유비쿼터스 홈네트워크를 위한 경량화된 침입탐지

유재학*, 이한성*, 정용화*, 최성백**, 양성현***, 박대희*
*고려대학교 컴퓨터정보학과
**조은시큐리티
***광운대학교 정보통신대학원

Lightweight Intrusion Detection for Ubiquitous Home Networks

Jaehak Yu*, Hansung Lee*, Yongwha Chung*, Sungback Choi**, Sunghyun Yang***, Daihee Park*
*Department of Computer & Information Science, Korea University
{dbzzang, mohan, ychungy, dhpark}@korea.ac.kr,
**Joeunsecurity, Inc.
sbc@joeuns.com,
***Graduate School of Information Communication, Kwangwoon University
shyang@daisy,kw,ac.kr

요 약

최근 들어, 유비쿼터스 홈네트워크에 대한 관심이 높아지고 있지만, 기능 구현에 초점을 맞추고 있다. 이러한 홈네트워크는 단일 서비스가 아닌 다양한 서비스 집합으로서의 성격이 강하므로 세분화된 보안 요구사항과 제한된 자원에서의 원활한 서비스를 위해 시스템 경량화는 필수적 요소이다. 이에 본 논문에서는 유비쿼터스 홈네트워크 환경에서 요구하는 보안성 및 경량화를 고려한 새로운 침입탐지 모델을 제안한다.

Ⅰ. 서론

홈네트워크는 유선(wired)과 무선(wireless)의 통합 네트워크를 기반으로 댁내의 가전기기와 유·무선 정보기기들을 상호 연결하여 다양하고 편리한 서비스를 제공하며 다양한 지능적인 상황 인지 기기들을 통해 사용자의 편의를 극대화시키기 위한 기술 집약적인 산업으로 성장하고 있다[1,2].

이러한 홈네트워크는 생활의 중심에서 편의성을 제공하는 서비스인 만큼 서비스를 제공받는 사용자 입장에서는 편의성만큼이나 그 역기능에 대한 우려가 클 수밖에 없다. 따라서 서비스의 성공적인 제공을 위해 초기부터 보안 기능을 구현하고, 사용자 정보의 보호 및 안정적인 서비스 제공을 위한 보안의 중요성이 더욱 강조되고 있다[2].

본 논문에서는 기존 유·무선 네트워크에서의 침입탐지에 대한 연구동향 중 특히 패턴분류(pattern classification) 및 함수 근사(function approximation) 등의 문제에서 우수한 성능을 보이는 SVM(Support Vector Machine)[3]을 침입탐지에 적용한 모델을 제안한다. 또한, 홈네트워크에서의 세분화된 보안 요구사항과 구축비용을 감안할 때 저가로 시스템을 구축하여 시스템이 경량으로 가볍게 동작해야할 뿐만 아니라 별도의 추가적인 장비나 프로그램의 도움 없이도 동작하는 침입탐지 시스템을 설계하였다. SVM을 이용한 홈 네트워크에서의 침입탐지 알고리즘 성능 측정을 위해 침입탐지 측면에서 탐지율과 FPR(False Positive Rate), FNR(False Negative Rate), 시스템 경량화 측면에서 명령어 접근 수(access count)와 캐시 미스(miss)율을 비교, 분석하였다[4].

HNMMIDS(Home Network Multi-step Multi-class Intrusion Detection System)로 명명된 새로운 침입탐지 시스템은 단일 클래스 SVM과 다중 클래스 SVM, 그리고 점증적 갱신의 클러스터링 알고리즘인 Kernel-ART[3]를 계층적으로 결합한 구조이다.

본 논문의 구성은 다음과 같다. 2장에서는 홈네트워크에서의 보안 요구사항 및 SVM에 대해 설명한다. 3장에서는 본 논문에서 제안하는 홈네트워크에서의 경량화 된 침입탐지 모델 및 다중 클래스 SVM에 대해 설명하고, 4장에서는 실험결과 및 성능분석을 기술한다. 마지막으로 5장에서는 결론 및 향후 연구과제에 대해 논한다.

