사이버 공격의 위험 지수 개발 및 피해 평가 우선 순위 산출 방안

윤현수*, 김용현**, 김동화**, 신동규*, 신동일*
*세종대학교 컴퓨터공학과
**국방과학연구소
e-mail : sug3nt7smd0djw@gmail.com

Development of Risk Index of Cyber Attack and Damage Assessment Priority Calculation Measures

Hyeonsu Youn*,YongHyun Kim**, DongHwa Kim**, Dongkyoo Shin*, Dongil Shin*
*Dept. of Computer Engineering, Sejong University
**Agency for Defense Development

요 약

정보통신 기술이 발전하면서 범국가적으로 사이버 환경은 사람들에게 없어서는 안 될 존재가 되었다. 이러한 사이버 환경은 간혹 악의적인 사이버 공격자로 인해 파괴되거나 손상된다. 본 논문에서는 사이버 공격에 대한 위험 지수를 개발 방안을 제시한다. 위험 지수에 대한 개발이 이루어진 후 위험 지수와 사이버 공격 횟수끼리 계산하여 값을 산출시킨다. 산출된 값은 곧 공격에 대한 중요도 점수로 표현되며, 이 값에 따라 사이버 전투 피해 평가 항목의 우선순위를 알 수 있게 된다.

1. 서론

정보통신 기술이 발전하면서 사이버 환경은 사람들에게 중요하게 인식되고 없어서는 안 될 존재가 되었다. 사이버 환경이 중요하게 인식되면서 대부분 국가의 정부들은 사이버 공격에 대한 관심이 높아졌다. 실제로 “사이버 공격을 당했다.” 라는 말은 이미 흔한 말이 되었다.

해외에서 발생한 대표적인 사이버 공격의 예로, 2017년 6월말에 우크라이나에서 발생한 이터널페트야(EternalPetya) 공격과 2017년 5월초에 전 세계에서 다발적으로 발생한 워너크라이(WannaCry) 공격이 대표적이다. 위 공격들은 랜섬웨어를 통해 공격한 사례인데 해결법이 미리 중요한 데이터를 정기적으로 백업, 백신을 최신 상태로 유지하거나 FSRM 기술로 랜섬웨어의 확장을 차단하는 방법 외에는 없다 [1]. 다음으로 국내에서 발생한 대표적인 사이버 공격의 예로, 2016년 9월에 발생했던 국군 사이버사령부 서버의 외부 해킹 사건이다. 국방부에 따르면 백신중계서버에 악성코드 감염 징후가 감지되었는데, 조사 진행 중 외부망과 내부망 일부 PC에서도 동일한 악성코드가 감염되었다고 발표했다 [2]. 이 사건으로 인해 군사 기밀 중 가장 중요한 1급 기밀인 ‘작전계획 5027’ 과 ‘작전계획 5015’ 가 유출되었다. 위의 사례들만 봐도 현재 사이버 공격에 노출된 국가가 입을 피해가 클 수 밖에 없다는 것을 알 수 있다.

이처럼 사이버 공격으로 인해 피해를 입은 국가는 해당 피해 규모가 얼마나 큰지, 어느 부분에 피해를 입혔는지 등을 인식해야 한다.

본 논문은 연구를 진행하는 동안 필요로 했던 기술이나 사전 지식을 설명하고 본격적으로 사이버 공격에 위험 지수를 적용하는 방법에 관한 내용과 사이버 공격 중 가장 먼저 평가해야하는 대상의 공격의 우선순위를 산출하는 방법에 대해 설명한다. 이후, 연구 결과에 대한 피드백과 향후에 어떻게 연구를 진행하면 좋을지에 대해 서술한다.

2. 관련 기술 및 연구

2.1 머신러닝 기술

머신러닝은 이미 실생활에서 많이 사용되어지는 추세이다. 머신러닝은 인공지능의 분야 중 하나로, 컴퓨터에게 데이터를 학습시키고 학습시킨 내용을 바탕으로 예측 및 분류와 같은 작업을 수행하도록 하는 것을 의미한다 [3]. 머신러닝 기술이 사용되어지는 이유로는 기존에 손이 많이 가고 복잡한 문제를 훨씬 간단한 코드로 해결이 가능하고 복잡하고 크기가 큰 데이터에서 의미 있는 내용을 찾아낼 수 있다는 것이다.

