IDS 알고리즘에 대한 탐지율 연구 비교

신경일*, 윤호상**, 신동일*, 신동규*
*세종대학교 컴퓨터공학과
**국방과학연구소
e-mail:sgi@gce.sejong.ac.kr

An Comparative Research of the Detection Rate of Intrusion Detection System Algorithms

Gyeong-Il Shin*, Hosang Yooun**, DongIl Shin*, DongKyoo Shin**
*Dept of Computer Engineering, Se-jong University
**Agency for Defense Development

요 약

인터넷의 발달로 인하여 네트워크 공격이 점차 발전되며 여러 가지 공격 기법들이 생겨나고 이러한 기법들은 혼합하여 사용하는 등 변칙적인 해킹기법들이 생겨나고 있다. 이로 인하여 침입 탐지 시스템(Intrusion Detection System, IDS)은 기존의 알려진 공격에 대해서만 탐지하고 변칙된 새로운 패턴의 공격을 탐지하지 못하는 경우가 생겨나고 있다. 이 문제에 적합한 해결책을 찾고자 여러 가지 알고리즘들이 연구되었고, 아직도 활발히 진행되고 있다. 본 글에서는 이러한 연구된 알고리즘들을 비교해 보았고 효율적인 방법을 제안한다.

1. 서론

최근 인터넷의 급격한 발달로 인하여 개인의 편리성뿐만 아니라 교육, 경제, 국방 등 다양한 곳에서 효율적으로 일들을 처리할 수 있게 되었다. 하지만 그로인하여 개인이나 집단이 네트워크를 통하여 많은 피해를 보기도 한다. 국내에서 발생한 대표적인 해킹 사례 중 하나로 2011년 3월에 국내 주요기관, 포털사이트 및 은행 등이 분산 서비스 거부 공격(Distributed Denial of Service attack, DDoS)을 받은 사건이 있겠다. 이 사건 이외에도 각종 해킹 관련된 사건들이 빈번하게 나타나면서 각 기관들은 막대한 손실을 입게 되면서 많은 매체를 통해 일반시민들에게 알려지게 되면서 사람들은 보안에 많은 관심을 가지게 되었다. 그로 인하여 공격들을 막기 위한 많은 연구들이 시작되었다. 또한 인터넷의 발달로 인하여 많은 해커들도 탄생하게 되면서 해킹의 기법들이 다양해지면서 여러 기법들을 혼합하여 새로운 공격 패턴들이 탄생하면서 공격을 탐지하지 못하고 있다. 본 글에서는 보안 기법 중에 사전에 공격을 탐지하여 방지하는 침입 탐지 시스템(Intrusion Detection System, IDS)에 대해서 관심을 가져, 네트워크 기반 침입 탐지 시스템(Network Intrusion Detection System, NIDS)에 대해 연구 해보았으며, 해킹 기법의 발달로 대응하여 변칙적인 공격 기법들을 탐지할 수 있도록 연구된 알고리즘들을 비교하여 효율적인 적용 방안을 제안한다.

2. 관련 연구

먼저 침입 탐지 시스템이란 사전에 공격들을 탐지하여 알려주는 것으로 침입을 탐지할 대상에 따라 호스트 침입 탐지 시스템(Host intrusion Detection System, HIDS)과 네트워크 침입 탐지 시스템(Network Intrusion Detection System, NIDS), 하이브리드 침입 탐지 시스템(Hybrid Intrusion Detection System)으로 크게 3가지 분류로 나뉘게 된다. HIDS는 주로 시스템 내부의 로그를 탐지하며, NIDS는 네트워크 기반의 공격 패턴 탐지를 하게 된다. Hybrid IDS의 같은 경우는 HIDS와 NIDS의 장점만을 취합한 것이다. 또한 탐지 방법에 따른 분류에 의한 비정상 행위(anomaly)탐지와 오용(misuse)탐지가 있다. 비정상행위 탐지 같은 경우는 정상적인 프로파일(profile)을 유지하다가 이에 어긋나는 행위가 발생할 때 탐지를 하여 알려주는 방법이고, 오용탐지는 알려진 취약점을 입력해두어 이미 알려진 공격이 시도될 때 이를 탐지하는 방법이다. NIDS의 경우 빠른 속도로 다양한 공격 기법들이 나오면서 이를 빠르게 대처할 수 있는 탐지 방법들이 필요하게 되었다. 그로 인하여 다양한 탐지 알고리즘들이 제안되었으며, 알고리즘 성능 평가를 할 때 주로 사용된 Dataset인 KDD Cup 1999에 의하면 주요 공격은 4가지로 분류된다 [1]. DoS(Denial of Service), R2L(unauthorized access from a remote machine), U2R(unauthorized access to local superuser privileges), Probe로 구별된다. KDD Cup training dataset에는 24개의 공격 유형이 포함되어 있고 test dataset에는 training dataset에 없는 14개의 유형이 추가로 포함된다.

