시스템 호출 기반의 사운덱스 알고리즘을 이용한 신경망과 N-gram 기법에 대한 이상 탐지 성능 분석☆

Anomaly Detection Performance Analysis of Neural Networks using Soundex Algorithm and N-gram Techniques based on System Calls

박 봉 구*
Park, Bong Goo

요 약

컴퓨터 네트워크의 확대 및 인터넷 이용의 급격한 증가에 따라 네트워크 서비스 품질의 보장과 네트워크의 관리가 어려울 뿐만 아니라 네트워크 보안의 취약성으로 인하여 해킹 및 정보유출 등의 위협에 노출되어 있다. 특히 시스템 침입의 보안 위협에 대한 능동적인 대처 및 침입 이후에 동일하거나 유사한 유형의 사건 발생에 대해 실시간에 대응하는 것이 중요하므로 침임 탐지 시스템에 대한 많은 연구가 진행되고 있다.

본 논문에서는 시스템 호출을 이용하여 이상 침입 탐지 시스템의 성능을 향상시키기 위해, 특징 선택과 가변 길이 데이터를 고정 길이 학습 패턴으로 변환 생성하는 문제를 해결하기 위한 사운덱스 알고리즘을 적용한 신경망 학습을 통하여 이상 침입 탐지의 연구를 하고자 한다. 즉, 가변 길이의 순차적인 시스템 호출 데이터를 사운덱스 알고리즘에 의한 고정 길이의 행위 패턴을 생성하여 역전파 알고리즘에 의해 신경망 학습을 수행하였다. 역전파 신경망 기법을 UNM의 Sendmail Data Set을 이용하여 시스템 호출의 이상 탐지에 적용하여 성능을 검증하였다.

Abstract

The weak foundation of the computing environment caused information leakage and hacking to be uncontrollable. Therefore, dynamic control of security threats and real-time reaction to identical or similar types of accidents after intrusion are considered to be important. As one of the solutions to solve the problem, studies on intrusion detection systems are actively being conducted.

To improve the anomaly IDS using system calls, this study focuses on neural networks learning using the soundex algorithm which is designed to change feature selection and variable length data into a fixed length learning pattern. That is, by changing variable length sequential system call data into a fixed length behavior pattern using the soundex algorithm, this study conducted neural networks learning by using a backpropagation algorithm. The backpropagation neural networks technique is applied for anomaly detection of system calls using Sendmail Data of UNM to demonstrate its performance.

☞ Keyword : Anomaly Intrusion Detection, Soundex Algorithm, System Calls, Neural Networks

1. 서 론

최근의 정보통신 기반구조는 컴퓨터 네트워크를 통한 연결로 다양한 서비스를 제공하고 있다. 특히 인터넷은 개방형 구조를 가지고 있어 서비스 품질 보장과 네트워크의 관리가 어렵고, 기반구조의 취약성으로 인하여 타인으로부터의 해킹 및 정보유출 등의 위협에 노출되어 있다. 불법 및 고의로 네트워크를 통한 컴퓨터 시스템에 접근하여 피해를 야기하는 문제에 대해 침입 차단, 인증 그리고 접근제어 등의 다양한 방법이 제공되고 있지만 역부족 상태이다. 보안 위협에 대한 능동적인 대처 및 침입 이후에 동일하거나 유사한 유형의 사

* 정 회 원 : 호남대학교 정보통신공학과 교수
bgpark@honam.ac.kr(제1저자)
[2005/04/12 투고 - 2005/04/19 심사 - 2005/07/22 심사완료]
☆ 본 논문은 호남대학교 교내 학술연구조성비의 지원에 의하여 수행되었음.

건 발생에 대해 실시간 대응할 수 있는 방법이 중요하게 되었다. 이러한 해결책으로서 침임 탐지 시스템에 대한 연구가 활발히 진행되고 있다.

침입 탐지 기법은 크게 이상 침입 탐지 기법과 오용 침입 탐지 기법으로 나눌 수 있다. 일반적으로 오용 탐지 방법이 많이 상업화되어 사용되지만 새로운 침입 패턴과 변형된 침입 패턴을 탐지할 수 없는 문제점이 있으며, 오용 탐지를 위한 공격 유형을 분석하여 오용 탐지 규칙 등의 인코딩 작업에 시간과 비용이 많이 소요되는 문제점을 갖고 있다. 해결책으로 정상 및 비정상 행위로부터 침입을 탐지하는 이상 침입 탐지 연구가 진행되고 있으나 아직은 연구 단계에 있으며 상업화되지는 못하고 있다.

초기 침입 탐지 시스템들은 이미 알려진 공격에 대한 징후를 수동으로 전문가가 인코딩하여 침입 여부를 판단하였다. 그러나 수동적인 방법에 의한 규칙 생성 및 확장은 매우 어려운 일이며, 그 효율성이 매우 떨어지는 방법이다. 이러한 문제를 해결하기 위하여 인공지능, 기계학습 및 데이터 마이닝 기법들을 침입 탐지에 이용하기 시작하였다. 지도학습에 기반을 둔 많은 침입 탐지 시스템은 학습과 침입 탐지 과정이 구분되어 있다. 따라서 침입 탐지를 위해서는 학습 과정이 반드시 필요하므로 시스템의 안정적 성능을 얻기까지 많은 비용이 들며, 학습을 위해 많은 양의 데이터가 필요하다. 이러한 방대한 학습 데이터의 수집 및 분류는 매우 어려운 일이며, 학습 데이터의 질에 의해 침입 탐지 시스템의 성능이 크게 좌우된다. 현재 침입 탐지에 사용되고 있는 많은 알고리즘은 방대한 데이터 처리 및 점증적 학습을 동시에 수행하기가 매우 어렵다. 따라서 실시간 침입 탐지를 위한 온라인 시스템의 구축이 어렵다. 또한 학습된 데이터 이외의 침입 유형에 대한 탐지 및 침입 유형에 대한 정보 제공도 어렵다[1,2,3,4].