머신러닝의 실제 적용 사례로는 크게 감정 분석, 사기 탐지, 로그 분석 및 위험 탐지 등으로 나눌 수 있다. 본 논문에서 집중해야 할 적용 사례는 실시간 위험 탐지와 로그 분석이다. 위험 탐지는 머신러닝을 이용한 위험 탐지방법 [4]을 사용하거나 이전에 연구했었던 CDC 프로그램 [2], 또는 대표적인 사이버 공격 데이터셋인 KDD [5]를 사용해도 무방하다. 머신러닝을 통해 위험을 탐지하는 방법은 최근 들어 많이

사용되어지고 있다. 예를 들어, 이상 행동 탐지 시스템을 사용하면 어떤 작업이 산출해야했던 결과가 아닌 예상과는 다른 결과가 산출되면 이것을 사용자에게 보고하는 방법이 머신러닝을 통해 이루어진다. 또한 CDC 프로그램은 시스템의 정보를 공격 받기 전과 공격 받은 후로 각각 저장시켜 비교하여 공격을 탐지해 준다. 이어서 KDD 데이터셋은 DARPA'98 IDS 평가 프로그램에 의해 필터링된 데이터를 기반으로 제작된 것이다 [5, 6, 7].

2.2 사이버 공격 유형

사이버 공격의 유형은 크게 Degradation, Interception, Interruption, Fabrication, Modification, Unauthorized use 6 가지로 나눌 수 있다 [8]. 아래의 표 1 은 사이버 공격 유형이 시스템에 어떤 영향을 미치는지 설명하고 본 연구의 진행 방향에 맞게 공격 유형의 수정과 보완이 이루어진다.

<표 1> 사이버 공격 유형 정리

공격 유형	시스템에 미치는 영향	비고
Degradation	시스템에서 이루어지는 작업속도가 현저하게 떨어지며 정보 생산의 정확성 등 시스템의 전반적인 성능이 하락	
Interruption	일정 시간 및 기간 동안 시스템의 특정 기능이나 모든 기능이 중단되어 복구되기 전까지 사용이 불가능	▲
Modification	시스템의 몇몇 정보나 모든 정보가 변경되어 사용자가 의도하지 않은 잘못된 결과가 발생	
Fabrication	사용자가 입력하지 않은 알 수 없는 거짓된 정보가 시스템에 입력되어 원하던 작업의 결과 확인을 방해	
Interception	시스템 상에서 사용자가 작업 중이었던 프로세스 및 작업 환경 등의 정보를 비 인가된 침입자에게 노출	▲
Unauthorized use	시스템 내 침입자가 삽입 해놓은 어떠한 프로그램 같은 것들로 인해 예기치 않았던 알 수 없는 결과가 발생	▲

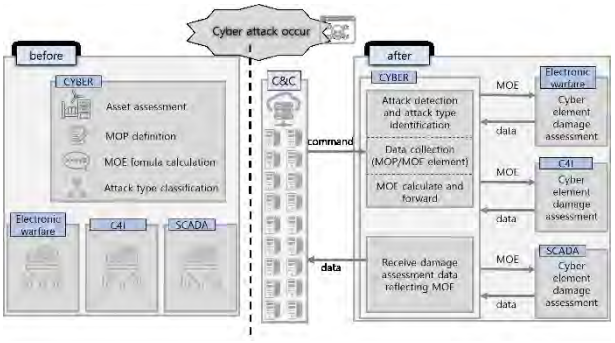
위와 같은 공격 유형에 대한 하위 공격 유형을 조사하였고 위의 6 가지 공격 유형들 중 하위 공격 유형이 될 수 있는 공격 유형들이 존재했다. Degradation, Fabrication, Modification, 이 세 가지 공격 유형은 Unauthorized use 공격에 포함된다. Unauthorized use 가 발생할 경우 Degradation 으로 인해 발생하는 시스템의 성능 하락이 발생할 수 있고 Fabrication 으로 인해 발생하는 거짓된 정보 입력으로 인한 잘못된 결과 확인, Modification 으로 인한 사용자가 의도하지 않았던 잘못된 결과 도출이 발생할 수 있다. 이에 근거하여 본 논문에서는 위 표의 6 가지 공격 유형 중 3 번 째 열에 표시가 되어 있는 Interruption, Interception, Unauthorized use 를 사용한다.

2.3 전투 피해 평가 척도

본 논문에서는 전투에서 사용하는 피해 평가 척도 중 잘 알려진 MOP 를 식별하기 위한 연구가 진행된다. MOP 는 성능지표라고 불리며 Measure of Effectiveness 라고 불리는 효과지표를 입력으로 하여 요구사항 분석을 통해 시스템 운용과 관련된 속성을 특징짓는 척도이다 [9]. 그러나 MOP 를 사이버 전투에 적용시키기엔 그와 관련한 문서들이나 기술 등이 상당히 부족하다. 따라서 본 연구의 진행을 통해 MOP 를 식별하기 위해 사이버전 피해 평가의 우선순위를 도출하면 사이버전에서 사용가능한 MOP 요소를 산출하는데 도움이 될 것으로 예상한다.