LBG(Linde-Buzo-Gray) 알고리즘은 EM(Expectation Maximization)기반의 K-means 알고리즘의 단점을 보완

하고자 만들어졌다. 비균일 이진 분할에 의한 클러스터링 알고리즘은 초기 값 없이 이진 분할에 의해 클러스터링을 수행하므로 수행 속도가 빠른데, LBG 알고리즘은 K-means와 비균일 이진 분할의 장단점을 통해 만들어져 네트워크 침입 탐지에 적용하였다. LBG 알고리즘은 Y.Linde, A.Buzo, R.Gray 세 명의 이름 첫 자를 따서 LBG라 명명되였고, Accuracy Rate는 99%로로 K-means와 비교해서 높은 정확도를 보여주지만 데이터양이 증가하면 할수록 K-means에 비해 수행 시간은 훨씬 오래 소요된다[2].

K-means 클러스터링(clustering) 알고리즘은 침입 탐지에 관련된 Dataset의 데이터마이닝에 많이 사용되었던 방법으로 비계층적 클러스터링이며, 분할영역 k개를 지정하여 데이터를 분할하는 방법으로 k개의 그룹에는 k개만큼의 중심점이 존재하고 이 중심점들과 데이터들 사이의 거리를 계산하여 데이터를 분류한다. k-means 클러스터링 방법은 최근 많이 일어나는 공격 기법 중 DDoS와 Witty Worm에 대해 공격 탐지를 능력을 평가했을 경우에 DDoS의 경우 97%, Witty Worm은 94% TPR(true positive rate)를 보여줬으며 각 2%씩의 FPR(false positive rate)을 보여주었다[3].

Y-means 클러스터링 알고리즘은 계층적 클러스터링 방법으로 각 데이터를 하나의 클러스터로 설정한 다음 데이터들 간에 거리를 계산하여 인접한 위치에 존재하는 데이터들끼리 그룹을 형성하는 방법이다. 기존 k-means 알고리즘에는 큰 데이터를 클러스터링 하는데 종속성과 퇴행성이라는 두 가지 단점이 존재했다. 그 단점을 보완한 것이 y-means 알고리즘이다. 데이터셋(dataset)을 적절한 수의 클러스터로 자동분할 할 수 있는데, 이것은 침입 탐지를 할 때 장점 중 하나가 될 수 있으며, 해당 알고리즘은 따로 가공하지 않고도 로그 데이터를 training data로 사용할 수 있다. DR같은 경우 89.98%가 나왔고 FR같은 경우 1%가 나왔다[4].

SVM(support vector machine)과 BIRCH Hierarchical 클러스터링 방법은 training 시간을 대폭 단축 할 수 있었고, SVM도 우수한 성능도 보여주었다. 획득 된 SVM 모델이 네트워크 트래픽보다 정확하게 분류 할 수 있도록 training 세트에서 중요하지 않은 기능을 제거하기 위해 feature selection이 적용되었고, Accuracy Rater 95.7%로 다른 방법과 비교해보았을 때 DoS 및 Probe 공격 탐지에 각 99.5%, 97.5%로 우수한 성능을 보였다[5].

