가우시안 혼합 모델을 이용한 네트워크 침입 탐지 시스템

박명언°* 김동국** 노봉남*
*전남대학교 정보보호 협동과정, **전남대학교 전자컴퓨터정보통신공학부
parkmo°@lsrc.chonnam.ac.kr, dkim@chonnam.ac.kr, bongnam@chonnam.ac.kr

Network Intrusion Detection System Using Gaussian Mixture Models

Myung-Aun Park°* , Dong-Kook Kim**, Bong-Nam Noh*
*Interdisciplinary Program of Information Security, Chonnam National University
**Dept. of Electronics, Computer and Information Eng., Chonnam National University

요 약

초고속 네트워크의 폭발적인 확산과 함께 네트워크 침입 사례 또한 증가하고 있다. 이를 검출하기 위한 방안으로 침입 탐지 시스템에 대한 관심과 연구 또한 증가하고 있다. 네트워크 침입을 탐지위한 방법으로 기존의 알려진 공격을 찾는 오용 탐지와 비정상적인 행위를 탐지하는 방법이 존재한다. 본 논문에서는 이를 혼합한 하이브리드 형태의 새로운 침입 탐지 시스템을 제안한다. 기존의 혼합된 방식과는 다르게 네트워크 데이터의 모델링과 탐지를 위해 가우시안 혼합 모델을 사용한다. 가우시안 혼합 모델에 기반한 침입 탐지 시스템의 성능을 평가하기 위해 DARPA'99 데이터에 적용하여 실험하였다 실험 결과 정상과 공격은 확연히 구분되는 결과를 나타내었으며, 공격 간의 분류도 상당 수 가능하였다.

1. 서 론

최근 인터넷의 급속한 발전에 따라 다양하고 고도화된 많은 해킹 기법들이 나타나게 되었다. 이를 탐지하기 위하여 많은 침입 탐지 시스템(Intrusion Detection System)들이 개발 되었다. 침입 탐지 시스템은 이미 알려진 침입 행위에 대한 정보를 이용하여 공격을 탐지하는 오용탐지(Misuse Detection)와 사용자의 정상 행위를 기반으로 정상적인 행동 패턴에 어긋나는 경우를 침입으로 탐지하는 비정상행위탐지(Anomaly Detection)로 나뉜다[1].

현재 침입탐지 시스템에서 가장 일반적인 형태는 네트워크 기반 오용탐지 시스템이다[2]. 오용탐지 시스템은 일반적으로 알려진 공격기법에 대한 탐지 능력만을 가지고 있고 전문가의 노력이 많이 요구된다는 단점을 가지고 있지만 확실한 공격에 대한 시그너처를 사용하는 Snort 같은 시스템의 경우 일반적으로 적은 False Alarm을 발생한다[3,6].

이러한 단점을 해결하기 위한 방안으로 정상 행위 프로파일링을 통한 비정상행위 탐지 기법에 대한 활발한 연구가 진행되고 있지만 모든 정상을 수집할 수 없는 한계와 정상과 공격을 구분 짓는 데 있어서 많은 어려움이 존재한다.

본 논문에서는 가우시안 혼합 모델(Gaussian Mixture Models : GMMs)을 네트워크 침입 탐지 시스템에 적용하여 정상과 공격을 구분 지을 수 있는지를 실험 및 분석하였다. 이를 위한 데이터는 이미 검증된 DARPA'99 DataSet의 2주 데이터를 사용하였다.

본 논문의 구성은 다음과 같다. 2절에서는 네트워크 데이터 셋의 특성에 대하여 제시하고, 3절에서는 GMMs의 사용을 위한 학습, 클러스터링, 테스트 방법에 대하여 설명한다. 4절에서는 구체적인 실험 및 분석 결과를 다루고 5절에서는 총괄적인 요약과 향후 연구방향을 제시한다.

2. 데이터 셋

본 논문에서 각 데이터는 TCP 세션 단위로 처리하여 사용하게 하였다. 네트워크 서비스는 TCP 연결을 통하여 이루어지며, 데이터는 여러 개로 쪼개져서 전송 될 수 도 있기 때문에 네트워크상의 패킷들은 각각의 패킷 단위보다 세션을 이룰 때 의미가 큰 정보를 가지고 있게 된다. 그러므로 본 논문에서는 세션단위로 패킷을 모아서 패킷끼리의 연산을 수행하였다.

