https://doi.org/10.33778/kcsa.2020.20.3.115

공개출처정보의 정량화를 이용한 인공신경망 기반 사이버위협 예측 모델★

이 종 관*, 문 미 남**, 신 규 용***, 강 성 록****

요 약

사이버공격은 최근 몇 년간 더욱 더 진화하고 있다. 이렇게 고도화, 정교화된 사이버위협에 대응하기 위한 최선의 대책 중 하나는 사이버 공격을 사전에 예측하는 것이다. 사이버위협을 예측하기 위해서는 많은 정보와 노력이 요구되며 최근 정보획득의 핵심인 공개출처정보(Open Source Intelligence, OSINT)를 활용한다면 사이버위협을 보다 정확히 예측할 수 있을 것이다. 공개출처정보를 활용하여 사이버위협을 예측하기 위해서는 공개출처정보로부터 사이버위협 데이터베이스의 구축과 구축된 DB에서 사이버위협을 평가할 수 있는 요소를 선정하는 것이 선행되어야 한다. 이를 위해 데이터마이닝 기법을 활용하여 DB를 구축하고, 축적된 DB 요소 중 핵심요소에 대한 중요도를 AHP 기법으로 분석한 선행연구를 기초로 하였다. 본 연구에서는 공개출처정보로부터 축적된 사이버공격 DB를 활용하여 사이버위협을 정량화할 수 있는 방안을 제시하고 인공신경망을 기반으로 한 사이버위협 예측 모델을 제안한다.

Cyber Threats Prediction model based on Artificial Neural Networks using Quantification of Open Source Intelligence (OSINT)

Jongkwan Lee*, Minam Moon**, Kyuyong Shin***, Sungrok Kang****

ABSTRACT

Cyber Attack have evolved more and more in recent years. One of the best countermeasure to counter this advanced and sophisticated cyber threat is to predict cyber attacks in advance. It requires a lot of information and effort to predict cyber threats. If we use Open Source Intelligence(OSINT), the core of recent information acquisition, we can predict cyber threats more accurately. In order to predict cyber threats using OSINT, it is necessary to establish a Database(DB) for cyber attacks from OSINT and to select factors that can evaluate cyber threats from the established DB. We are based on previous researches that built a cyber attack DB using data mining and analyzed the importance of core factors among accumulated DG factors by AHP technique. In this research, we present a method for quantifying cyber threats and propose a cyber threats prediction model based on artificial neural networks.

Key words : Open Source Intelligence(OSINT), Artificial Neural Network, Cyber Threats, Prediction model

접수일(2020년 6월 29일), 수정일(2020년 9월 14일), 게재확정일(2020년 9월 27일)

★ 본 논문은 2018년 국군사이버사령부(11-1290000-000742-01)와 육군사관학교 화랑대연구소의 지원에 의해 연구되었음

* 육군사관학교 컴퓨터학과(주저자)
** 육군사관학교 수학과(교신저자)
*** 육군사관학교 컴퓨터학과
**** 육군사관학교 심리경영학과

1. 서 론

IT의 급격한 발전은 초연결, 초지능의 기능을 산업 전 분야에 확산시켰으며 우리 삶에 많은 변화를 가져왔다. 인공지능, 클라우드, IoT(Internet of Things), 빅데이터 등의 기술을 이용하여 기존에 없던 새로운 서비스들이 창출되고 있다. 그런데 이러한 발전의 근간을 흔드는 것이 사이버 위협이다. 이에 많은 연구자 및 보안업체에서는 이를 사전에 예방하기 위한 다양한 연구를 진행하고 있다. 특히 정보의 생산, 유통, 처리 과정에서 발생되는 다양한 데이터의 분석을 통해 사이버 공격을 예측하는 연구가 최근 각광을 받고 있다.

