Support Vector Machine을 이용한 DoS 탐지에 관한 연구

김종호° 서정택 문종섭
고려대학교 정보보호 대학원, 국가보안기술연구소
{sky45k°, jsmoon}@korea.ac.kr, seojt@etri.re.kr

An Approach for DoS Detection with Support Vector Machine

Jongho Kim° Jungtaek Seo Jongsub Moon
Center for Information Security Technology, Korea University, ETRI

요 약

서비스 거부 공격은 그 피해의 규모에 비해 방어하기가 무척 어려우며 충분히 대비를 한다 해도 알려지지 않은 새로운 서비스 거부 공격 기법에 피해를 입을 위험성이 항상 존재한다. 또한 최근 나타나고 있는 서비스 거부 공격 기법은 시스템 자원을 고갈시키는 분산 서비스 거부 공격(DDoS)에서 네트워크의 대역폭을 고갈시킴으로서 주요 네트워크 장비를 다운시키는 분산 반사 서비스 거부 공격(DRDoS)으로 진화하고 있다. 이러한 공격 기법은 네트워크 트래픽의 이상 징후로서만 탐지될 뿐 개별 패킷으로는 탐지가 불가능하여 공격 징후는 알 수 있으되 자동화된 대응이 어려운 특징이 있다.

본 논문에서는 이미 알려진 공격뿐 아니라 새로운 서비스 거부 공격 패킷을 탐지하기 위하여, 패턴 분류 문제에 있어서 우수한 성능을 보이는 것으로 알려져 있는 Support Vector Machine(SVM)을 사용한 실험을 진행하였다. 테스트 결과, 학습된 공격 패킷에 대해서는 정확한 구분이 가능했으며 학습되지 않은 새로운 공격에 대해서도 탐지가 가능함을 보여주었다.

1. 서 론

최근의 인터넷 환경은 급속한 네트워크의 보급과 함께 그에 따른 정보보호 문제를 발생시키고 있으며 이러한 네트워므 보안 문제에 대한 해결책으로 여러 가지 보안 솔루션들이 사용되고 있다. 그러나, 기존의 보안 솔루션들은 이미 알려진 공격에 대한 탐지와 대응을 위한 것이며, 새로운 기법의 공격에 대해서는 한계를 보이고 있다. 인터넷 환경의 대중화와 함께 취약점의 발견되면 수일 내에 공격 코드의 개발이 이루어지고 있음을 감안할 때, 새로운 형태의 공격에 대한 대응책의 마련은 중요한 연구 과제라고 하겠다.

특히 서비스 거부 공격의 경우, 공격자는 다른 공격 기법에 비해 많은 노력을 기울이지 않고도 공격 대상에게 막대한 피해를 입힐 수 있으며, 근원지의 주소를 스푸핑하거나 분산 반사 서비스 거부(DRDoS) 공격[4][5] 등을 사용하여 쉽게 자신을 은폐할 수 있는 특성이 있다. 또한 분산 반사 서비스 거부 공격은 정상적인 서비스를 운영하고 있는 서버를 에이전트로 활용하기 때문에 해커들이 손쉽게 이용할 수 있으며 공격의 근원지를 추적하기가 어렵고 공격을 막을 수 있는 방법이 그리 용이하지 않다. [4][5]

본 논문에서는 TCP/IP를 이용한 서비스 거부 공격의 탐지를 위하여 은닉 채널, SPAM 필터링 등 패턴분류 문제에 있어서 우수한 성능을 보이는 것으로 알려진 Support Vector Machine[1][2]을 이용하였으며 학습된 공격 기법 및 새로운 형태의 공격 기법의 탐지에 관한 실험을 진행하였다.