GMMs에 적용하기 위하여 네트워크 세션 데이터에서 네트워크 헤더 정보를 구성하는 다음과 같은 기본적인 5가지 속성을 추출하였다.

$$O = \{ \text{Source Port, Destination Port, IP Flag, TCP Flag, Data Size} \}$$

네트워크의 특성에서 각 속성 정보들의 특징을 보면 IP Flag와 TCP Flag 의 경우는 범주(categorical)가 개별적으로 정해져 있는 경향이 있고 Data Size의 경우는 연속적인(continuous) 경향이 있다. 이 처럼 현재 사용하고자 하는 속성 정보 값들은 서로 다른 형태의 경향을 가지고 있다. 이러한 정보를 GMMs 형태에서 최적의 해를

구하기 위해서는 각 데이터 간의 범주를 비슷하게 하여야한다[7]. 그러므로 이를 위하여 다음과 같은 정규화를 통한 전처리 작업을 하였다.

값 = (본래값 - 평균값) / 표준편차 (2.1)

이를 통하여 임의적으로 각 데이터들의 편차를 줄임으로서 데이터 유형이 동일하다는 가정을 통하여 하나의 벡터로 가정하고 처리를 할 수 있게 되었다.

3. 가우시안 혼합 모델링

가우시안 혼합 모델(Gaussian Mixture Models)은 통계적 패턴 인식 기반의 음성인식이나 얼굴인식 등의 생체인식 시스템에 많이 사용되고 있는 모델이다[4]. GMMs의 특징은 정보량 기반의 클러스터링은 특정한 프로토타입이나 형태와 상관없이 정보량의 안정화에 따라 클러스터링을 수행한다는 것이다. 또한 GMMs은 단일 상태를 갖는 HMMs(Hidden Markov Models)라고 할 수 있으며, 이는 HMM의 학습 방법을 조금의 수정을 통하여 GMMs에도 적용 할 수 있음을 의미한다.[5]

본 논문에서는 포괄적으로 쓰이는 형태를 취하기 위하여 음성인식에서 널리 사용되고 있는 형태로 모델링 하여 적용하였다. 본 실험을 위한 GMMs 공식은 다음과 같다.

D-차원 요소 벡터 o_t 는 완전한 관측 시퀀스

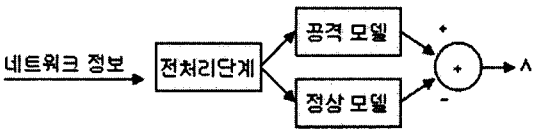
$O = [\ o_1, \ldots, o_t, \ldots, o_T\ ]$ 의 일부일 때,

$$P(o_t|\theta) = \sum_{m=1}^{M} w_m N(o_t|\mu_m, \Sigma_m) \quad (3.1)$$

$$N(o_t|\mu_m, \Sigma_m) = \quad (3.2)$$

여기서 T는 속성을 5개를 사용하였으므로 5개가 되며, o_1부터 차례대로, 전처리 작업을 거친 { Source Port, Destination Port, IP Flag, TCP Flag, Data Size } 가 된다. 학습된 결과물이라고 할 수 있는 파라미터 벡터 $\theta=\{w_m, \mu_m, \Sigma_m\}$, $m=1,\ldots,M$로 나타낸다. M은 최종 Mixture Order로서 학습 후, 최종 클러스터 수를 나타낸다. 그리고 W_m은 가중치 값(weight)을 μ_m은 평균 값(mean)을 Σ_m는 공분산(covariance)을 말한다.

본 논문에서 학습이라 함은 전처리를 거친 속성 값 벡터 O를 가지고 최적의 파라미터 벡터를 구하는 것을 말한다. 각 네트워크 데이터는 세션단위의 정보를 사용하였고 정상 네트워크 패킷들은 1개의 군으로 공격들은 각 공격별 이름과 요일로 나누어 학습하였다.