사이버위협 예측 범주는 (1) 공격투영(Attack projection) 및 공격의도 인지(Attack Intention Recognition), (2) 침입 예측(Intrusion Prediction), (3) 네트워크 보안 상황 예측(Network Security Situation Forecasting)으로 구분할 수 있다. 먼저, 공격투영은 2001년 Geib와 Goldman의 연구[3]에서 처음 사용된 용어로 다중 목적, 관찰의 실패 또는 관찰되지 않은 행동 등과 같은 문제를 식별하고 공격계획을 인지하기 위해 도입된 개념이다. 궁극적으로 공격자가 다음 단계에서 무엇을 할지를 예측하고자 하는 것이다. 즉, 공격자의 최종 공격 목적이 무엇인지를 예측하고자 하는 것이다[4]. 둘째로 침입 예측은 가장 일반적인 사이버 위협 예측 범주로서 어떤 종류의 사이버공격이 언제, 어디서 일어날 지를 예측하는 것이다. 이는 취약점 예측, 공격확산 예측, 다단계 공격 예측 등의 다양한 공격형태를 포함한다[5]. 마지막으로 네트워크 보안 상황 예측은 우리가 통제하는 시스템에 대한 보안에 중점을 둔 예측으로 공격 투영 및 침투 예측이 공격 자체 또는 공격자에 초점을 둔 것과 대조된다. 네트워크 보안상황 예측의 핵심은 네트워크 보안상황을 정량화하는 것이다. 그 결과 입력값 및 예측값이 숫자의 형태여서 대부분의 보안 상황 예측 모형은 연속형 모형으로 연구되고 있다[6]. 본 연구에서는 공개출처정보의 특성으로 인해 특정 공격에 대한 단계 또는 목적보다 일반적인 네트워크 상황의 사이버위협을 예측하기에 적합하여 사이버위협 예측 범주 중 네트워크 보안 상황 예측에 초점을 두고, 사이버 위협을 정량화하는 방법과 이를 활용하여 사이버 위협을 예측하는 방법론을 제시할 것이다.

예측 범주를 달성하기 위해 그래프 이론, 시계열 분석, 딥러닝 및 머신러닝 등과 같은 다양한 방법론이 사용되고 있다. 사이버위협 예측 모형에 적용된 방법론을 분류하는 다양한 방법이 존재하나 일반적으로 배경지식에 따라 이산형 모형, 연속형 모형, 머신러닝 모형 등으로 구분할 수 있다. 먼저, 이산형 모형 기반의 방법론에는 그래프(Graph) 모형[7], 베이지안 네트워크(Bayesian Network) 모형[8], 마코브(Markov) 모형[9], 게임이론(Game Theory) 모형[10] 등이 있다. 이산형 모형 기반 방법론의 핵심은 공격단계를 수립하고 각각의 공격단계를 상태로 정의한 후 상태전이를 통해 다음 공격 상황을 예측하는 방법으로 공격투영 및 침입 예측을 위해 많이 사용되고 있다. 연속형 기반 방법론은 시계열(Time Series) 분석[11]과 그레이 모형(Grey Model)[12] 등이 있으며, 이와 같은 방법론을 이용하여 공격의 조합, 공격의 정도, 공격의 횟수 등의 예측이 가능하다. 마지막으로 머신러닝 기반의 방법론에는 인공신경망(Artificial Neural Network) 모형[13], SVM(Support Vector Machine) 모형[14], 데이터 마이닝(Data Mining) 모형[15] 등이 있다. 본 연구에서는 최근 급격히 발전하고 있는 인공신경망을 활용하여 시계열 데이터를 예측하는 방법을 제안한다.

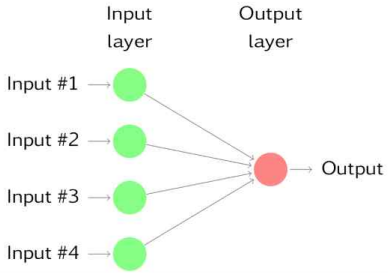
본 연구는 사이버 공격 데이터베이스와 사이버 위협 평가요소의 중요도 분석 연구결과[1, 2]를 토대로 (1) 사이버 위협 정량화 방법을 결정하고, (2) 인공신경망을 활용하여 사이버위협 예측 모형 및 적용결과를 제시한 뒤, (3) 마지막으로 향후 활용 방안을 모색한다.

본 논문의 구성은 다음과 같다. 먼저 2장에서는 사이버위협 예측 모형 방법론인 인공신경망(ANN)을 설명한다. 3장에서는 사이버위협 예측 모형의 개발을 위한 선행 조건인 사이버위협 정량화 방안

을 설명한다. 4장에서는 사이버위협 예측 모형 적용 결과를 제시한다. 마지막으로, 5장에서는 본 연구의 결론 및 향후 연구방향을 제시한다.