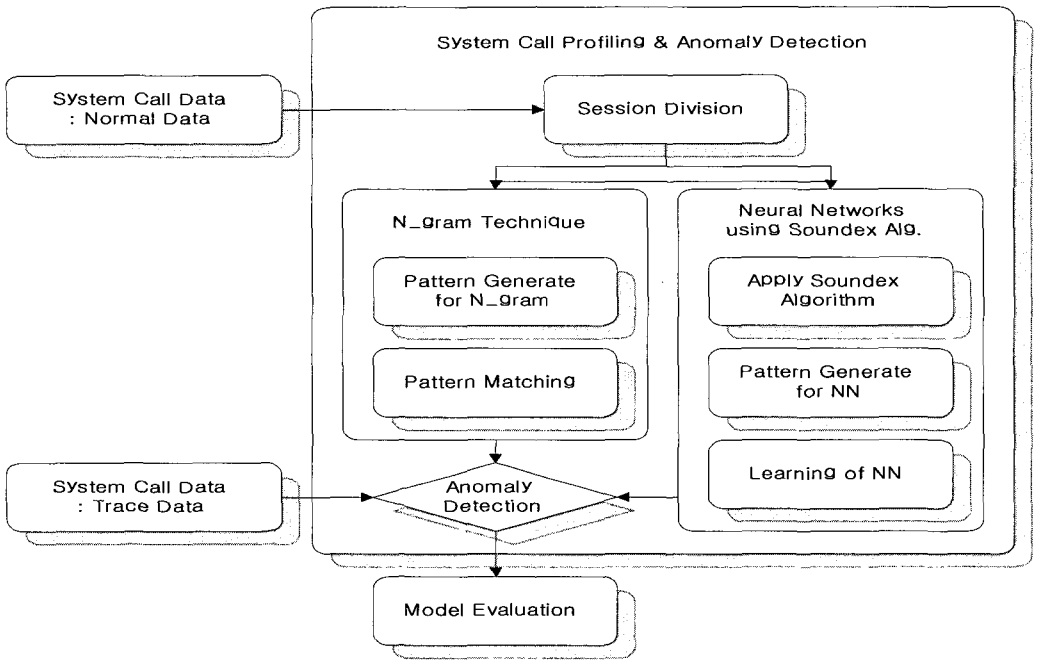
호스트 기반의 이상 침입 탐지 기법은 열거형 방법, 빈도 기반 방법, 데이터 마이닝 접근 방법 그리고 유한 상태 기계 방법으로 분류할 수 있다. 열거형 순차 방법은 정상 행위를 경험적으로 추적하여 알려지지 않은 패턴을 모니터링하여 이상 탐지한다. 빈도 기반의 방법은 다양한 이벤트의 빈도 분포를 기준으로 하여 침입을 탐지하며, 데이터 마이닝 접근 방법은 정상 행위 데이터로부터 발생하는 공통의 원소로부터 특징을 추출하고, 규칙 집합으로 기술함으로써 침입 탐지가 가능하도록 한다. 또한 유한 상태 기계 방법은 기계 학습 기법으로 프로그램을 추적하여 인식하는 유한 상태 기계를 구축하여 이상 침입을 탐지하는 방법이다[5].

본 논문에서는 지도학습 신경망 기반의 침입 탐지 시스템에서 학습에 사용되는 가변 길이의 시스템 호출 데이터의 문제점을 해결하기 위하여 사운덱스 알고리즘을 적용하고자 한다. 사운덱스 알고리즘에 의해서 가변 길이의 데이터를 고정 길이의 패턴 변환으로 간결한 학습 알고리즘과 학습을 위한 복잡도를 줄일 수 있다. 호스트 기반의 이상 침입을 탐지하기 위해서는 먼저 프로세스 아이디에 의한 세션을 구분하고, 시스템 호출을 이용하여 호스트의 행위 패턴을 사운덱스 알고리즘에 의해 가변 길이를 고정 길이 패턴으로 변환하여 생성한다. 정상적인 행위 패턴을 이용하여 정상 행위를 프로파일링하고, 신경망의 지도 학습에 의해서 비정상적 행위를 탐지하고자 한다.

2. 관련 연구

2.1 사운덱스 알고리즘

항공 회사와 같이 전화로 고객 업무를 처리하는 경우 발음이 부정확하거나 다른 고객의 이름을 검색하는 경우가 종종 발생한다. 이런 문제가 아니더라도 데이터베이스 안에 저장된 고객의 수가 많은 경우에는 고객의 이름을 하나씩 확인해 보는 선형 검색 방법은 지나치게 많은 시간을 필요로 한다. 이러한 문제점을 해결하기 위해서 마가렛 오델(Margaret K. Odell)과 로버트 러셀(Robert C. Russell)이 사운덱스 알고리즘을



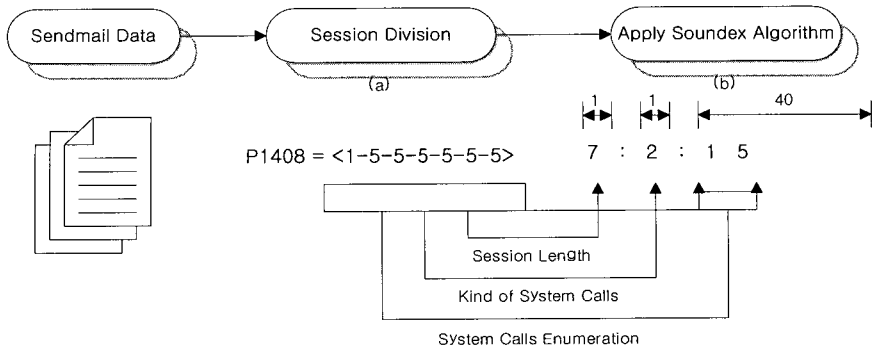
〈그림 1〉 사운덱스 알고리즘을 이용한 신경망과 N-gram 기법의 수행 절차

개발하였다. 사운덱스 알고리즘은 미군의 개인 기록부터 인구 통계 조사까지 사용되었다. 또한 여러 가지 소프트웨어의 철자 확인기(spell checker) 엔진 속에도 포함되어 활용되고 있으며, Ancestor Search 웹 사이트에서도 사운덱스 알고리즘을 사용하고 있다.