3. 사이버전 피해 평가를 위한 사전 준비

사이버 공격이 발생한 후 피해를 평가를 하기 위해서는 아래 그림 1처럼 현재 보유한 자산의 평가를 하여 위험 지수를 매겨져 있어야한다. 이후 C&C 서버의 명령을 받아 사이버 공격의 종류를 분류하고 MOP, MOE 요소에 대한 데이터를 수집하여 MOE를 계산해야한다. 그 후 각종 환경의 전투에서 쓰이는 장비에 대한 MOE의 피해 평가 데이터를 받아 C&C 서버로 전송해야한다.



(그림 1) 사이버 전투 피해 평가 과정

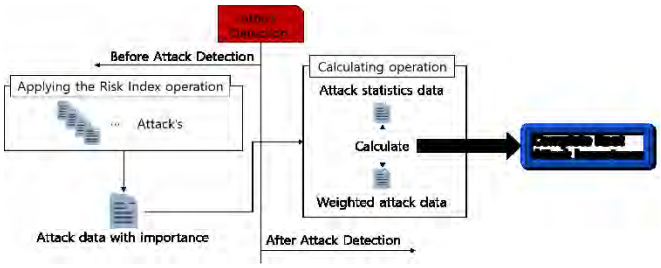
본 연구에서는 그림 1의 과정 중 Attack type Identification과 Data collection 과정에 중점을 둔다. 이 두 가지 단계가 진행되기만 하면 MOP가 식별되어 MOE 산출 작업을 마지막으로 사이버전의 피해 평가가 원활히 이루어지게 된다. 중점을 둔 두 과정을 좀 더 구체적으로 표현한 것은 그림 2와 같다. 공격이 탐지되기 전 먼저 각 공격에 대한 위험 지수 적용 작업이 이루어져야 한다. 위험 지수 적용을 위해 2장에서 결정했던 3가지 공격 유형을 세분화시켜 하위 공격 방법들까지 트리화 시킬 필요가 있다. 트리화가

완료된 공격에 대해 각각 위험도를 설정하여 위험한 공격일수록 높은 점수를 매긴다. 점수 매김 작업이 완료되면 그림 3의 Root Attack Category의 항목인 세 가지 공격에 앞선 작업의 완료로 이루어진 공격 점수로 위험 지수를 적용시켜준다.

$$\frac{\sum_1^n Attack's(n)}{n}$$

(수식 1) 공격 점수의 위험 지수 계산법

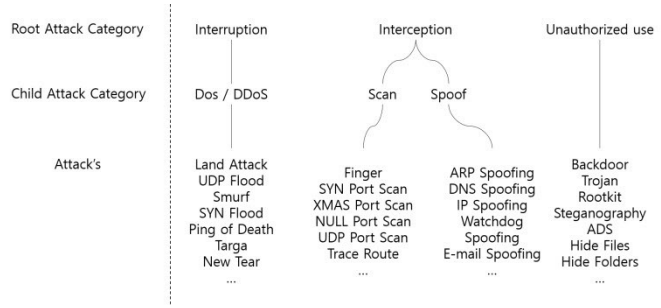
수식 1에 의해 위험 지수가 계산되면 그것을 데이터 형태로 저장시켜 준다. 이 후, 사이버 공격이 탐지되면 미리 계산해 놓았던 각 공격에 대한 위험 지수 데이터와 탐지된 공격과 그 횟수에 대한 통계를 계산하면 Root Attack의 중요도를 산출해낼 수 있다. Root Attack의 중요도가 산출되면 이를 통해 사이버 공격에 대한 피해 평가의 우선순위를 정할 수 있게 된다.



(그림 2) Root Attack 의 중요도를 정하는 과정

3.1 사이버 공격 위험 지수 적용 방안

방어자 차원에서 사이버 공격을 당하게 되면 그로 인해 시스템 상의 정보가 변조되거나 어떠한 작업의 결과가 잘못 도출될 가능성이 있다. 사이버 공격을 분류하기 위해서 먼저 2장에 있는 표 1의 공격 유형을 트리화하여 세분화할 필요가 있다. 아래의 그림 3는 공격 유형을 트리화하여 세분화 시킨 결과이다. 사이버 상에서 이루어지는 모든 공격을 포함시키기에는 그 수가 너무 많아 대표적인 공격만을 하위 공격 유형으로 설정해놓았다.