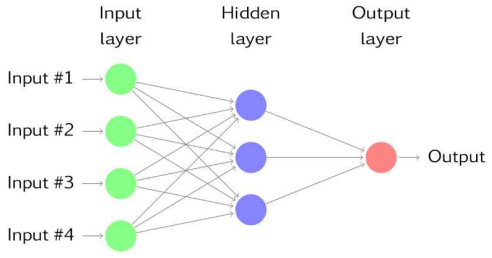
2. 인공신경망 모델

신경망은 뉴런(neuron)들 간의 계층화된 네트워크라 할 수 있다. 입력값은 최하단 계층을 형성하고, 결과값은 최상단 계층을 형성한다. 그리고 다수의 뉴런으로 구성된 중간 계층이 있을 수 있다. 가장 단순한 신경망은 은닉 계층이 없는 것으로 이것은 선형 회귀(linear regression)와 동일하다. (그림 2-1)은 4개의 뉴런으로 구성된 입력층과 1개의 뉴런으로 구성된 출력층의 신경망을 나타내며 이것은 앞서 설명한 바와 같이 선형 회귀와 동일하다. 그리고 입력값에 곱해지는 계수를 가중치라고 하고, 결과값(예측값)은 입력값의 선형 결합으로 구해진다. 다시 말해, 입력값에 가중치가 곱해진 결과들의 합이 출력층의 출력값이 되는 것이다. 가중치는 MSE(Mean Square Error)와 같은 비용함수(cost function)를 최소화하는 학습 알고리즘(learning algorithm)에 의해서 선택된다[16].



(그림 2-1) 은닉계층이 없는 신경망의 예

만약 (그림 2-1)에 중간계층(은닉계층)이 추가되고 비선형함수가 사용되면 해당 신경망은 비선형망이 된다. (그림 2-2)는 각 계층의 뉴런들이 이전 계층으로부터 입력값을 받는 다중계층 전방향 신경망(FFNN: Feed Forward Neural Network)을 나타낸다. 특정 계층에서 뉴런들의 출력값은 다음 계층의 입력값이 된다. 각 입력값들은 적절한 가중치가 곱해지고 선형적으로 더해진 후 시그모이드(sigmoid) 함수와 같은 비선형 함수에 의해서 수정된다. 이것은 극단적인 입력값이 전체 신경망에 미치는 영향을 줄여 신경망이 이상값(outlier)에 강인한(robust) 특성을 갖도록 하기 위함이다.



(그림 2-2) 다중계층 feed-forward 신경망

신경망의 파라미터인 가중치와 바이어스는 학습 데이터로부터 학습된다. 가중치는 초기에 임의의 값으로 선택되고 관측된 데이터에 따라서 지속적으로 갱신된다. 결과적으로 신경망에 의해서 생성된 결과값(예측값)에는 임의성(randomness)이 존재하게 된다. 따라서 신경망은 보통 임의의 여러 시작점(즉, 가중치)에서 학습을 하고 그 결과값들의 평균을 최종 결과값으로 한다. 한편, 은닉계층의 수와 각 은닉계층에 존재하는 뉴런의 수 등은 학습에 의해서 최적화될 수 없고 사용자에 의해서 사전에 정의되어야 한다[17, 18].

시계열 데이터에서 이전 시간의 데이터들은 신경망의 입력값으로 사용될 수 있는데, 이것이 신경망 AR 또는 NNAR 모델이라고 한다[19]. 본 연구에서는 한 개의 은닉 계층을 가진 FFNN를 고려하며, 은닉 계층에 k개의 뉴런이 있고 p개의 지연된 입력값을 가지는 신경망을 $NNAR(p,k)$로 표현한다. 예를 들어, $NNAR(9,5)$ 모델은 은닉 계층에 5개의 뉴런이 있고, 출력값 y_t를 예측하기 위해 9개의 과거값($y_{t,1}$, …, $y_{t,9}$)들이 입력값으로 사용됨을 나타낸다. $NNAR(p,0)$은 $ARIMA(p,0,0)$과 동일하다. 다만, NNAR에서는 데이터가 정상적(stationary)일 필요는 없다. 바로 다음 값을 예측하기