Ⅱ. 연구배경

1. 홈네트워크에서의 보안 요구사항 분석

기존 침입탐지 시스템을 비롯한 보안시스템을 홈네트워크 환경에 직접 적용이 어렵기 때문에 기존의 보안시스템과는 별도로 홈네트워크 환경을 위해 다음과 같은 점들을 고려해야 한다: 1) 대부분의 홈네트워크 사용자는 시스템에 대한 전문지식을 가지고 있지 않다. 따라서 시스템 갱신을 포함한 많은 부분이 자동화 되어야 한다; 2) 훈련된 보안 관리자가 운영하는 것이 아니므로 사용이 간편하고 정책설정이 쉬운 사용자 인터페이스를 제공하여야 한다; 3) 홈네트워크 구축비용을 감안할 때 저가로 시스템을 구축할 수 있어야 하며 따라서 시스템이 경량으로 가볍게 동작하여야 한다; 4) 홈네트워크의 침입이 정보가전의 동작에 영향을 미치기 전에 처리할 수 있도록 실시간성이 보장되어야한다; 5) 기존의 보안 제품과 달리 홈네트워크 기기는 각 가정에 보급되는 것을 목표로 하기 때문에 대량생산, 대량공급 및 운영 편의성이 보장되어야한다.

홈네트워크 환경에서의 접근경로는 크게 댁외망과 댁내망을 통한 접근으로 나누어진다. 댁외망은 외부 유선망과 외부 무선망이 현재 사용하고 있는 인터넷 망을 통해 접근하는 방법이고, 댁내망은 PLC나 802.11 등의 프로토콜 등을 이용한 접근 방법이다. 그림 1은 홈네트워크 환경에서의 접근형태를 보여주고 있다.

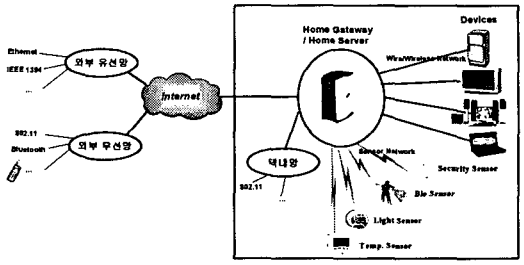

그림 1. 홈네트워크 환경에서의 접근경로 분석

2. 단일 클래스 SVM기반의 다중 클래스 SVM

d-차원 입력공간상에 존재하는 K-개의 데이터 집합 $D_k = \{ x_i^k \in R^d \mid i = 1, \ldots, N_k \};\ k = 1, \ldots, K$ 이 주어졌을 경우, 각각의 클래스를 분류하기 위한 다중 클래스 분류기는 각 클래스의 학습 데이터를 포함 하면서 체적을 최소화 하는 구체(sphere)를 구하는 문제로 정의 되며, 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.

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

$$\begin{aligned} f_k(x) \;=\; & R_k^2 - \Big[1 - 2 \sum_{i=1}^{N_k} \alpha_i^k k_k(x_i^k, x) \\ & + \sum_{i=1}^{N_k} \sum_{j=1}^{N_k} \alpha_i^k \alpha_j^k k_k(x_i^k, x_j^k) \Big] \ge 0 \end{aligned} \tag{2}$$

서로 다른 특징 공간상에서 정의되는 단일

SVM의 출력 $f_k(x)$값은 각 클래스의 특징 공간상의 경계로부터 해당 테스트 데이터와의 절대 거리를 의미함으로, 서로 다른 특징 공간상의 절대거리를 비교하여 소속 클래스를 결정하는 것은 바람직하지 않다. 따라서 특징 공간상의 절대거리 $f_k(x)$를 특징 공간상에서 정의되는 구형체의 반경 R_k로 나눔으로서 상대적 거리 $\hat{f}_k(x) = f_k(x)/R_k$를 계산하고, 상대거리가 가장 큰 클래스를 입력 데이터 x의 소속 클래스로 결정한다.

$$\text{Class of } x \equiv \arg\max_{k=1,\ldots,K} \hat{f}_k(x)$$

$$\equiv \arg\max_k \left[\left\{ R_k^2 - \left(1 - 2\sum_{i=1}^{N_k} \alpha_i^k k_k(x_i^k, x) + \sum_{i=1}^{N_k}\sum_{j=1}^{N_k} \alpha_i^k \alpha_j^k k_k(x_i^k, x_j^k)\right)\right\} / R_k \right] \quad (3)$$

3. Kernel-ART

Kernel-ART는 개념 벡터의 특성상 메모리 측면의 효율성이 높을 뿐만 아니라 클러스터링 수행 후, 각 클러스터의 내용 요약을 위해 별도의 대표 벡터 계산 없이 개념 벡터를 바로 사용하여 클러스터의 레이블링(labeling)을 수행할 수 있다는 장점을 가지고 있다[3].