GA(generic algorithm)과 함께 퍼지 규칙을 이용한 것으로 경계가 모호한 상태인 침입 탐지에 적합하다고 제안한 연구결과도 있다. 퍼지논리란 불분명한 상태, 모호한 상태를 참 혹은 거짓의 이진 논리에서 벗어나 다치성을 표현하는 논리로, 경계가 잘 정의 되어 있지 않은 네트워크상에서 공격인지 아닌지 예측하기가 어려운데, 이 문제를 해결하기 위해 퍼지 규칙을 사용하여 침입 차단활동을 결정해 오작동을 줄일 수 있다. 해당 방법은 98.95%로 높은 탐지율을 보여주었지만, 7% 오탐율로 비교적 높은 수치가 나왔다[6].

GA와 SVM을 동시에 이용한 연구에서는 SVM 기반 침입 탐지 시스템을 개선하기 위해 GA를 사용하였다. 많은 연구들에서는 SVM을 분류기로 이용을 하고 있으며, SVM 기반 탐지속도와 처리속도 측면에서 IDS의 성능을 향상 시켰지만 최근 네트워크가 빨라짐에 따라 처리량 증가로 인해 개선이 필요로 했다. 또한 감사 데이터의 특징수가 많아지면서 방대한 양의 감사 데이터를 처리해야 하기 때문에 IDS 탐지율이 떨어 질 수 있다. 따라서 해당 논문은 GA와 SVM의 융합을 통해 최적의 feature set을 선택할 수 있을 뿐만 아니라 SVM 분류자에 대한 최적의 파라미터(parameter)를 파악 할 수 있다. GA는 대형 검색 공간에서 검색 할 때 가장 강력한 알고리즘 중 하나로 최적화 할 수 있는 함수의 형태의 수학적 제약이 거의 없다. GA와 SVM(Support Vector Machine)을 이용한 방법으로 해당 방법은 GA와 퍼지 규칙을 이용한 방법보다 조금 더 높은 99%의 탐지율을 보여주었다[7].

SOM(self-organizing map) 방법은 입력 벡터를 훈련 집합에서 일치되도록 가중치를 조정되는 인공 신경세포 격자에 기초한 비지도 학습 알고리즘이다. 해당 논문의 실험결과 3계층에서 89%의 탐지율과 4.6%의 오탐율을 보여주었다[8].

베이시안 네트워크(bayesian network)와 k-means 클러스터링을 이용한 방법은 99%의 높은 정확도를 보여주었다. 베이시안 네트워크는 몇몇 관측치만으로도 불확실한 침입의 가능성을 예측할 수 있는 추론 기능을 주로 이용하여 정상 또는 비정상을 판단한다. 또한, 학습 기능을 통해 변화하는 공격들의 특징이나 패턴을 추출해 낼 수 있기 때문에 다양한 학습 알고리즘의 적용을 통해 탐지 정확도를 높일 수 있다. 초기 네트워크가 생성되면 데이터 입력을 받게 되는데, 입력데이터는 test와 training 데이터의 두 가지가 있다. training 데이터로 판단되는 경우 이를 분석하여 k-means 클러스터링 기반의 노드 구간 범위와 가중치를 산출하고 산출된 가중치는 학습에 적용되어 보다 견고한 결과를 산출할 수 있게 된다. test 데이터로 판별된 데이터들은 침입 탐지를 수행하며 공격으로 의심되는 경우에는 관리자에게 이를 보고하고, 이상 데이터는 데이터 저장소에 보관되어 향후 학습용 데이터로 이용 될 수 있다[9].

LSTM(long short-term memory) 알고리즘은

RNN(recurrent neural network) 알고리즘의 종류 중 하나로 RNN의 장기의존성 문제를 해결하기 위해 고안된 방법이다. 해당 방법은 98.88%의 탐지율을 보여 주었으며, 정확도 또한 96.93% 높은 수치가 나왔는데, 오탐율이 무려 10.04%나 나타난다[10].

3. 본론

본 글에서는 공격을 탐지하는 방법론들을 조사하였고 그 중에서 9가지 유형의 논문들에서 제시한 방법들을 비교하여 보았다. 또한, 조사된 논문들에서 제시된 실험치들 중 가장 우수한 결과 값을 가져와 비교를 하였다.