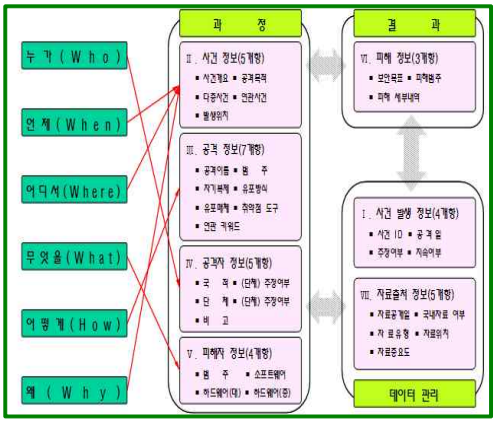
위해, 과거 입력값들이 사용되며 두 번째 다음 값을 예측하기 위해 첫 번째 예측된 값이 다시 입력값으로 사용된다. 이러한 과정은 요구되는 예측범위를 모두 계산할 때까지 반복된다.

3. 사이버위협 정량화 방안

이 장에서는 사이버위협 예측 모형을 개발하기 위해 사이버공격 데이터베이스를 활용한 사이버위협 정량화 방법을 제시한다.

3.1 사이버공격 DB 구조

(그림 3-1)에서 보듯이 본 연구의 선행연구로부터 구축된 사이버공격 데이터베이스는 사건발생 정보, 사건 정보, 공격 정보, 공격자 정보, 피해자 정보, 피해 정보, 그리고 자료출처 정보 등 7가지 범주 33개 변수로 구성되어 있다[1].



(그림 3-1) 사이버공격 데이터베이스 구성[1]

사이버위협을 정량화하기 위해 본 연구의 선행연구에서 사이버 위협에 직접적으로 영향을 미치는 ① 공격목적, ② 공격범주, ③ 공격대상, ④ 공격 지속성, ⑤ 공개출처정보의 빈도 등 총 5개 요소를 사이버 위협 평가요소로 선정하였고, 계층분석적 의사결정방법(AHP)을 적용하여 각 평가요소의 중요도를 분석하였다[2].

3.2 사이버위협 요소별 정량화

사이버위협을 정량화하기 위한 많은 연구에서 사이버위협을 공격범주, 취약점, 자산의 곱으로 정의하고 있다. 하지만 본 연구에서는 공개출처정보를 활용한 데이터베이스를 기초한 사이버위협을 정량화하기 때문에 특정 자산에 대한 사이버공격이 아닌 광범위한 사이버공격에 대한 위협도 계산을 위한 정량화 방법이 요구된다. 따라서 사이버위협을 정량화하기 위해 선행연구에서 제시한 평가요소 5개 항목을 사이버위협 정량화 변수로 지정하고, 각 변수에 대한 가중치를 활용하여 사이버위협 정량화 계산식을 다음과 같이 정의한다.

$$Z = X \times (w_1 Y_1 + w_2 Y_2 + w_3 Y_3 + w_4 Y_4)$$

여기서, Z는 사이버위협, X는 피해정도, Y_1은 공격목적, Y_2는 공격범주, Y_3은 공격대상, Y_4는 공격 지속성이다. w_i는 AHP 기법을 활용하여 계산된 Y_i의 중요도로 <표 3-1>과 같다.

<표 3-1> 중요도 계산결과

구분	w_1	w_2	w_3	w_4
중요도	2.98	1.07	2.47	1

공격목적(Y_1)은 ① 경제적 목적, ② 데이터 탈취, ③ 업무 방해, ④ 시스템 파괴, ⑤ 정치적 목적, ⑥ 기타와 같이 총 6가지 범주로 구분하였고, 각 항목에 대한 점수는 AHP 기법에서 산출된 비율을 이용하여 <표 3-2>와 같이 값을 산정하였다. 단, 목적이 식별되지 않을 경우 기하 평균값을 적용하였다.

<표 3-2> 공격목적 정량화 값

구분	값	구분	값	구분	값
경제적 목적	2.38	업무 방해	1.72	정치적 목적	3.28
데이터 탈취	3.22	시스템 파괴	4.82	기타	1.00

공격범주(Y_2)는 ① 스파이웨어, ② 랜섬웨어, ③ 봇 넷, ④ 백도어, ⑤ 와이퍼, ⑥ 기타 등 총 6가지 범주로, 공격대상(Y_3)은 ① 불특정다수, ② 일반업체, ③ 금융권, ④ 정부기관, ⑤ 사회기반시설, ⑥ 기타 등 총 6가지 범주로 구분하였고 각 항목에 대한 점수는 각각 <표 3-3>, <표 3-4>와 같다.