(그림 3) 공격 유형 세분화 및 트리화

그림 3의 좌측 하단을 보면 Attack's 행에서 공격명이나 방법을 확인할 수 있다. Attack's에 해당하는 공격이 발견되면 각각 공격의 Root Attack Category인 Interruption, Interception, Unauthorized use로 분류를 해야 한다. 따라서 KDD 데이터셋을 머신러닝 알고리즘을 적용시켜 각각의 공격을 탐지하거나 공격 탐지 프로그램으로 공격을 탐지한 뒤, 탐지한 각 공격에 대한 위험 지수를 설정해야한다.

3.2 사이버전 피해 평가 우선순위 도출

사이버 공격에 대한 위험 지수 적용이 성공적으로 이루어지면 그 다음 작업으로 방어자가 당한 공격방법과 그 횟수를 통계적으로 알아내야한다. 예를 들어 Interception 공격은 Land Attack, Smurf, UDP Flood 등의 공격의 상위 공격법이 되므로 Land Attack 등의 공격이 들어오면 이 공격이 발생한 횟수를 알아내야한다는 것이다. 아래 수식 2의 ar은 각 공격에 대한 위험 지수 데이터(수식 1)이고 sc는 통계 데이터에서 산출한 공격 횟수를 의미한다.

$$Importance\ score = \sum_1^n ar(n) * sc(n)$$

(수식 2) 중요도 점수 산출 계산법

수식1과 수식2의 모든 계산과정을 완료하면 이제 그림 3의 Root Attack의 중요도 점수를 구할 수 있고 중요도 점수를 보고 평가 순위도 구할 수 있게 된다.

4. 실험 결과

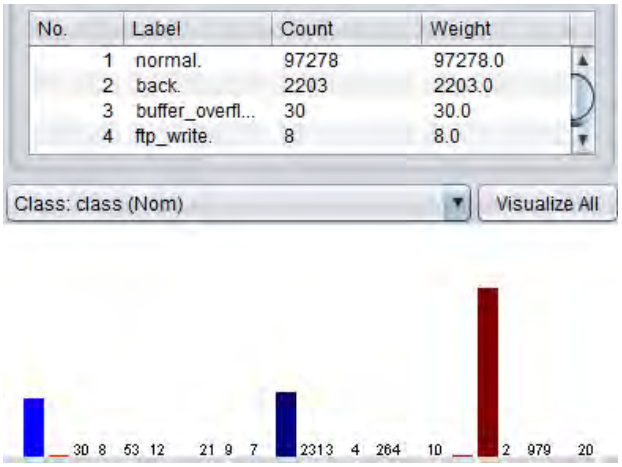
실험한 내용은 다음과 같다. 먼저 그림 3처럼 공격 유형에 대한 트리화를 적용시켜 하위 공격방법들을 세분화 시켰다. 이어 이러한 공격 방법들로 인해 시스템의 어떤 부분이 피해를 입는지, 피해 정도에 따라 위험 지수는 얼마나 되는지를 설정해 보았고 그에 대한 결과의 예는 아래의 표 2와 같다.

<표 2> 공격에 대한 위험 지수 설정의 예

Attack's 항목	공격 목표	위험 지수
Smurf	cpu 과부하 및 가용성 침해	Ⅱ
ARP Spoofing	타겟 시스템에서 이루어지는 트래픽을 가로채거나 변조하여 전달	Ⅲ
SYN Flood	서버의 자원 접근 방해 및 가용성 침해	Ⅱ
Backdoor	보안 액세스 컨트롤 우회 및 시스템 강제 엑세스	Ⅲ
Trace Route	네트워크 경로 추적 및 포트 스캔	Ⅰ
⋮	⋮	⋮

위험 지수는 로마 숫자 (Ⅰ, Ⅱ, Ⅲ) 로 나누었다. 숫자가 높을수록 위험하다는 것을 의미하고 Ⅰ부터 Ⅲ까지 차례대로 1점부터 3점으로 점수를 부여했다.

각각의 위험 지수 및 점수는 R. Ostler의 연구 [10]에 근거하여 부여하였다. 이런 식으로 Attack's의 항목에 위험 지수의 점수를 모두 부여해 준 뒤, 수식 1을 적용시켜 그림 2의 공격 탐지 전에 이루어지는 위험 지수 설정 및 적용 작업을 통해 중요도가 적용된 공격 데이터를 생성한다. 이 다음 작업으로 그림 2의 공격 탐지 후, 방어자 차원에서 공격 받은 횟수에 대한 데이터를 산출해야한다. 이에 의해 본 연구에서 KDD의 네트워크 탐지 데이터셋과 [2]에서 사용한 CDC 프로그램을 사용하여 통계를 내보았다. 빠르게 통계를 확인하기 위하여 아래의 그림처럼 WEKA 프로그램을 사용하여 차트를 참고하였다.