2. 서비스 거부 공격 탐지 방안

SVM을 이용한 서비스 거부 공격 탐지에 관한 기존 연구는 주로 1998 DARPA 침입 탐지 평가 프로그램의 데이터를 이용하여 수행되어 왔다.[6][7][8] 그러나 여기서 사용되었던 서비스 거부 공격 기법들은 매우 고전적인 것으로, 현재에는 대부분 패치가 이루어져 거의 사용되고 있지 않다. 또한 기존의 연구에서는 학습 시 사용된 공격 기법과 테스트 시 사용한 공격 기법이 동일하다는 단점이 있다.

본 논문에서는 기존의 공격 도구들을 사용하여 생성된 패킷을 이용하여 SVM의 학습 데이터로 사용하고, 학습 시 사용된 공격 도구와 2000년 이후 주로 사용되던 분산 서비스 거부 공격 도구 및 최근 공격 기법인 분산 반사 서비스 거부 공격 도구로 생성된 패킷을 이용하여 3개 부분으로 테스트를 진행하였다.

2.1 공격용 패킷 분석

서비스 거부 공격을 위해서 공격자는 일반적으로 헤더 부분에 비정상적인 데이터를 가진 패킷을 전송하여 이를 처리하기 위하여 시스템의 자원을 소비시키도록 한다. 지금까지 알려진 공격들이 사용하는 TCP/IP 헤더의 필드는 다음과 같다.

- bonk : IP id, IP flags, Fragment offset
- jolt : IP id, IP flags, Fragment offset
- land : Source/Destination IP/Port
- nestea : IP id, IP flags, Fragment offset
- newtear : IP id, IP flags, Fragment offset
- syndrop : Seq num, IP id, IP flags, Fragment offset
- TearDrop : IP id, IP flags, Frag. offset
- WinNuke : TCP flags, option field

표 1. DoS 공격에 주로 사용되는 IP 헤더 필드

Ver	Len	Service Type	Total length
Identification		Flag	Fragmentation Offset
Acknowledgement number			
TTL	Protocol	Header checksum	
Source IP address			
Destination IP address			
Option			

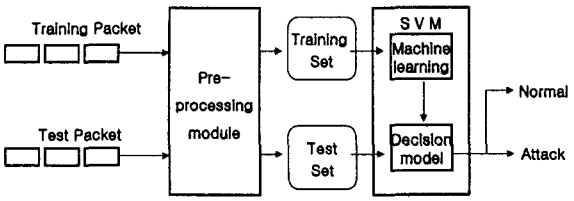
표 2. DoS 공격에 주로 사용되는 TCP 헤더 필드

Source Port number							Destination Port number
Sequence number							
Acknowledgement number							
Offset	Reserve	U	A	P	R	S	F	Window Size
Checksum							Urgent pointer
Option and Padding							

2.2 탐지 방안

본 장에서는 TCP/IP를 이용한 서비스 거부 공격을 탐지하기 위하여 SVM 학습방안을 제안한다. 제안하는 학습방안은 SVM이 주어진 데이터 집합에 대하여 이진 분류기의 역할을 한다는 기본 개념에 착안한 것으로서 정상 패킷들과 비정상 패킷들의 집합들을 구분하기 위해 여러 SVM 커널 함수들을 적용하여 최적의 분리 경계면을 찾아내는 것이다. 이러한 최적 분리 경계면을 찾는 것은 결국 서비스 거부 공격을 위한 TCP/IP 패킷을 탐지하는 것과 같은 결과를 가져오게 된다.

그림 1. SVM 학습을 통한 탐지 방안



SVM 학습을 통한 탐지방안의 구조는 그림 1과 같다. 먼저 학습과 테스트를 위한 패킷을 수집하고, 이 패킷에 대하여 전처리 과정을 거쳐 feature를 선택하여 Training Set과 Test Set을 구성한다. Training Set은 SVM의 학습 데이터로 사용되고 Test Set에 대한 공격 탐지를 수행한다.