공격지속성(Y_4)은 지속성이 있는 경우 2를 지속성이 없는 경우에는 1의 값을 설정하였다. 피해정도(X)는 높음과 낮음으로 구분하였고, 일반적으로 피해정도가 클수록 공개출처정보의 빈도가 증가하므로 피해정도를 나타내는 변수로 공개출처정보의 빈도수를 활용하였다. 빈도가 모든 빈도수의 기하평균보다 클 경우 피해정도가 높음으로 하고 작을 경우 낮음으로 하였다. 각 경우의 값은 높음의 경우 2, 낮음의 경우 1의 값을 갖는다.

<표 3-3> 공격범주 정량화 값

구분	값	구분	값	구분	값
스파이웨어	3.23	봇 넷	3.27	와이퍼	1.80
랜섬웨어	2.87	백도어	4.50	기타	1.00

<표 3-4> 공격대상 정량화 값

구분	값	구분	값	구분	값
불특정다수	1.21	금융	2.62	사회기반시설	5.39
일반업체	1.32	정부기관	3.59	기타	1.00

4. 사이버위협 예측 모형 적용 결과

4.1 데이터 구성 및 모델링

입력 데이터는 크게 2주 단위로 종합된 평균 위협도와 누적 위협도를 사용한다. 인공신경망을 이용한 위협도 예측을 위해 입력데이터의 개수(즉, 미래 위협도를 예측하기 위해 사용되는 과거값의 개수), 은닉 뉴런의 개수를 변화시켰으며, 전체 데이터 중 1월에서 10월까지의 데이터를 훈련데이터로 사용하였으며, 11월과 12월 데이터는 모델의 성능을 측정하기 위한 시험데이터로 사용하였다. 모델의 검증을 위한 지표로는 MSE(Mean Square Error)를 사용하였으며, 단순히 과거 값의 평균을 예측값으로 추정하는 경우와 비교하여 상대적인 성능을 평가하였다.

위협도 예측 정확도를 평가하는 지표로 R_{MSE}를 다음과 같이 정의한다.

$$MSE_{ARIMA} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \overline{Y})^2$$

$$MSE_{ANN} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \widehat{Y}_i)^2$$

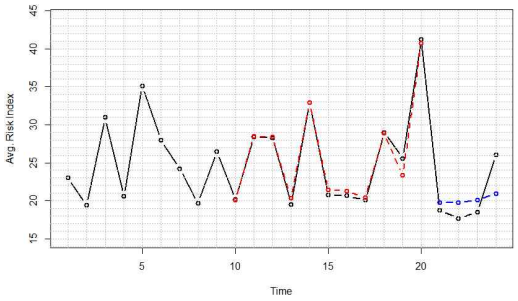
$$R_{MSE} = \frac{MSE_{ANN}}{MSE_{ARIMA}}$$

위 식에서 n은 예측한 데이터의 수를 나타낸다. Y_i는 i번째 시험데이터의 값을 의미하고 $\overline{Y}$은 시험데이터들의 평균을 의미한다. 그리고 $\widehat{Y}_i$은 인공신경망에 의한 i번째 예측값이다. 즉, R_{MSE}는 인공신경망을 통한 예측값과 시험데이터와의 MSE를 과거 데이터의 평균값과 시험데이터와의 MSE로 나눈 값으로 정의한다. 따라서 계산한 결과가 1보다 클수록 인공신경망 모델에 의한 예측값의 정확도가 낮은 것이고 1보다 작을수록 정확도가 높은 것이라 할 수 있다.