표 1에서 제시된 탐지율 관련 식들은 각 논문들의 수식을 참고하여 대표적인 용어로 바꾸어 표기하였다. 먼저 Detection rate는 탐지율로 탐지한 공격 건수를 전체 공격 건수로 나눈 값이며, False rate 는 오탐율로 공격을 정상으로 판단한 건수를 공격으로 판단한 전체수로 나눈 값이다. Accuracy rate는 정확도로 각 분류별 정확도를 합하여 평균을 구한 값이며, True Positive rate는 침입을 침입이라 판정한 것이고, False Positive rate는 침입이 아닌데 침입이라 판단한 값이다. 또한 F-Measure는 단일한 측정치를 위해 침입으로 판단되는 클러스터 내에 실제로 얼마나 많은 침입데이터를 포함하는지를 평가하는 정확도(Precision)와 침입데이터의 전체 패킷중 얼마나 많은 패킷이 알고리즘에 의해 검출되었는지 평가하는 재현율(Recall)로 계산한 것으로 가중치 β=1로 설정되었다. T / N (True/Negative)은 비정상 중 정상 탐지율, F/N (False/Negative) 비정상 중 비정상 탐지율, F/P (False/Positive) 정상 중 비정상 탐지율, T/P (True/Positive) 정상 중 정상 탐지율이다.

$$F = \frac{(1+\beta)^2 * R * P}{\beta^2 (R+P)}$$

표 1을 보면 각 논문의 실험 결과를 비교해본 결과 정확도는 99%로 LBG, Beyesian Network와 K-means Clustering을 사용한 두 방법이 가장 높았고, 탐지율은 99%로 GA와 SVM을 이용한 방법이 가장 높았다. 그리고 오탐율은 1%로 Y-means Clustering이 가장 좋은 결과가 나왔지만 LSTM 같은 경우는 상당히 높은 수치인 10.04%다. 침입이 아닌데 침입이라고 판단된 경우는 0.7%로 SVM과 BIRCH hierachical Clustering이 가장 결과 값이 높았다. 전체적으로 결과 값을 비교해본 결과 모두 높은 탐지율을 보여주고 있다. 하지만 대부분 가장 많이 사용되는 Dataset인 KDD Cup 99 dataset을 사용하여 얻은 결과 값이라서 현시점의 공격 탐지에 우수한 결과를 보여줄 것이라고 예상하기는 어렵다. KDD Cup 99같은 경우 Knowledge Discovery Data Mining 대회에서 1999년도에 쓰인 데이터이기 때문에 굉장히 오래된 데이터이다. 또한 KDD Cup 99 데이터셋은 training 및 test 데이터에 공통적으로 많은 중복된 레코드가 포함되어 있고, training과 test 데이터가 각각 78%와 75%가 데이터가 중복된다. KDD Cup 99 데이터셋 중복성으로 학습 알고리즘은 잦은 공격을 빈번한 공격 기록에 편향 시키고 가끔씩 발생하지만 유해한 기록에 대해서는 잘못된 분류를 하는 경우가 발생한다. 또한 데이터의 중복성으로 인하여 실험결과가 대체적으로 높은 결과 값을 가지게 된다. 이러한 단점에도 불구하고 KDD Cup 99 Dataset을 쓰는 이유는 침입 탐지 관련 Dataset이 많이 공유가 되어있지 않기 때문이고, 많은 연구들에서 대부분 KDD Cup 99 dataset을 이용하다 보니 결과들의 수치가 대부분 비슷하게 제시되었다.

4. 결론

각 논문의 실험 결과를 비교해본 결과 대체적으로 높은 탐지율을 보였고 그 중 LBG방법과 Bayesian Network + K-mean Clustering 방법이 99%의 정확도로 가장 높은 탐지율을 보여주었다. 그리고 LSTM은 98.88%의 높은 탐

