A Method to Find Feature Set for Detecting Various Denial Service Attacks in Power Grid

전력망에서의 다양한 서비스 거부 공격 탐지 위한 특징 선택 방법

DongHwi Lee, Young-Dae Kim, Woo-Bin Park, Joon-Seok Kim, Seung-Ho Kang†
이동휘, 김영대, 박우빈, 김준석, 강승호†

Department of Information Security, Dongshin University, 185 Geonjae-Ro, Naju, Jeonnam 58245, Korea
† kinston@gmail.com

Abstract

Network intrusion detection system based on machine learning method such as artificial neural network is quite dependent on the selected features in terms of accuracy and efficiency. Nevertheless, choosing the optimal combination of features, which guarantees accuracy and efficienty, from generally used many features to detect network intrusion requires extensive computing resources. In this paper, we deal with a optimal feature selection problem to determine 6 denial service attacks and normal usage provided by NSL-KDD data. We propose a optimal feature selection algorithm. Proposed algorithm is based on the multi-start local search algorithm, one of representative meta-heuristic algorithm for solving optimization problem. In order to evaluate the performance of our proposed algorithm, comparison with a case of all 41 features used against NSL-KDD data is conducted. In addtion, comparisons between 3 well-known machine learning methods (multi-layer perceptron., Bayes classifier, and Support vector machine) are performed to find a machine learning method which shows the best performance combined with the proposed feature selection method.

인공신경망과 같은 기계학습에 기반한 네트워크 침입탐지/방지시스템은 특징 조합에 따라 탐지의 정확성과 효율성 측면에서 크게 영향을 받는다. 하지만 침입탐지에 사용 가능한 여러개의 특징들 중 정확성과 효율성 측면에서 최적의 특징 조합을 추출하는 특징 선택 문제는 많은 계산량을 요구한다. 본 논문에서는 NSL-KDD 데이터 집합에서 제공하는 6가지 서비스 거부 공격과 정상 트래픽을 구분해 내기 위한 최적 특징 조합 선택 문제를 다룬다. 최적 특징 조합 선택 문제를 해결하기 위해 대표적인 메타 휴리스틱 알고리즘 중 하나인 다중 시작 지역탐색 알고리즘에 기반한 최적 특징 선택 알고리즘을 제시한다. 제안한 특징 선택 알고리즘의 성능 평가를 위해 NSL-KDD 데이터를 상대로 41개의 특징 모두를 사용한 경우와 비교한다. 그리고 선택된 특징 조합을 사용했을 때 가장 높은 성능을 보여주는 기계학습 방법을 찾기위해 3가지 잘 알려진 기계학습 방법들 (베이즈 분류기와 인공신경망, 서포트 벡터 머신)을 사용해 성능을 비교한다.

Keywords : Network intrusion detection system, machine learning, feature selection, local search algorithm, power grid

I. 서론

컴퓨터 응용 범위의 확대와 유무선 통신망의 폭발적인 보급에 따라 정보 유통의 방법 및 규모에 이전과는 비교할 수 없는 변화가 도래하였다. 특히 전력망과 정보통신기술이 융합하면서 전력생산 및 송배전의 영역 및 규모, 시간 등에서 효율성과 생산성을 높이는데 많은 기여하였다. 하지만 컴퓨터의 응용 범위 확대 및 통신망의 보급은 이와 같은 긍정적 측면과 함께 악의적인 사용자들에 의한 정보 유출, 침해 사고 등 부정적인 측면 또한 발생시켰고 스마트 그리드와 같은 새로운 전력망의 구축 및 보급에 많은 걸림돌이 되고 있다. 이러한 부정적 측면을 해결하기 위한 한 방법으로 네트워크에서 이루어지는 악의적 사용자에 의한 비정상적 사용을 탐지/방지하고자 하는 다양한 노력이 있어왔다. 하지만 기존의 방법들은 보안 전문가에 의해 선택된 시그니처와 몇 가지 규칙에 기반하고 있는 방법들이 대부분이다. 이러한 방법들은 전문가에 대한 의존성이 크고 기존에 알려지지 않은 공격 방법들에 대한 탐지 및 방어에는 약점을 보이고 있다. 이런 문제점을 해결하기 위한 방법으로 기계가 공격 방법들의 일정한 패턴을 학습하고 이를 토대로 공격 탐지 및 종류를 구분해 내는 기계학습 기반의 침입탐지/방지 시스템에 대해 연구자들의 관심이 높다 [1]-[3].