4.2 평균 위협도 예측

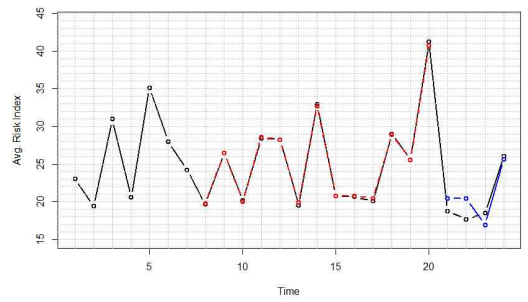
위협도 예측모델을 구성하기 위해 입력뉴런의 개수와 은닉뉴런의 개수를 각각 1~10개까지 변화시켜 그 성능의 차이를 비교하였다. 입력뉴런이 9개이고, 은닉뉴런이 1개일 때 $R_{MSE}=0.2542051$로 가장 우수한 성능을 나타내는 것을 확인하였다. 일반적으로 입력뉴런의 수와 은닉뉴런의 수에 따른 성능 상관관계를 파악하는 것은 쉽지 않다. 이는 훈련데이터의 패턴에 따라 성능이 좌우되기 때문이다. (그림 4-1)은 최적의 성능을 보일 때의 2주

단위 평균 위협도 예측 결과를 나타낸다.



(그림 4-1) 2주 단위 평균 위협도 예측
(#IN= 9, #HN= 1)

(그림 4-1)에서 검은색은 실제 데이터의 값을 나타내고, 붉은 색은 훈련 데이터로 학습한 결과이다. 파란색은 인공신경망 모델을 통해 예측한 값이다. 그림에서 알 수 있는 바와 같이 예측값의 결과가 실제 데이터의 값과 매우 유사함을 알 수 있다. 위협도가 급격히 감소한다는 것을 비교적 잘 예측하고 있는 것을 확인할 수 있다.



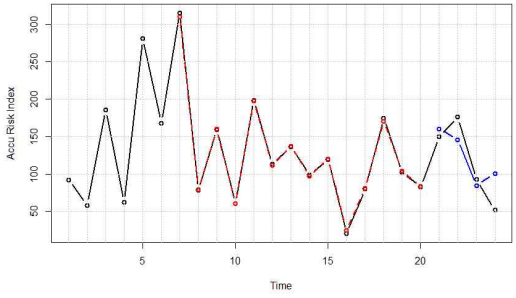
(그림 4-2) 2주 단위 평균 위협도 예측
(#IN=7, #HN=2)

(그림 4-2)는 입력뉴런의 수가 7일 때 은닉뉴런의 수가 2인 경우의 위협도 예측 결과를 나타낸다. 다양한 조건에서 인공신경망을 구성하여 실험한 결과 일반적으로 은닉뉴런의 수가 증가할수록 훈련데이터의 추세를 매우 정확하게 반영하고 예측 결과도 양호하다. 하지만 은닉뉴런의 수가 일정 수준 이상으로 증가하면 과적합(overfitting)으로 예측 결과의 정확도는 오히려 크게 떨어지게 된다.

4.3 누적 위협도 예측

앞에서 평균위협도 예측과 유사하게 2주 단위의 누적 위협도의 예측모델을 구성하기 위해 입력뉴런의 개수와 은닉뉴런의 개수를 각각 1~10개까지 변화시켜 그 성능의 차이를 비교하였다. 입력뉴런이 6개이고, 은닉뉴런이 4개일 때 $R_{MSE}=0.5085299$로 가장 우수한 성능을 나타내는 것을 확인할 수 있다.

(그림 4-3)은 최적의 성능을 보일 때의 위협도 예측 결과를 나타낸다. 결과에서 보듯이 마지막 예측값을 제외하고 비교적 정확하게 예측하는 것을 확인할 수 있다.



(그림 4-3) 2주 단위 누적 위협도 예측
(#IN=6, #HN=4)

4.4 사이버위협 예측 모형 분석

앞서 수행한 다양한 실험을 통해 확인한 바와 같이 본 연구에서 제시한 모형의 성능은 훈련데이터의 충분성에 의해 크게 좌우된다. 추가적으로 제시한 모형을 2주 단위와 월단위 두 가지 경우로 적용한 경우 2주 단위 위협 예측이 월단위 위협 예측에 비해 정확도가 높은 것을 확인할 수 있었다. 이는 2주 단위 훈련데이터가 월 단위에 비해 2배로 많아 그 만큼 충분한 학습이 가능했기 때문이다. 따라서 보다 정확한 위협 예측을 위해서는 많은 학습데이터 수집이 필수적이다.