<그림 1> likelihood ratio 기반의 침입탐지 시스템

학습 과정을 통하여 그림 1의 정상 모델과 공격 모델을 구하게 된다. 테스트 수행 시에는 정상 모델은 음수의 라이클리후드 값을 가지게 되고 공격 모델은 양수의 값을 가지게 된다. 또한 판별이 필요한 네트워크 정보는 공격인지 정상인지를 판단하기 위하여 다음과 같은 공식을 사용하여 최대 라이클리후드를 구하여 가장 가까운 쪽으로 판별하게 된다.

$$P(O \mid \theta) = \prod_{t=1}^{T} P(O_t)\ ,$$

$$\max\left(\frac{P(O_t|\theta)}{P(O_t|\theta)}\right)\begin{cases} \geq\theta \\ <\theta \end{cases} \quad (3.3)$$

GMMs를 라이클리후드 함수로 이용하게 되면 식 3.2와 3.3 처럼 매우 잘 알려진 형태의 통계적 모델을 이용하기 때문에 계산이 매우 간단진다는 장점이 있다. 계산은 은 간단하지만 실제 실험에서 사용되어 지는 값은 매우 큰 수가 구하여 지므로 오버플로(overflow)를 방지하지 위하여 log를 취하여 계산하였다. 그리고 새로운 공격이 추가 될 경우 공격 전체를 학습할 필요가 없다. 이전에 학습된 것들을 각각을 별개로 사용할 수 있기 때문에 추가된 새 공격 데이터만 학습을 하고 탐지 식에 적용할 수 있다.

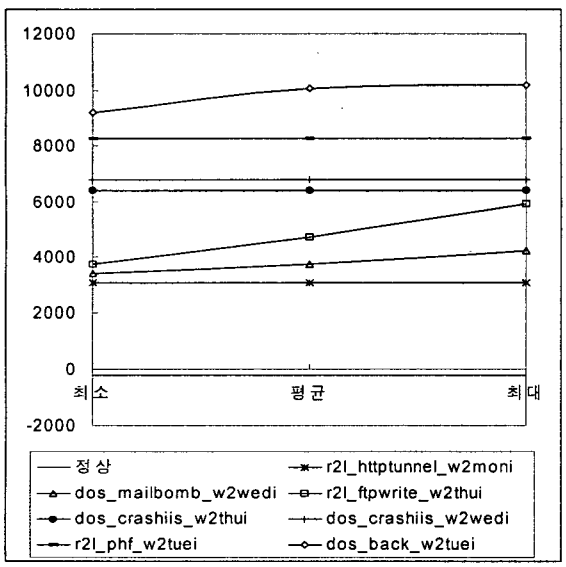
4. 실험 및 분석

본 연구에서는 기본적인 학습과 실험을 위하여 DARPA'99 데이터 셋의 주별 데이터의 일부를 사용하였다. 데이터 셋은 2주 데이터만을 사용하였는데, 이는 2주 데이터의 정상을 가지고 학습 한 것이 공격을 분류해 내는데 있어서 효과적이라고 판단하였기 때문이다. 정상 학습에 사용한 데이터는 2주차 월요일, 금요일의 inside dump이며, 정상 테스트용 데이터는 2주차 화요일과 수요일 inside dump를 사용하였고 공격군은 2주차의 각 공격을 일자별로 학습용과 실험용을 나누어 사용하였다.

GMMs를 사용한 학습에서 학습된 파라미터 벡터 값들을 구하는데 Mixture Order가 결과에 영향을 줄 수 있다. 이를 알아보기 위하여 64, 128, 256, 512, 1024로 나누어 학습을 실시하였다. 하지만 Mixture Order를 다른 값으로 바꾸어 한 것과는 기대했던 것만큼 많은 차이가 나진 않았다. 표 1은 Mixture Order를 256으로 하여 테스트 데이터의 Detection을 위하여 값을 구하여 본 것이다. probes_portsweep_w2thui는 2주차 목요일 inside dump 중 probes 부류에 해당하는 portsweep 이라는 공격을 말한다.