(그림 4) 방어자 차원에서 받은 공격에 대한 통계

위와 같은 통계 데이터와 앞서 산출했던 표 2의 위험 지수 데이터끼리 계산하여 Root Attack의 중요도 점수를 산출해낼 수 있다. 아래의 표 3은 각 Root Attack에 대한 중요도를 수식 2로 산출해낸 결과를 보여준다.

<표 3> Root Attack 의 중요도 점수 산출 결과

Root Attack Category	중요도 점수	평가 순위
Interruption	13,209.142	**1**
Interception	6,063.000	**2**
Unauthorized use	5,664.857	**3**

앞서 위험 지수 점수가 높을수록 위험했으므로 중요도 점수가 높으면 사이버 전투에서의 피해 평가에서 가장 먼저 평가 대상이 되야한다는 결론이 나왔다.

5. 결론 및 향후 연구 방향

이번 연구를 진행함으로써 기존에 존재하지 않았던 각 공격에 대한 위험 지수 계산방법, 그 지수를 점수화 하는 방법 및 이와 같은 계산으로 사이버전의 피해 평가 시 우선적으로 평가해야 할 공격에 대한 피해의 순위를 알 수 있게 되었다. 현재 공격을 탐지하여 데이터셋 형태로 저장하는 프로그램이나 방법이 부족하여 오래된 공격 탐지 데이터셋이나 심플한 프로그램을 사용했다는 점이 살짝 아쉽다. 또한 이번 연구에서는 평가해야 할 공격의 우선 순위를 설정하는 방법만을 제안했지만 향후 연구에서는 공격으로 인해 발생한 시스템의 피해 부분을 보고 우선순위를 정할 수 있도록 하면 더욱 좋은 연구가 될 것으로 기대된다.

ACKNOWLEDGEMENT

본 연구는 방위사업청과 국방과학연구소의 지원으로 수행되었습니다(UD160066BD).

참고문헌

[1] Gandhi., Krunal A., "Survey on Ransomware: A New Era of Cyber Attack." International Journal of Computer Applications, vol. 168, no. 3, 2017.

[2] H. S. Youn., Y. H. Kim., et al., "System Information Comparison and Analysis Technology for Cyber Attacks." Proc. Korea Information Processing Society, vol. 24, no. 1, pp. 198-200, Apr. 2017.

[3] Witten, Ian H., et al., "Data Mining: Practical machine learning tools and techniques." 4th ed., Morgan Kaufmann, 2016.

[4] R. Sommer., V. Paxson., "Outside the closed world: On using machine learning for network intrusion detection." In: Security and Privacy (SP), 2010 IEEE Symposium on. IEEE, pp. 305-316, 2010.

[5] Tavallaee., Mahbod., et al., "A detailed analysis of the KDD CUP 99 data set." Computational Intelligence for Security and Defense Applications, 2009. CISDA 2009. IEEE Symposium on. IEEE, pp. 1-6, 2009.

[6] S. J. Stolfo., W. Fan., W. Lee., et. al., "Costbased modeling for fraud and intrusion detection: Results from the jam project," discex, vol. 02, p. 1130, 2000.

[7] R. P. Lippmann., D. J. Fried., I. Graf, et. al., "Evaluating intrusion detection systems: The 1998 darpa off-line intrusion detection evaluation," discex, vol. 02, pp. 1012, 2000.

[8] Musman., Scott., et al.. "Computing the impact of cyber attacks on complex missions." Systems Conference (SysCon), 2011 IEEE International. IEEE, pp. 46-51, Apr. 2011.

[9] T. K. Kim., H. J. Lee., W. S. Cho., et. al., "Research on Matrix of Measurement of Effectiveness(MOEs) and Measurement of Performance(MOPs) for Cyber Threat and Defense Behavior on Cyberwarfare Simulation." Conf. Korean Institute of Industrial Engineers, pp. 3114-3118, Apr. 2016.

[10] R. Ostler., "Defensive cyber battle damage assessment through attack methodology modeling. " AIR FORCE INST OF TECH WRIGHT-PATTERSON AFB OH GRADUATE SCHOOL OF ENGINEERING AND MANAGEMENT, 2011.