은닉뉴런의 수가 증가할수록 훈련데이터의 추세를 매우 정확히 반영하며 예측값의 정확도도 일정

수준 증가한다. 하지만 일정 수 이상으로 은닉뉴런을 증가시킬 경우 예측값의 변화에 큰 영향이 없거나 오히려 예측값의 정확도를 크게 떨어뜨린다. 따라서 적절한 은닉뉴런의 수를 선정해야 한다. 또한, 시계열 데이터를 분석하는데 있어서 은닉계층의 수는 하나면 충분하다고 알려져 있다. 실험결과 은닉계층의 수가 하나임에도 불구하고 훈련데이터의 추세를 매우 정확하게 반영하므로 불필요하게 은닉계층의 수를 증가시킬 필요는 없다고 판단된다.

입력뉴런의 수가 정확도에 미치는 영향은 데이터가 내포하고 있는 패턴에 따라 다르게 나타난다. 미래의 값을 추정하는데 과거의 값 중 최근의 값의 영향이 크게 반영될 수도 있고, 반대로 그렇지 않을 수도 있다. 이것은 수집된 데이터에서 식별되는 패턴에 의해서 결정된다.

5. 결론 및 향후 연구방향

본 논문은 공개출처정보로부터 선정된 사이버위협 평가요소를 토대로 사이버 위협 정량화 방안을 제시하였고, 인공신경망을 이용하여 평균 위협도 및 누적 위협도를 예측하였다. 위협도 계산을 위해서 사이버 위협을 평가할 수 있는 5가지 요소를 이용한 정량화 공식을 제안하였다. 또한 제안된 공식을 토대로 인공신경망 모델을 이용하여 사이버 위협도를 계산하였다. 그리고 인공신경망의 입력뉴런의 수와 은닉뉴런의 수의 변화에 따른 예측결과를 제시하였고, 앞에서 살펴 본 바와 같이 평균 위협도와 누적 위협도예측이 비교적 잘 이루어짐을 확인하였다.

하지만 보다 성능이 향상된 예측 모형 개발을 위해서는 공개출처정보 데이터베이스를 보다 많이 축적하여 충분한 훈련데이터를 확보해야 한다. 또한 공개출처정보의 가치에 따라 제안하는 모델의 성능이 좌우되므로 공개출처정보의 신뢰성을 판단하는 절차가 별도로 추가될 필요가 있다.

한편, 본 연구에서 제안한 사이버위협 예측 모형은 공개출처정보로부터 도출된 사이버 위협도에 대한 예측이기 때문에 특정 시설 또는 특정 공격단계에 대한 예측이 제한적이다. 이를 함께 적용할 수 있도록 하기 위해서는 IDS 탐지결과 등과 같은 공개출처정보 이외의 데이터를 추가적으로 수집하여 공개출처정보 데이터베이스와 연동하고, 사이버위협도 정량화 방법을 지속적으로 보완한다면 보다 현실적이고 정확한 예측 모형 개발이 가능할 것이다. 따라서 추후에 기업, 정부, 연구소, 대학 등이 연계하여 공개형 사이버 공격 데이터베이스를 구축하고, 사이버위협과 인포콘(INFOCON) 연동모델 개발 등이 공동으로 이루어진다면 이를 기초로 다양한 사이버 위협을 실시간으로 예측하고 이를 예방할 수 있으리라 생각한다.

참고문헌

[1] Kuyoung Shin, Jinchel Yoo, Changhee Han, et al., "A study on building a cyber attack database using Open Source Intelligence(OSINT)", Convergence Security Journal 19(2), pp. 113-133, 2019.

[2] Sungrok Kang, Minam Moon, Kyuyoung Shin, Joogkwan Lee, "A study on Priority Analysis of Evaluation Factors for Cyber Threats using Open Source Intelligence(OSINT)", Convergence Security Journal 20(1), pp. 49-57, 2020.

[3] C. W. Geib and R. P. Goldman, "Plan recognition in intrusion detection systems," in DARPA Information Survivability Conference amp; Exposition II, 2001. DISCEX '01. Proceedings, 2001.

[4] A. A. Ahmed and N. A. K. Zaman, "Attack intention recognition: A review," IJ Network Security, vol. 19, no. 2, pp. 244 - 250, 2017.

[5] A. A. Ahmed and N. A. K. Zaman, "Attack intention recognition: A review," IJ Network Security, 2017.