사운덱스 알고리즘은 네 가지의 규칙으로 이루어진다. 규칙 1은 이름의 첫 번째 글자를 저장하고, 첫 번째 글자를 제외한 나머지 글자 중에서 a, e, i, o, u, w, y를 모두 제거한다. 규칙 2는 이름 안에 존재하는 글자들에게 다음과 같은 번호를 부여한다. {b, f, p, v : 1}, {c, g, j, k, q, s, x, z : 2}, {d, t : 3}, {l : 4}, {m, n : 5}, {r : 6}. 규칙 3은 원래 이름에서 서로 인접하여 연속으로 나타나는 글자는 맨 앞에 하나만 남기고 나머지는 제거한다. 규칙 4는 최종적인 결과를 '글자, 숫자, 숫자, 숫자'의 형태로 맞추기 위해서 숫자가 세 개 이상이면 나머지는 생략하고, 세 개 미만이면 뒤에 0을 붙여서 형태를 맞춘다[6].

2.2 N-gram 기법

프로그램 행위 기반 침입 탐지 기법의 전제는 대부분의 공격은 프로그램 결함이나 버그로 인하여 발생할 수 있으며 프로그램의 정상적인 사용과는 그 행위가 다르다는데 있다. 그러므로 프로그램의 행위가 적합하게 표현될 수 있다면 침입 탐지를 위한 행위 특성으로 활용될 수 있다. 프로그램의 정상행위를 자동적으로 추출하고 정의하기 위한 대표적인 연구는 뉴멕시코 대학의 Forrest 연구팀에서 개발한 N-gram 기법이다. 이 기법은 면역학의 개념을 침입 탐지에 적용한 사례이다[7,8]. N-gram 기법은 프로그램에 의해 발생되는 일정 길이의 순차 시스템 호출들, 즉 N-gram 또는 스트링으로 프로파일 데이터베이스를 구축한다. 프로파일 데이터베이스가 구축된 후, 프로그램이 발생시킨 시스템 호출들 중에서, 특정 길이의 일련의 시스템 호출들이 프로파일에 존재하지 않는다면 비정상 행위로 간주하여 개수를 샌다.



〈그림 2〉 사운덱스 알고리즘을 이용한 행위 패턴 생성 과정

세션 내의 총 스트링 개수에 대해 비정상 행위로 간주된 스트링의 개수의 비율이 매우 크다면, 그 세션을 비정상으로 판정한다. Forrest 연구팀에서는 이 기법을 UNIX 프로그램 중에서 루트 권한으로 실행되는 주요 데몬 프로그램들, 즉 Sendmail, ftpd, inetd 등에 적용하여 높은 탐지율을 보였다. 그러나 이 기법은 프로그램마다 매우 큰 프로파일이 필요하다는 문제점이 있다[9,10].

3. 사운덱스 알고리즘을 이용한 신경망

시스템 호출 기반의 이상 침입 탐지에 사운덱스 알고리즘을 적용한 신경망과 N-gram 기법을 적용하여 성능을 비교하기 위한 절차는 그림 1과 같이 구성한다.

시스템 호출을 이용한 이상 침입을 탐지하기 위해서는 먼저 프로세스 아이디에 의한 세션을 구분한다. 세션은 행위를 나타내는 단위이며, 하나의 세션이 하나의 행위 패턴으로 변환된다. 가변 길이의 정상적인 시스템 호출 데이터를 이용하여 고정 길이의 정상 행위 패턴을 생성하여 정상 행위 프로파일을 구축한다. 정상 행위 패턴을 이용하여 신경망의 지도 학습을 수행하여 이상 침입 탐지를 수행한다.

정상 행위 프로파일을 구축하기 위해서는 N-gram 기법을 적용한 프로파일과 사운덱스 알고리즘과 신경망을 적용한 프로파일을 구축하여 두 모델간의 성능을 비교 분석한다.

3.1 사운덱스 알고리즘을 이용한 행위 패턴

호스트 기반의 침입 탐지에는 호스트의 시스템 호출 정보를 이용하여 침입을 탐지한다. 본 논문에서는 시스템 호출 정보를 이용하여 사운덱스 알고리즘에 의해 정상 행위를 고정 길이 패턴으로 프로파일링하여 신경망 학습에 의한 이상 침입을 탐지한다.

사운덱스 알고리즘을 이용하여 행위 패턴 생성 과정은 먼저, Sendmail DataSet을 프로세스 아이디(PID)에 의한 호출된 시스템 호출 번호를 필터링하여 세션을 구성한다. 그리고 구성된 세션을 사운덱스 알고리즘을 이용하여 고정 길이의 행위 패턴을 생성하며 그림 2와 같이 나타낸다. 그림 2의 (a)와 (b) 과정의 산출물은 각각 그림 3과 4에 해당한다.

정상 행위의 프로파일을 구축하기 위해서는 하나의 행위를 기술할 수 있는 표현법이 필요하다. 본 논문에서 사용하는 행위 표현법은 세션의 시작과 끝은 각각 <와 >으로 표시하고, 시스템 호출간의 구분은 '-'에 의해 구분한다. 호스트의 행위를 표현하기 위하여 UNM의 Sendmail Data Sets[13]의 세션 구분의 결과를 그림 3과

〈표 1〉 특징 선택에 의한 패턴 벡터의 구성

특징 선택	형태	내용
세션의 길이	Integer	세션의 시스템 호출 갯수
시스템 호출 종류	Integer	세션의 시스템 호출의 종류
시스템 호출의 나열	Sequential Integer	시스템 호출의 종류별 발생 순서 나열

〈표 2〉 학습에 사용될 패턴 수

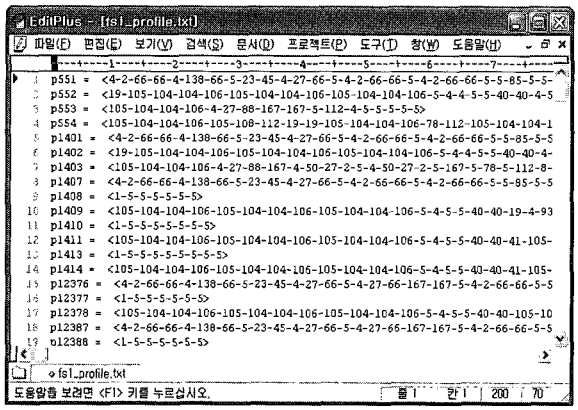
패턴 클래스		세션의 수	패턴 벡터의 수	시스템 호출의 종류	패턴 벡터 수/세션의 평균
훈련	정상	199	228,181	53	1,147
	침입	15	4,186	48	279
추적		10	2,569	43	257
합계		224	234,936		1,049

같이 나타낼 수 있다. 그림 3에 표현된 호스트의 행위 패턴들을 모아서 정상 행위 프로파일 구축에 사용된다.