<표 1> 침입 탐지 방법론별 수치 비교

Method	Dataset	Accuracy rate	Detection rate	False rate	True Positive rate	False Positive rate	기타
LBG [2]	KDD cup 99	99%					F-Measure : 99%
K-means Clustering [3]	DDoS : Guido, Willy Worm : IWSEC Working Group				DDoS : 97% Willy Worm : 54%	DDoS : 3% Willy Worm : 2%	
Y-means Clustering [4]	KDD cup 99		88.96%	1%			
SVM + BIRCH hierachical Clustering [5]	KDD cup 99	95.7%				0.7%	
GA(Genetic Algorithm) [6]	KDD cup 99		98.95%	7%			
GA (Genetic Algorithm) + SVM(Support Vector Machine) [7]	KDD cup 99		99%				
SOM(Self Organizing Feature Map) [8]	KDD cup 99		88%	4.8%			
Beyesian Network + K-mean Clustering [9]	KDD cup 99	99%					T/N -> 23.35%, F/N -> 2.57%, F/P -> 0.48%, T/P -> 76.04%
LSTM(Long Short-term Memory) [10]	KDD cup 99	96.93%	98.88	10.04%			

지율을 보여주었지만, 오탐율 또한 너무 높아서 이 부분을 개선할 수 있는 방법을 찾아야 할 것 같다.

하지만 각 논문들에서 쓰인 Dataset같은 경우 오래된 데이터로 최근 발생하는 변칙적인 공격기법들을 탐지했을 때는 또 다른 결과를 가져올 것이다. 그리고 KDD Cup 99 Dataset의 문제점 중 중복성으로 인하여 논문들의 실험 결과가 대체적으로 높은 결과 값을 보여준다. 이러한 단점을 개선하기 위해 기존 KDD Cup 99의 중복된 레코드를 제거한 향상된 NSL-KDD와 UNB ISCX에서 제공해주는 Dataset을 이용하는 것을 제안한다. 최근 변칙된 공격기법들이 무수히 생기는 시점 최신화된 Dataset이 필요하다고 생각하기에 Snort로 IDS 환경을 구성하여 직접 다양한 공격을 통해 쌓인 네트워크 패킷을 통해 침입 탐지 Dataset을 만들어 각 종 방법론들을 실험해 보는 것 또한 제안한다.

Acknowledgement

본 연구는 방위사업청과 국방과학연구소의 지원으로 수행되었습니다(UD160066BD).

참고문헌

[1] Knowledge Discovery and Data Mining.(1999). KDD Cup 1999 Task Description[Online]. Available: http://kdd.ics.uci.edu/databases/kddcup99/task.html.

[2] 박성철, and 김준태. "LBG 알고리즘 기반 데이터마이닝을이용한 네트워크 침입 탐지율 향상." 한국지능정보시스템학회논문지 15 (2009): 23-35.

[3] 신동혁, et al. "K-평균 클러스터링을 이용한 네트워크 유해트래픽 탐지." 한국통신학회논문지 41.2 (2016): 277-284.

[4] Guan, Yu, Ali A. Ghorbani, and Nabil Belacel. "Y-means: A clustering method for intrusion detection." Electrical and Computer Engineering, 2003. IEEE CCECE 2003. Canadian Conference on. Vol. 2. IEEE, 2003.

[5] Horng, Shi-Jinn, et al. "A novel intrusion detection system based on hierarchical clustering and support vector machines." Expert systems with Applications 38.1 (2011): 306-313.

[6] Gomez, Jonatan, and Dipankar Dasgupta. "Evolving fuzzy classifiers for intrusion detection." Proceedings of the 2002 IEEE Workshop on Information Assurance. Vol. 6. No. 3. New York: IEEE Computer Press, 2002.

[7] Kim, Dong Seong, Ha-Nam Nguyen, and Jong Sou Park. "Genetic algorithm to improve SVM based network intrusion detection system." Advanced Information Networking and Applications, 2005. AINA 2005. 19th International Conference on. Vol. 2. IEEE, 2005.

[8] Gunes, H., et al. "On the capability of an SOM based intrusion detection system." Proceedings of the international joint conference on neural networks. 2003.

[9] 최보민, 이정식, and 한명묵. "베이지안 네트워크 개선을 통한 탐지율 향상의 IDS 모델." 한국지능시스템학회논문지 24.5 (2014): 495-503.

[10] Kim, Jihyun, et al. "Long Short Term Memory Recurrent Neural Network Classifier for Intrusion Detection." Platform Technology and Service (PlatCon), 2016 International Conference on. IEEE, 2016.