[6] M. Abdlhamed, K. Kifayat, Q. Shi, and W. Hurst, "Intrusion Prediction Systems". Cham: Springer International Publishing, 2017.

[7] N. Polatidis, E. Pimenidis, M. Pavlidis, and H. Mouratidis, “Recommender systems meeting security: From product recommendation to cyber-attack prediction,” in Engineering Applications of Neural Networks. Cham: Springer International Publishing, 2017.

[8] K. Huang, C. Zhou, Y. C. Tian, S. Yang, and Y. Qin, “Assessing the physical impact of cyberattacks on industrial cyber-physical systems,” IEEE Transactions on Industrial Electronics, vol. 65, no. 10, 2018.

[9] A. Bar, B. Shapira, L. Rokach, and M. Unger, “Identifying Attack Propagation Patterns in Honeypots Using Markov Chains Modeling and Complex Networks Analysis,” in Software Science, Technology and Engineering (SWSTE), 2016 IEEE International Conference on.IEEE, 2016.

[10] M. Abdlhamed, K. Kifayat, Q. Shi, and W. Hurst, “A system for intrusion prediction in cloud computing,” in Proceedings of the International Conference on Internet of Things and Cloud Computing, ser. ICC ’16.New York, NY, USA: ACM, 2016.

[11] G. Werner, S. Yang, and K. McConky, “Time series forecasting of cyber attack intensity,” in Proceedings of the 12th Annual Conference on Cyber and Information Security Research, ser. CISRC ’17. New York, NY, USA: ACM, 2017.

[12] Y.-B. Leau and S. Manickam, “A Novel Adaptive Grey Verhulst Model for Network Security Situation Prediction,” International Journal of Advanced Computer Science & Applications, vol. 1, no. 7, 2016.

[13] F. He, Y. Zhang, D. Liu, Y. Dong, C. Liu, and C. Wu, “Mixed Wavelet-Based Neural Network Model for Cyber Security Situation Prediction Using MODWT and Hurst Exponent Analysis,” in Network and System Security. Cham: Springer International Publishing, 2017.

[14] G. K. Jayasinghe, J. S. Culpepper, and P. Bertok, “Efficient and effective realtime prediction of drive-by download attacks,” Journal of Network and Computer Applications, vol. 38, pp. 135 - 149, 2014.

[15] Y.-H. Kim and W. H. Park, “A study on cyber threat prediction based on intrusion detection event for apt attack detection,” Multimedia Tools and Applications, vol. 71, no. 2, pp. 685-698, Jul 2014.

[16] Goodfellow, Ian, et al. “Deep learning,” Vol. 1. Cambridge: MIT press, 2016.

[17] Abiodun, Oludare Isaac, et al. "State-of-the-art in artificial neural network applications: A survey." Heliyon, Vol. 4. No. 11, 2018.

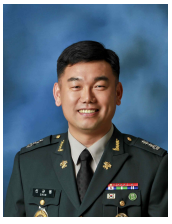
[18] Wang, Lin, et al. "Optimal forecast combination based on neural networks for time series forecasting." Applied soft computing 66, pp. 1-17, 2018.

[19] Singh, Navneet, Asheesh Singh, and Manoj Tripathy. "Selection of hidden layer neurons and best training method for ffnn in application of long term load forecasting." Journal of electrical engineering, Vol. 63, No.3, pp. 153-16, 2012.

〔저자소개〕

이 종 관 (Jongkwan Lee)
2000년 3월 육군사관학교 학사
2004년 3월 한국과학기술원 석사
2011년 3월 아주대학교 박사
email : jklee64@kma.ac.kr

문 미 남 (Minam Moon)
2001년 3월 육군사관학교 학사
2006년 2월 고려대학교 수학석사
2015년 8월 텍사스 A&M 대학교
수학박사
email : minammoon23@gmail.com



신 규 용 (Kyuyong Shin)
1996년 3월 육군사관학교 학사
2000년 2월 한국과학기술원 석사
2009년 12월 (美)노스캐롤라이나
주립대학교(NCSU) 박사
email : kyshin@kma.ac.kr

강 성 록 (Sungrok Kang)
1996년 3월 육군사관학교 학사
2001년 2월 연세대학교 석사
2010년 8월 (美)오리건주립대(OSU)
박사
email : ksr6452@mnd.go.kr