Ⅲ. 제안하는 시스템 구조

HNMMIDS(home network Multi-step Multi-class Intrusion Detection System)로 명명된 새로운 침입탐지 시스템은 다음의 평가 기준들을 모두 만족하는 차원에서 설계되었다: 1) 홈네트워크에서의 침입유형 분류; 2) 시스템에서 학습되지 않은 새로운 공격 유형의 신속한 발견; 3) 탐지된 공격 유형에 대한 세부적 정보의 제공; 4) 빠르고 효율적인 학습 및 갱신으로 인한 경제적인 시스템의 유지/보수; 5) 시스템의 점증성(incrementality) 및 확장성.

그림 2는 HNMMIDS의 구조를 나타내며 다음의 네 단계에 걸쳐 수행되는 바, 각 단계에 대한 보다 상세한 설명은 다음과 같다.

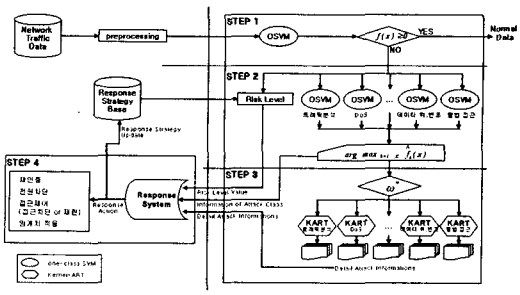

그림 2. HNMMIDS 구조도

제 1단계 : 정상데이터에 의해 학습된 단일 클래스 SVM은 정상 데이터와 공격 데이터에 대한 일차적인 분류를 빠른 속도로 수행한다. 학습 시 정상데이터만을 요구함으로 학습을 위한 별도의 공격 데이터를 준비할 필요가 없으며, 학습 속도 또한 매우 빠르다. 학습된 단일 클래스 SVM은 비정상 탐지모델로써 시스템에서 학습되지 않은 새로운 공격(novel attack)을 탐지하며, 시스템 운영 시 정상데이터에 대한 추가적인 처리과정은 없고, 공격데이터가 탐지되면 침입대응 시스템(intrusion response system)에 1차적 경고를 발생시킨다.

제 2단계 : 1단계에서 공격으로 분류된 데이터들을 다중 클래스 SVM에 의해 하나의 공격 유형으로 분류하고 침입대응 시스템에 공격 유형에 대한 추가 정보를 제공한다. 또한 새롭게 발견된 공격 데이터를 시스템에 반영할 경우, 시스템 전체를 재학습 시킬 필요 없이 해당 클래스의 분류기만을 점증적 갱신의 학습방법으로 재학습한다.

제 3단계 : 해당 공격에 대한 보다 세부적인 정보가 요구될 경우, 각 공격 종류별 클러스터링이 Kernel-ART에 의해 수행되며 각 클러스터의 내용 요약을 위해 별도의 대표 벡터 계산 없이 개념 벡터를 사용함으로 공격 별 레이블링(labeling)을 제공한다.

제 4단계 : 공격 데이터에 대한 risk level 정보와 공격 유형별 정보, 해당 공격에 대한 세부적인 정보를 이용해 재인증, 전원차단, 접근제어 등의 대응 방법을 결정하여 피해를 최소화 한다.