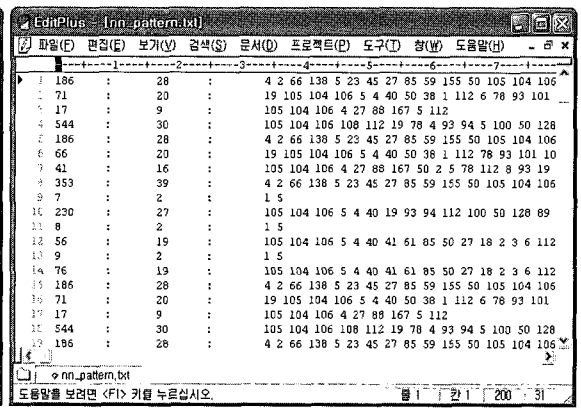
시스템 호출 정보를 이용하여 프로세스 아이디에 의해서 세션 별로 분류하면, 세션의 크기가 고정적이지 않고 가변적이다. 세션에 사용된 시스템 호출 종류가 최소 2, 최대 40 종류이며, 세션의 크기가 최소 7에서 최대 31927로 매우 가변적이다. 가변 길이의 데이터는 데이터 처리도 어렵지만, 신경망 학습에 학습 패턴으로 적용하기도 어렵다.

본 논문에서는 시스템 호출 데이터로 이루어진 가변적인 세션에 사운덱스 알고리즘을 적용함으로써 세션 정보를 유지하면서 고정된 행위 패턴의 프로파일을 구축한다. 세션을 이루는 가변의 행위 데이터를 신경망 학습에 적용하기 위해서는 먼저, 패턴을 생성하기 위한 특징 선택이 필요하다. 본 논문에서는 3개의 특징을 선택하였다. 패턴 벡터의 선택으로는 세션의 크기, 시스템 호출의 종류, 그리고 나열로 표 1과 같이 특징 선택하였고 그림 4와 같이 생성하였다.

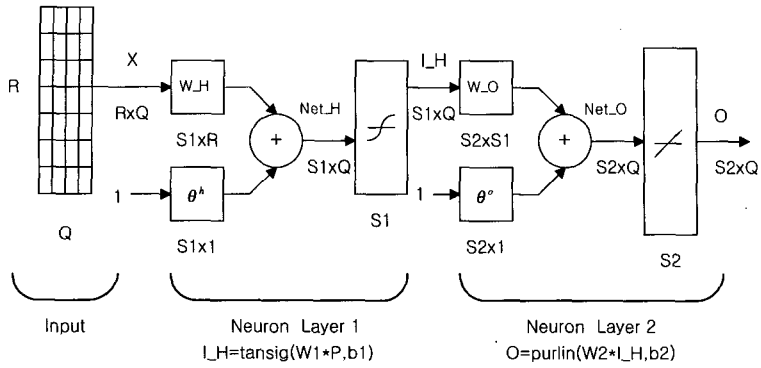
특징 선택 중에서 나열 필드는 40개의 항목으로 구성된다. 모든 행위 패턴을 대상으로 조사한 결과 시스템 호출은 182 종류이나, 행위 데이터에 사용된 시스템 호출 종류는 53 종류이었다. 세션 중에서 가장 많이 쓰인 시스템 호출의 종



〈그림 3〉 시스템 호출을 이용한 행위 세션 생성



〈그림 4〉 사운덱스 알고리즘을 이용한 패턴 벡터 생성



〈그림 5〉 2 계층의 역전파 모델

류는 40 종류이었다. 모든 행위 패턴들은 최대가 40이거나 이보다는 작을 것이다. 그러므로 나열 항목의 크기를 40으로 설정하였다.

표 2는 시뮬레이션에 사용된 학습 패턴에 대한 패턴의 특징 정보를 나타낸 것이다. 신경망 학습에 사용될 훈련 패턴은 정상 패턴이며, 학습 후의 검증에 사용될 추적 패턴으로 구분된다. 훈련 패턴은 다시 정상과 비정상 패턴으로 구분된다. 정상 패턴에 사용된 시스템 호출은 53 종류, 비정상 패턴은 48 종류, 그리고 테스트 패턴은 43 종류가 사용되었다. 정상 패턴의 세션은 199개, 비정상 패턴의 세션은 15개 그리고 테스트 패턴의 세션은 10개로 구성된다. 훈련 패턴은 199개의 정상 행위 패턴을 사용하여 신경망 학습을 수행한다.

3.2 역전파 신경망 학습

신경망은 두뇌 활동의 메커니즘을 수학적으로 재현한 인공지능의 한 분야이다. 신경망은 인간의 두뇌를 모방하여 지적능력을 학습을 통하여 컴퓨터의 지식 베이스로 구축하고, 구축된 지식 베이스를 이용하여 주어진 자료를 추론하고 그 결과를 예측하고 설명하는 기능을 말한다.

신경망이 주어진 자료의 특성을 학습하는데 사용되는 학습 알고리즘에는 여러 가지가 있으나 그 중에서 오차를 최소화시켜 나가는 역전파(Backpropagation) 방법이 흔히 사용된다. 역전파 알고리즘은 최소자승 알고리즘의 비선형적 확장으로 볼 수 있는 가장 많이 쓰이는 지도학습 기법이다. 즉, 입력계층의 각 노드에 입력패턴을 주면 이 신호는 각 노드에서 변환되어 은닉계층에 전달되고 계산과정을 거쳐 출력계층에서 신호를 출력하게 된다. 이때 출력값과 목표값을 비교하여 둘 사이의 차이, 즉 오차를 줄여나가는 방향으로 가중치를 반복적으로 조정해 나가는 방법이 역전파 신경망 학습법이다.