3. 서비스 거부공격 탐지 실험 및 결과

3.1 실험 환경

실험을 위하여 먼저 SVM을 학습하기 위한 학습 데이터 집합과 테스트 데이터 집합을 구성하였으며 이때 사용되는 데이터들의 경우 tcpdump를 사용하여 수집하였다. 이상탐지의 특성상 학습 시 사용되지 않은 데이터로 테스트를 하는 것이 정확한 결과를 얻을 수 있으므로, 정상 데이터의 경우 학습용 패킷과 테스트용 패킷을 다르게 하여 데이터를 수집하였으며 http, ftp, telnet, smtp, pop3 가 서비스 되고 있는 서버에서 수집하였다. 비정상 패킷은 분리된 네트워크에서 서비스 거부 공격 도구를 사용하여 수집하였다. 학습 및 테스트 데이터를 수집하기 위하여 사용한 공격 도구 및 공격 기법은 표 3과 같다.

표 3. 공격 패킷 수집시 사용된 도구 및 공격 기법

구 분	공격 도구	공격 기법
학습 집합	targa2 (DoS)	bonk, jolt, land, nestea, newtear, syndrop, teardrop, winnuke, 1234, saihyousen, oshare
테스트 집합	targa2 (DoS)	bonk, jolt, land, nestea, newtear, syndrop, teardrop, winnuke, 1234, saihyousen, oshare
	TFN2K (DDoS)	targa3-IP stack penetration
	pHorgasm (DRDoS)	DRDoS

학습 및 테스트에 사용되는 패킷들의 전처리 과정에 사용된 feature 값들은 2.1에서 설명된 바와 같이 서비스 거부 공격에서 주로 사용되는 헤더정보 중에서 다음의 표 4와 같이 구성하였다. 이때 차원 하나의 값은 패킷의 헤더의 16진수를 10진수로 변환하여 사용하였다.

표 4. 헤더에서 전처리되어 수집된 features

mode	feature 개수	feature description
1	5	IP id IP flags IP fragment offset TCP offset TCP flags
2	7	**mode 1에서 사용된 feature +** TCP sequence number TCP acknowledge number
3	10	**mode 2에서 사용된 feature +** Urgent pointer 2/Option field
4	12	**mode 3에서 사용된 feature +** 2/Option field

Train Set 은 모두 10000개의 패킷을 사용하여 전처리하였으며 각각 정상 패킷 5000개와 서비스 거부 공격 패킷 5000개를 사용하였다. Test Set은 모두 2000개의 패킷을 사용하여 전처리하였으며 학습 시 사용되지 않은 정상 패킷 1000개와 학습 시 사용되지 않은 서비스 거부 공격 패킷 1000개를 사용하였다.

표 5. SVM 학습 데이터 집합

DataSet / 패킷종류	Training Set 1 - No Sliding(10000개)
정상 패킷	개별 TCP/IP 패킷 (5000개)
공격 패킷	targa2에서 생성된 패킷(5000개)

표 6. SVM 테스트 데이터 집합

DataSet / 패킷종류	Test Set (4000개)
정상 패킷	개별 TCP/IP 패킷 (1000개)
공격 패킷	targa2에서 생성된 패킷(1000개)
	TFN2K에서 생성된 패킷(1000개)
	pHorgasm에서 생성된 패킷(1000개)

본 실험에서는 mySVM 공개 도구를 사용하여 실험하였으며 이때 SVM 탐지 성능 비교를 위한 SVM 커널로는 linear와 polynomial 커널을 사용하였고 polynomial 커널에서 degree는 3으로 고정하여 실험하였다. 실험에서 사용된 mySVM 공개 도구는 SVM 분류나 회귀 예측 알고리즘을 구현한 검증된 도구 중의 하나로서 독일 Dortmund 대학에서 개발하였으며, 현재 SVM을 이용한 응용 연구에 가장 널리 사용되는 도구로서 알려져 있다.[10]

3. 2. 실험 결과

SVM을 이용한 TCP/IP 서비스 거부 공격 탐지 실험은 커널 함수로서 linear, polynomial을 사용하여 각각 진행하였다.