<표 1> Mixture Order 256에 대한 각 데이터 likelihood 분포

분 류	최대	최소	평균값	세션수
정상	-223.79	-223.91	-223.85	83182
probes_portsweep_w2thui	15486.2	6873.87	9561.95	15
probes_portsweep_w2tuei	43081.7	43057.2	43072.7	11
probes_satan_w2wedi	15554.7	4460.07	8834.36	26
dos_neptune_w2thui	52357.7	14544.6	25288.5	10240
dos_land_w2thui	24267.3	24267.3	24267.3	1
dos_crashiis_w2wedi	6762.08	6762.08	6762.08	1
dos_crashiis_w2thui	6410.25	6410.25	6410.25	1
dos_back_w2tuei	10192.4	9220.85	10081.2	40
dos_mailbomb_w2wedi	4253.2	3423.38	3773.28	500
r2l_phf_w2tuei	8242.53	8242.53	8242.53	1
r2l_ftpwrite_w2thui	5895.64	3758.92	4713.83	3
r2l_httptunnel_w2moni	3084.41	3084.41	3084.41	1



<그림 2> 각 공격별 결과 그래프

정상의 최대 값과 공격들의 최소 값들을 비교하여 보면 정상과 공격은 확연하게 차이가 나는 것을 알 수 있다. 이는 θ값을 0으로 놓아도 분별이 가능 하다는 것을 말해 준다. 이를 그래프 형태로 나타내어 분포를 살펴보면 그림 2에서 나타나는 것처럼 각 값들은 서로 겹치는 부분이 거의 없게 되어 각 공격에 대한 분류까지 가능할 것으로 보인다. 하지만 dos_neptune_w2thui, probes_portsweep_w2thui, probes_satan_w2wedi와 같은 형태는 다른 공격과의 큰 차이를 보이는 형태의 값을 갖진 못했고 이는 공격이라고 탐지는 가능하나 다른 형태의 공격으로 오인하여 탐지하게 될 가능성이 있음을 말한다.

하지만 식 3.2의 $(O_t-\mu_m)$를 보면 알 수 있듯이 현재 네트워크에서 수집된 패킷의 값들은 즉각적으로 가우시안 PDF(probability density function)의 결과 값에 변화를 주며 식 3.3에서 θ값과 바로 비교하여 판별하기 때문에, 실시간 탐지도 가능할 것으로 보인다.

5. 결론

본 연구에서 접근한 침입 탐지 시스템은 정상 패턴과 공격 패턴을 동시에 사용하는 하이브리드 형태의 접근 방식으로 분류 될 수 있겠지만, 기존의 방식보다 다음과 같은 장점을 지닌다.

우선 공격이 늘어나게 되더라도 해당 공격만 따로 학습하여 테스트 시에 적용할 수 있다. 둘째로 잘 알려진 가우시안 혼합 모델을 사용함으로서 테스트 속도 또한 빠르다고 할 수 있다. 마지막으로 실시간 수집된 데이터를 바로 테스트 하여 임계값에 따라 공격 여부를 판별 해 낼 수 있다. 또한 제안한 이 방법은 현재 매우 단순한 형태의 속성만을 사용하므로 속성 추가를 통하여 더 복잡하고 정확한 시스템을 구축 가능 할 것이다.

본 논문은 GMMs과 최대 라이클리후드를 사용한 네트워크 침입 탐지 모델을 보임으로서 각 공격이 정상과 분명한 차이를 가짐을 보였다. 앞으로 실시간 탐지와 알려지지 않은 공격 탐지에 있어서 얼마 만큼에 성과가 있는지를 알아보는 것이 중요한 과제이다.

6. 참고 문헌

[1] T. F. Lunt, "A Survery of Intrusion Detection Technique," Computer & Security. Vol 12. No.4, January 1993.

[2] 박종영, "네트워크 침입탐지를 위한 밀도함수 기반 아웃라이어 탐지 기법", 한국컴퓨터종합학술대회 2005 논문집 Vol. 32, No.1, p148-150, 2005

[3] S. E. Smaha, "Tools for Misuse Detection", In proceedings of ISSA' 93, Crystal City, VA, 1993

[4] Conrad Sanderson, "Likelihood normalization for face authentication in variable recording conditions", IEEE ICIP 2002, I-301~I-304, 2002

[5] Jonas Richiardi, "Gaussian Mixture Models for On-line Signature Verification", ACM WBMA'03, November 8, 2003

[6] www.snort.org

[7] Jonas Samuelsson, "Multiple description coding based on Gaussian mixture models", IEEE Signal Processing Letters VOL. 12, p449-452 June 2005

본 연구는 정보통신부 대학 IT연구센터 육성,지원사업의 연구결과로 수행 되었습니다