2 계층 역전파 신경망의 소프트웨어에 의한 학습 모델을 그림 5에 나타낸다.

역전파 신경망의 학습은 입력 x와 은닉계층의 가중치 w의 곱의 합에 은닉계층의 편의(bias) θ를 더하여 순입력 net으로 식(1)과 같이 사용되며, 그림 5에서 계층 1의 Net_H 항목에 해당된다. 윗 첨자 h와 o는 은닉계층과 출력계층을 나타내며, 아래 첨자 p와 j는 원소의 위치를 나타낸다.

$$\neq i_{pj}^{h} = \sum_{i=1}^{N} w_{ji}^{h} x_{\pi} + \theta_{j}^{h} \quad (1)$$

순입력 net에 의한 은닉계층의 전달함수 출력 i가 식(2)와 같이 계산되며, 그림 5에서 계층 1의 I_H 항목에 해당된다.

$$i_p^h = f_j^h(\neq t_{pj}^h) \quad (2)$$

은닉계층의 출력을 출력계층의 입력으로 하고, 출력계층의 가중치의 곱의 합에 출력계층의 편의가 더하여 출력계층의 순입력으로 식(3)과 같이 사용되며, 그림 5에서 계층 2의 Net_O 항목에 해당된다.

$$\neq t_{pk}^o = \sum_{j=1}^{L} w_{kj}^o i_{pj} + \theta_k^o \quad (3)$$

순입력에 의한 출력 계층의 전달함수 출력 o가 식(4)와 같이 계산되며, 그림 5에서 계층 2의 O 항목에 해당된다.

$$o_{pk} = f_k^o(\neq t_{pk}^o) \quad (4)$$

식 (1)과 식 (2)는 그림 5의 계층 1의 은닉계층의 기능을 나타내며, 식 (3)과 식 (4)는 그림 5의 계층 2의 출력계층의 기능을 나타낸다.

다음 단계는 교사 신호에 의해서 학습이 이루어지는 단계이다. 교사 신호는 출력계층의 오차를 이용하며, 역전파라는 이름에서 의미하듯 오차를 뒤로 전파시킨다. 즉, 출력계층의 오차를 이용해서 은닉계층의 오차를 계산한다.

출력계층과 은닉계층의 오차는 식 (5)와 (6)에 의해서 계산된다.

$$\delta_{pk}^o = (y_{pk} - o_{pk}) f_k^{o'}(\neq t_{pk}^o) \quad (5)$$

$$\delta_{pj}^h = f_j^{h'}(\neq t_{pj}^h) \sum_k \delta_{pk}^o w_{kj}^o \quad (6)$$

학습은 출력계층과 은닉계층의 가중치 수정에 의해서 이루어진다. 가중치의 수정은 학습율 η와 식 (5)와 식 (6)을 교사 신호로 이용하여 식 (7)과 (8)에 의해서 계산된다.

$$w_{kj}^o(t+1) = w_{kj}^o(t) + \eta \delta_{pk}^o i_{pj} \quad (7)$$

$$w_{ji}^h(t+1) = w_{ji}^h(t) + \eta \delta_{pj}^h x_i \quad (8)$$

출력계층의 출력값과 목표치의 오차가 신뢰수준에 도달하거나 제약조건을 만족할 때까지 학습은 계속 이루어진다.

본 논문에서는 사운덱스 알고리즘에 의해서 3개의 필드, 42 항목의 정상 행위 패턴을 이용하여 이상 침입 탐지를 위한 역전파 신경망의 학습을 수행한다. 학습이 완료되면 추적 데이터에 의해서 이상 침입 탐지 성능을 측정하고 N-gram 기법과 비교한다.

4. 시뮬레이션

사운덱스 알고리즘을 이용한 신경망의 이상 침입 탐지 시뮬레이션은 UNM의 Sendmail Data Set을 이용하였고, 시뮬레이션 툴은 Perl과 Matlab을 이용하였다. UNM의 Sendmail Data Set의 정상 행위 데이터는 bounce.int.gz, bounce-1.int.gz, bounce-2.int.gz, plus.int.gz, queue.int.gz, sendmail.log.int.gz을 사용하였고, 침입 행위 데이터는 sm-10763.int.gz, sm-10801.int.gz, sm-10814.int.gz, sm-280.int.gz, sm-314.int.gz 파일을 사용하였다. 그리고 성능 비교를 위한 추적 데이터는 fwd-loops-1.int.gz, fwd-loops-2.int.gz, fwd-

〈표 3〉 정상 데이터의 윈도우 크기에 의한 중복 제거한 패턴의 수

N	패턴 수	중복 제거 패턴 수	N	패턴 수	중복 제거 패턴 수
3	809997	440	7	226987	910
4	227584	570	8	226788	996
5	227385	693	9	226640	1076
6	227186	811	10	326179	1153

〈표 4〉 침입과 추적 데이터의 윈도우 크기에 중복 제거한 이상 패턴의 수

N	탐지 패턴의 수		N	탐지 패턴의 수	
	Intrusion	Trace		Intrusion	Trace
3	21	18	7	72	55
4	176	177	8	78	64
5	55	38	9	84	73
6	66	46	10	90	80

〈표 5〉 정상 행위의 윈도우 크기에 의한 세션 별 이상 탐지 패턴의 수

PID	Session Length	N							
		3	4	5	6	7	8	9	10
162	532	18	28	38	46	55	64	73	80
163	102	8	14	19	20	21	22	23	24
170	489	7	13	18	23	31	38	45	50
182	533	20	29	38	45	52	58	65	71
183	24	5	8	10	11	11	11	11	11
206	533	18	28	38	46	55	64	73	80
207	102	8	14	19	20	21	22	23	24
107	160	0	2	3	4	6	7	7	7
119	77	7	14	17	20	23	25	27	29
144	17	0	0	0	0	0	0	0	0
Total	2,569	91	150	200	235	275	311	347	376
탐지율		80%	90%						

loops-3.int.gz, fwd-loops-4.int.gz, fwd-loops-5.int.gz 파일을 Perl을 이용하여 시스템 호출 데이터의 세션을 구분하고, 정상 행위, 침입 행위 그리고 추적 패턴들을 생성하였다.

