야> 지능 데이터베이스, 데이터마이닝