표 7. SVM을 이용한 공격 탐지 결과
(FP : false positive, FN : false negative)

kernel	mode	FP	FN		
			targa2 (DoS)	TFN2K (DDoS)	pHorgasm (DRDoS)
linear	1 (5 feature)	0.3	33.0	30.4	100.0
	2 (7 feature)	4.8	0.0	60.7	18.2
	3 (10 feature)	3.2	0.0	60.0	84.5
	4 (12 feature)	0.8	34.0	64.0	62.4
poly	1 (5 feature)	4.2	0.0	46.5	98.5
	2 (7 feature)	0.0	0.0	48.2	50.2
	3 (10 feature)	0.0	0.0	33.2	96.8
	4 (12 feature)	3.9	0.0	25.8	97.4

표 7은 각 mode별 TCP/IP의 서비스 거부 공격에 대한 SVM의 탐지 결과이다. 학습 패턴에 따른 탐지결과를 분석하면 전반적으로 5% 미만의 false positive를 보이고 있으며 학습된 targa2 공격에 대해서는 우수한 탐지 성능을 보이고 있다. TFN2K의 경우 polynomial 커널에서의 탐지성능이 좀 더 향상되었으며 pHorgasm은 mode 2에서의 탐지 성능이 가장 좋게 나타나고 있다. polynomial 커널을 적용했을 때에는 mode 2와 3에서 false positive가 0%로 나타났으며 학습된 공격 기법의 경우 정확히 탐지해냈고 DDoS와 DRDoS는 mode2에서 약 50% 탐지하였다.

4. 결론 및 향후 연구 방향

본 논문에서는 TCP/IP를 이용한 서비스 거부 공격에 대하여 SVM을 이용한 탐지방안을 제안하였으며 비록 TCP/IP를 이용한 공격으로 한정하였지만 최근 주로 사용되는 서비스 거부 공격 도구를 사용하여 실험을 하였다. 실험을 통하여 학습된 공격 기법은 정확하게 탐지하며 새로운 형태의 공격에 대해서도 일정부분 탐지가 가능함을 알 수 있었다.

향후 성능 향상 변수로 사용될 수 있는 커널의 종류와 제약 조건 인자에 대한 고려를 통하여 보다 정확한 실험 결과를 도출할 수 있을 것으로 사료되며, TCP/IP 이외의 프로토콜을 사용하는 서비스 거부 공격에 대한 탐지 실험도 필요할 것이다.

참 고 문 헌

[1] Vapnik. V., "The Nature of Statistical Learning Theory" Springer-Verlag, NewYork, 1995

[2] C. Campbell, N. Cristianini, "Simple Learning Algorithms for Training Support Vector Machines" Technical report, University of Bristol, 1998

[3] Pontil. M., Verri. A., "Properities of Support Vector Machines" A. I. Memo No. 1612:CBCL paper No. 152, Massachusetts Institute of Technology, Cambridge, 1997

[4] Vern Paxson, "An Analysis of Using Reflectors for Distributed Denial-of-Service Attacks" ACM Computer Communications Review(CCR), 2001

[5] Steve Gibson, "Distributed reflection denial of service", http://grc.com/dos/drdos.htm, Feb. 2002

[6] KDD Cup 1999 Data, http://kdd.ics.uci.edu/databases/kddcup99/task.htm

[7] S. Mukkamala. G. Janowski, A. H. Sung. "Intrusion Detection Using Neural Networks and Support Vector Machines" Proceedings of IEEE IJCNN. pp.1702-1707. May 2002.

[8] D. S. Kim, J. S. Park, "Network-Based Intrusion Detection with Support Vector Machines", ICOIN 2003, pp.747-756, 2003.

[9] Denning D. "An Intrusion-Detection Model" IEEE Transactions on SE, Number 2. pp. 222-250, Feb. 1997

[10] Joachmims T. "mySVM - a support Vector Machine" University Dortmund. 1999