4.1 N-gram 기법

정상 행위 데이터에 N-gram 기법을 적용하여 정상 행위의 패턴을 생성한다. N-gram 기법의 윈도우 크기를 나타내는 N을 변화시키면서 추적 데이터에 의한 이상 탐지율을 비교 분석한다.

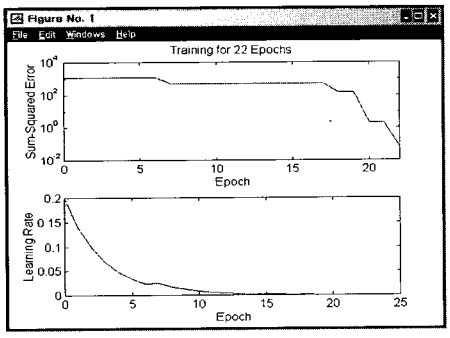
먼저, N-gram 기법에 의한 정상 행위를 프로파일하면 표 3과 같은 결과를 얻을 수 있다. 윈도우의 크기 N을 증가시키면 패턴의 수는 일부 영역을 제외하고는 단조 감소 경향을 보이며, 패턴의 중복을 제거하면 패턴의 수는 증가하는 경향을 보였다.

정상 행위 데이터를 기준으로 침입 데이터의 15개 세션과 추적 데이터의 10개의 세션을 비교하여 정상 행위에 존재하지 않는 중복 제거한 패턴의 수를 표 4에 나타낸다.

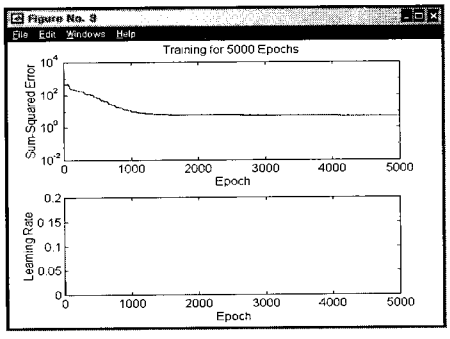
윈도우 크기 N이 커짐에 따라 탐지되는 패턴의 수도 증가하지만, 중복 제거된 패턴의 수도 단조 증가하였다. 특별하게 N이 4인 경우에 침입 데이터와 추적 데이터의 중복 제거한 패턴의 수가 갑자기 증가하였으며, 가장 큰 값을 보였다.

정상 데이터의 N의 크기를 변경하면서 10개의 세션인 추적 데이터를 이용해 이상 탐지를 수행하였다. 표 4에서 윈도우 크기가 4인 경우에 많은 패턴 종류의 수를 탐지하였지만, 표 5에서 세션의 탐지율은 윈도우 크기가 10인 경우에 대체적으로 많은 패턴 수를 탐지하였다.

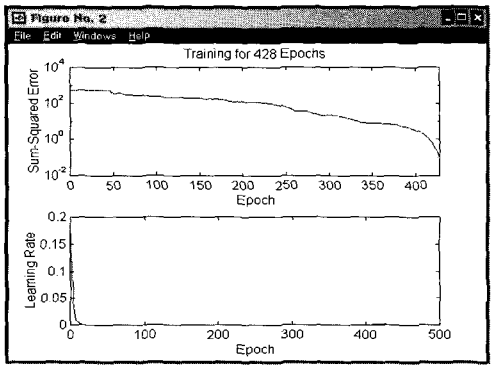
N-gram의 윈도우의 크기 N을 3부터 10까지 윈도우의 크기가 커짐에 따라 이상 탐지된 패턴의 수도 점차 단조 증가하였고, 윈도우 크기가 3에서 4로 변경되면, 탐지율은 80%에서 90%로 상승하였다. $N=3$인 경우에 추적 데이터의 10개 세션에 대해 PID 107과 PID 144의 두 세션이 미탐지 되었으나, $N=4$인 경우에는 PID 144



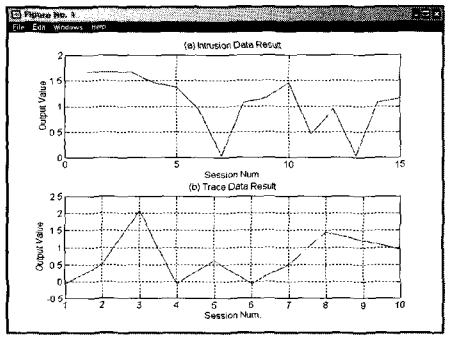
〈그림 6〉 은닉계층의 뉴런이 10인 경우



〈그림 7〉 은닉계층의 뉴런이 30인 경우



〈그림 8〉 은닉계층의 뉴런이 12인 경우



〈그림 9〉 정상 행위 데이터에 의한 침입과 추적 데이터의 역전파 신경망의 출력값

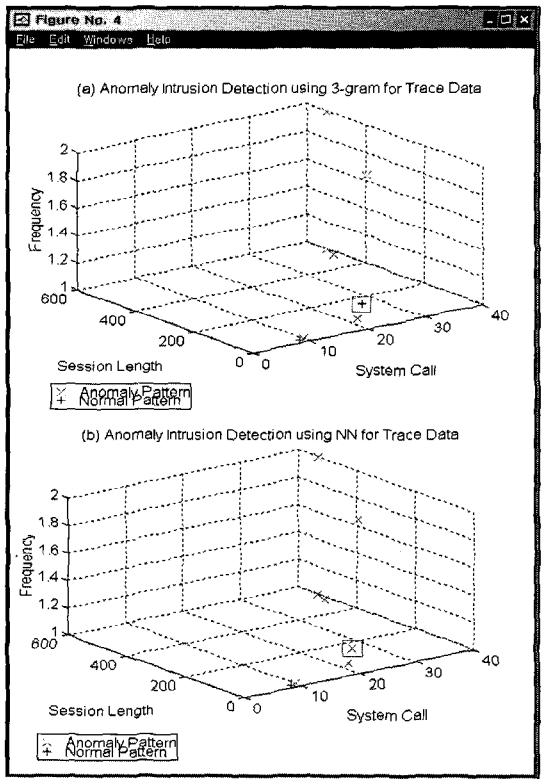
의 세션만 미탐지 되었다. 직관적으로 N-gram 기법에서는 윈도우 크기가 4인 경우에 가장 높은 탐지율 90%이고, N이 증가하여도 탐지율은 상승하지 않으므로, Occam's Razor 원리에 의해서 가장 효율적인 윈도우 크기이다.