Ⅳ. 실험 결과 및 성능 분석

본 논문에서 제안한 침입탐지 시스템의 성능 검증을 위해 여러 공격유형 중 DoS 공격을 분산화, 자동화시켜 더욱 발전시킨 DDoS(Distributed Denial of Service) 공격 방법을 실험 대상으로 삼았다. 실험을 위해 정상적인 트래픽 5,000개와 DDoS의 대표적 툴인 Trinoo, TFN, TFN2K, Stacheldrath로 공격 트래픽을 각 1,000개씩 생성하고[5], 정상 트래픽 1,000개와 각 공격별 트래픽 200개씩을 선택하여 학습하였다. 또한, 홈네트워크 환경에서의 제한된 자원 이용을 고려하기 위해 SVM[6] 알고리즘의 수정 전·후를 분석하고 검증하기 위해 simplescalar[7]를 기반으로 한 panalyzer[8]로 그 성능을 실험하였다.

아래의 표 1은 SVM 수정 전·후의 탐지율과 FPR, FNR을 나타낸다. sigma 값(σ value)은 커널의 범위를 결정하거나 입력 공간에서 고려되는 데이터를 결정하는데 사용한다.

σ	수정 전 SVM			수정 후 SVM		
	탐지율	FPR	FNR	탐지율	FPR	FNR
0.1	81.17	66.11	0.11	82.15	63.43	0.11
0.9	99.68	32.25	0.22	99.89	28.79	0.21
1.7	96.58	15.48	6.78	97.17	14.76	6.77

표 1. 공격 트래픽 탐지율 비교

FPR은 정상 트래픽을 공격 트래픽으로 오 판정한 비율을 나타내며 이는 시스템에 큰 영향이 미치지 않지만, FNR은 공격 트래픽을 정상 트래픽으로 판단하기 때문에 보안상 커다란 문제가 될 수 있다. 이에 본 논문에서는 비교적 정확한 탐지율과 FNR 값이 최소가 되는 sigma 0.9 값을 사용하였다. 또한, 캐시 크기가 32Kbyte일 때를 가정한 캐시 미스율이 4.5%에서 3.37%로 감소됨을 확인하였다.

Ⅴ. 결론

본 논문에서는 기존의 유·무선 네트워크 환경과는 다른 홈네트워크 환경에서의 발생 가능한 보안 취약성을 분석하고, 이를 바탕으로 보안 요구사항들을 제시하였으며, 제한된 자원에서의 원활한 서비스를 위해 시스템 경량화에 기여하였다. 또한, 홈네트워크 환경에서의 다중 클래스 SVM은 점증성과 확장성, 빠르고 효율적인 학습 및 갱신 등 기존의 다중 클래스 SVM에 비해 알고리즘의 복잡도를 향상시킨 새로운 방법론이다.

본 연구의 향후 연구 과제로는 제안된 홈네트워크 환경에서의 침입탐지 시스템에 의해 탐지된 공격들의 세부 정보를 침입방지 시스템의 정책 수립과 지능적이고 적응적인 침입대응 방법론에 관한 연구가 향후 연구과제로 요구된다.

Acknowledgments

본 연구는 산업자원부 및 한국산업기술평가원의 성장동력기술개발사업의 연구결과로 수행되었습니다.

[참고문헌]

[1] 김도우, 한종욱, 주홍일, 이윤경, "디지털홈 환경에서의 보안 프레임워크 연구", 한국해양정보통신학회 춘계종합학술대회, pp. 724-727, 2004.

[2] Bo Zhou, Qi Shi, Madjid Merabti, "Real-Time Intrusion Detection in Ubiquitous Networks with a String-Based Approach", Workshop on Ubiquitous Application and Security Service, LNCS 3983, pp. 352-359, 2006.

[3] 이한성, 송지영, 김은영, 이철호, 박대희, "다중클래스 SVM기반의 침입탐지 시스템", 퍼지 및 능 시스템학회 논문지, Vol. 15, No. 3. pp. 277-281, 2005.

[4] Guy H, Johnny S.K Wong, Vasant H, Les M, Yanxin W, "Lightweight Agents for Intrusion Detection", The Journal of Systems and Software, pp. 109-122, 2003.

[5] Paul J. Criscuolo, "Distributed Denial of Service - Trinoo, Tribe Flood Network, Tribe Flood Network2000", CIAC-2319, 2000.

[6] http://svmlight.joachims.org/

[7] http://www.simplescalar.com/

[8] http://www.eecs.umich.edu/~panalyzer/