4.2 신경망 학습의 과적합과 미적합

신경망 학습을 위해 은닉 계층의 뉴런 수를 10에서 40까지 변경하여 학습율과 에러를 그림 6에서 그림 8까지 나타낸다. 신경망 학습의 단점인 미적합(underfitting)과 과적합(overfitting)[14]을 벗어나기 위해서는 은닉 계층의 뉴런 수를 결정하여야 한다. 과적합은 학습에 학습 데이터이외에 잡음도 학습하는 경우를 의미하며, 미적합은 학습이 제대로 이루어지지 않은 상태를 의미한다.

그림 6은 은닉 계층의 뉴런 10개에 의한 신경망의 미적합 상태를 나타내고, 그림 7은 은닉 계층의 뉴런 30개에 의한 과적합 상태를 나타낸다. 그림 8은 은닉계층의 뉴런이 12개이며, 428 Epoch에 의해서 학습이 이루어진 상태를 나타낸다.

신경망 학습을 위해 은닉계층의 뉴런의 수를 12로 설정하고 199개의 정상 행위 패턴을 학습 데이터로 사용하였다. 정상 행위 패턴은 시스템 호출 데이터를 사운덱스 알고리즘에 의해 42개 항목의 학습 패턴으로 생성하였고, 신경망 학습의 오차율 0.01, 학습율 0.2와 Epoch수를 5000



〈그림 10〉 N-gram과 제안한 방법의 이상 탐지 결과

번 이하로 학습을 수행하였다. 그림 9는 학습된 신경망에 침입 데이터와 추적 데이터를 입력하여 이상 침입을 탐지한 결과이다.

그림 10은 N-gram과 사운덱스 알고리즘을 이용한 신경망 기법의 이상 탐지 결과를 나타낸 것

〈표 6〉 제안한 방법과 N-gram의 비교분석

항목		N-gram			사운덱스 알고리즘을 이용한 신경망	분석 (N-gram : 신경망)
		3	4	10		
중복	패턴 수	809,997	227,584	326,179	199	4070, 1144, 1639 : 1
	데이터량	9,608KB	3,473KB	12,076KB	22KB	437, 158, 549 : 1
중복 제거	패턴 수	440	570	1,153	41	11, 14, 28 : 1
	데이터량	5KB	7KB	34KB	5KB	1, 1.4, 6.8 : 1
탐지율 (탐지세션/전체세션)		8/10	9/10	9/10	9/10	0.8, 0.9, 0.9 : 0.9

이다. 그림 10의 (a)는 N-gram 기법의 윈도우 크기가 3인 경우의 이상 탐지 결과를 나타낸 것이며, 추적 데이터의 10개 세션 중에서 8개의 세션을 탐지하였다. 그림 10의 (b)는 그림 9의 (b)의 사운덱스 알고리즘을 이용한 신경망에 의한 탐지 결과를 추적 데이터의 특징 벡터 분포에서 나타낸 것이다. 추적 데이터의 10개 세션 중에서 9개의 세션을 탐지하여 탐지율 90%를 보였다.

4.3 사운덱스 알고리즘을 이용한 신경망과 N-gram 기법의 이상 탐지 비교

UNM의 Sendmail 데몬의 시스템 호출 데이터 집합에 대해 사운덱스 알고리즘을 이용한 신경망과 N-gram 기법으로 이상 탐지를 시뮬레이션한다. 시스템 호출 데이터를 사운덱스 알고리즘에 의한 42개 항목의 학습 패턴으로 변환하고, 신경망 학습의 오차율 0.01, 학습율 0.2와 Epoch수를 5000번 이하로 학습을 수행하였다. 그리고 N-gram 기법은 윈도우의 크기를 3에서 10까지 변경하여 이상 행위 탐지를 수행하여 표 6과 같이 결과를 비교 분석하였다. 사운덱스 알고리즘을 이용한 신경망 기법은 N-gram 기법과 비교하여 프로파일링을 위한 공간이 평균 2284분의 1로 축소되어서 공간 복잡도 측면에서 월등하게 우수하였다.

MDL(Minimum Description Length)[15]은 에러 손실 $L(D|H)$와 복잡도 손실 $L(H)$로 구성된다. MDL은 작은 값을 갖는 모델이 효율적인 모델이 된다. 에러 손실은 "1 - 이상 탐지율"로 정의하고, 복잡도 손실은 "1 - {자료 기술 공간에 대한 정보 점유율}"로 정의한다. 표 7은 제안한 방법과 N-gram 기법을 MDL에 의해서 비교 분석한다.

MDL에 의해서 N-gram 기법의 윈도우 크기 $N=4$인 경우에 가장 효율적으로 판명되었다. 그러나 제안한 방법과 N-gram의 $N=4$인 경우와 비교하면 탐지율의 성능은 같지만, 모델 복잡도 측면에서 제안한 방법이 더 효율적인 것으로 나타났다.

N-gram의 윈도우 크기가 3, 4 그리고 10인 경우의 탐지율은 80%, 90% 그리고 90%이었으며, N-gram의 윈도우가 커질수록 패턴을 기술하기 위한 많은 공간과 패턴 처리를 위한 시간이 증가하게 된다. 그러나 사운덱스 알고리즘과 신경망을 이용한 경우의 탐지율은 90%이었다. 신경망을 이용한 이상 탐지 경우와 비교하면, N-gram의 윈도우 크기가 3인 경우와 비교하면, 제안한 방법이 탐지율과 복잡도 측면에서 절대적으로 우수하였다. 윈도우 크기가 4부터 10까지의 이상 탐지 성

〈표 7〉 제안한 방법과 N-gram의 MDL 비교분석

항목	N-gram			사운덱스 알고리즘을 이용한 신경망
	3	4	10	
에러 손실	0.2	0.1	0.1	0.1
복잡도 손실	0.9970	0.9999	1.0000	0.9997
MDL	1.1970	1.0999	1.1000	1.0997

능을 같은 수준을 유지하면서 절대적으로 신경망 기법이 시간과 공간 복잡도 측면에서 우수하였다.

5. 결 론

본 논문에서는 기계학습 기법인 지도학습 신경망을 이용한 이상 침입 탐지 시스템에 사용될 가변길이 데이터 문제점을 해결하기 위하여 사운덱스 알고리즘을 적용하였다. 사운덱스 알고리즘에 의한 가변 길이의 시스템 호출 데이터를 고정 길이 패턴의 변환으로 신경망의 학습 알고리즘이 간결해지고 침입 탐지를 위한 학습에 공간과 시간 복잡도를 해결하였다. 호스트 기반의 이상 침입을 탐지하기 위해서는 먼저 세션을 구분하고, 호스트의 행위 패턴을 사운덱스 알고리즘에 의해 가변 길이 데이터를 고정 길이 패턴으로 변환하여 생성한다. 정상적인 행위 패턴을 지도학습 기법인 역전파 신경망을 이용하여 정상 행위를 학습하여 이상 행위를 탐지하였다. 학습에 사용될 가변 길이 데이터 처리의 어려움을 해결하여 학습 알고리즘이 간결하고, 학습을 위한 공간과 시간 복잡도를 줄이는 효과를 가져와 이상 침입 탐지의 성능을 향상시켰다.

시뮬레이션에 사용한 데이터는 UNM의 Sendmail Data Sets을 이용하여 호스트의 Sendmail 데몬의 정상 행위를 세션의 길이, 시스템 호출의 종류, 그리고 사용된 시스템 호출을 나열하는 3개 특징 선택을 하여 42 항목의 패턴 벡터를 생성하여 뉴로-퍼지를 이용한 이상 침입 탐지를 수행한다. 그리고 신경망 학습의 오차율 0.1, 학습율 0.2와 Epoch수를 5000번 이하로 학습을 수행하였다.

신경망과 윈도우 크기가 3인 N-gram 기법에서는 제안한 방법의 이상 탐지율이 더 높았으나 크기가 4부터 10까지의 경우에는 N-gram 기법과 동등한 탐지율 90%를 보였다. 그러나 MDL에 의한 알고리즘 수행의 시간과 공간 복잡도 측면에서는 사운덱스 알고리즘을 이용한 신경망의 이상 침입 탐지가 더 우수하였다.

참 고 문 헌

[1] Leonid Portnoy, "Intrusion detection with unlabeled data using clustering", Undergraduate Thesis, Columbia University, 2000.

[2] Jack Marin, Daniel Ragsdale, and John Shurdu, "A Hybrid Approach to the Profile Creation and Intrusion Detection", Proceedings of DARPA Information Survivability Conference and Exposition, IEEE, 2001.

[3] Nong Ye, and Xiangyang Li, "A Scalable Clustering Technique for Intrusion Signature Recognition", Proceedings of 2001 IEEE Workshop on Information Assurance and Security, 2001.

[4] Wenke Lee, Salvatore J. Stolfo, Philip K. Chan, Eleazar Eskin, Wei Fan, Matthew Miller, Shlomo Hershkop, and Junxin Zhang, "Real Time Data Mining - based Intrusion Detection", IEEE, 2001.

[5] Christina Warrender, Stephanie Forrest, Barak Pearlmutter, "Detecting Intrusion Using System Calls : Alternative Data Models", 1998.

[6] Soundex Algorithm, http://www.archives.gov/research_room/genealogy/census/soundex.html

[7] S. Forrest, S. Hofmeyr, A. Somayaji and T. Longstaff, "A sense of self for unix processes", In IEEE Symposium on Security and Privacy, pp.120-128, 1996.

[8] Steven A. Hofmeyr, Stephanie Forrest, Anil Somayaji, "Intrusion Detection using Sequences of System Calls", Journal of

Computer Security, Vol.6, pp.151-180, August 18, 1998.

[9] A. K. Ghosh, A. Schwarzbard and M. Shatz, "Learning program behavior profiles for intrusion detection", Proceedings of the 1st USENIX Workshop on Intrusion Detection and Network Monitoring, April, 1999.

[10] A. K. Ghish, J. Wanken and F. charron, "Detecting anomalous and unknown intrusions against programs", Proceedings of the 1998 Annual Computer Security Applications Conference(ACSAC '98), 1998.

[11] A. Wespi, M. Dacier and H. Debara, "Intrusion detection using variable-length audit trail patterns", Recent Advances in Intrusion Detection(RAID 2000), pp. 110-129, 2000.

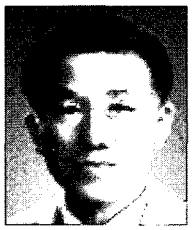
[12] W. Lee and S. Stolfo, "Learning Patterns from Unix Process Execution Traces for Intrusion Detection", AAAI Workshop : AL Approaches to Fraud Detection and RISK Management, pp.50-56, July, 1997.

[13] UNM의 Sendmail Data Sets, http://cs.unm.edu/~immsec/data/synth-sm.html

[14] Simon Haykin, "Neural Networks: A Comprehensive Foundation", IEEE Press, pp.179-181, 1994.

[15] Christopher M. Bishop, "Neural Networks for Pattern Recognition", Oxford Press, pp. 429-433, 1995.

◐ 저 자 소 개 ◑



박 봉 구 (Bong-Goo Park)

1973년 공주사범대학 수학교육학과 졸업(학사)
1982년 원광대학교 대학원 수학과 졸업(석사)
1987년 조선대학교 대학원 수학과 졸업((박사)
1984~ 현재 호남대학교 정보통신공학과 교수
관심분야 : 정보보안
E-mail : bgpark@honam.ac.kr