기계학습 방법에 기반한 침입탐지/방지 시스템의 개발과 관련해 가장 중요한 요소 중 하나는 공격의 종류를 표현하는 특징 조합의 발굴이다. 네트워크의 패킷이나 시스템의 로그로부터 다양한 특징들이 발굴되고 제시되고 있으나 제시된 특징들을 객관적이고 공정하게 평가하기 위한 공유된 데이터 집합에 대한 요구가 관련

Manuscript received February 9, 2016, revised April 15, 2016, accepted April 25, 2016
ISSN 2465-8111(Print), 2466-0124(Online), DOI http://dx.doi.org/10.18770/KEPCO.2016.02.02.311

연구자 사회에 끊임없이 제기되어 왔다. 이러한 요구에 부응해 침입 탐지/방지 시스템들의 객관적이고 공정한 비교가 가능하도록 KDD'99 데이터 집합이 MIT Lincoln 연구실에 의해 제공되었다 [4]. 이후 많은 연구자들이 자신들이 제안한 침입 탐지 방법들의 성능 평가를 위해 KDD'99 데이터를 사용하였다. 하지만 KDD'99 데이터는 규모가 너무 크고 중복된 데이터가 많으며 특정 공격 방법에 데이터가 편중되어 있는 등 많은 문제점을 가지고 있어서 그대로 사용하기에는 많은 제약이 있다. 이러한 KDD'99 데이터의 문제점들을 해결하고 다양한 침입 탐지/방지 시스템 간의 공정한 비교를 위해 NSL_KDD 데이터가 Tavallaee 등 [5][6]에 의해 제공되었다.

이들 두 데이터 집합은 다양한 종류의 공격을 표현하기 위해 총 41개의 특징을 사용하고 있다. 41개의 특징들 모두를 사용해 개별 공격들을 표현하고 이를 학습시켜 실시간 침입 탐지에 사용하기에는 특징 벡터의 크기가 너무 커서 부적절하다는 문제가 제기되었다. 이러한 문제의식하에 공격 유형을 판별하는데 실제적으로 중요한 특징들만을 조합하여 사용하려는 연구들이 최근 관심을 모으고 있다. 하지만 제시된 방법들은 개별적 특징들의 정보 획득 (information gain) [7] 이나 의존성 비율 (dependency ratio) [8], 상관계수 [9] 등과 같은 공격 유형과의 상관성을 분석하고 특성에 순위를 매겨 낮은 순위의 특성들을 제거하는 방식에 기반하고 있다. 하지만 특징들의 조합적 특성은 개별 특징들의 단순 합과는 다른 창발적 (emergent) 효과를 시스템에 가져올 수 있다는 점에서 이러한 접근 방법에 많은 문제가 있음을 알 수 있다. 대부분의 기존 연구들이 개별 특성들의 상관성 분석에 의존하는 가장 큰 이유는 주어진 특징들로부터 선발 가능한 특징 조합의 수가 너무 많아 (예를 들어 41개의 특징인 경우 조합 가능한 특징의 종류는 2^{41}-1개나 된다.) 모든 특징 조합을 하나씩 실험을 통해 평가하기가 현실적으로 불가능하기 때문이다.

이러한 특징의 최적 조합 문제를 해결하기 위한 방법으로 메타 휴리스틱 알고리즘에 기반한 방법이 Kang 등 [10]에 의해 제시되었다. 이 연구는 대상이 되는 공격의 범주를 서비스 거부 공격 (DoS: Denial of Service)에 한정 하고 그 종류에 상관없이 DoS 공격 여부만을 판별하는 침입탐지/방지 시스템을 위한 최적 특징 추출 알고리즘을 제안하였다. 제안한 방법은 41개의 특징 조합 모두를 사용한 경우와 정확성 측면에서 별차이가 없는 20여개의 특징만으로 구성된 특징조합을 찾아내는데 성공하였다. 하지만 해결하고자 하는 문제가 DoS 공격 유무의 판별이라는 2클래스 문제라는 점에서 공격 방법에 따라 대처 방법도 달리해야하는 현실을 생각할 때 많은 한계가 있다.

본 논문은 KDD'99 데이터 집합과 NSL_KDD 데이터 집합이 포함하고 있는 6가지 DoS 공격을 구분하고 정상 트트래픽으부터 이들을 탐지하기 위한 최적 특징 선택 알고리즘을 제시하고자 한다. 제시하는 방법은 다중 시작 지역탐색 알고리즘을 사용해 특징 조합을 선택하는 한편, 발굴한 특징 조합과 함께 높은 정확성를 보여주는 기계학습 방법을 찾고자 베이지 분류기, 인공신경망, 서포트 벡터 머신 등 잘 알려진 기계학습 방법들을 사용해 실험하였다.

본 논문의 구성은 다음과 같다. 2장에서는 NSL_KDD 데이터 집합의 구성 및 특성에 대해서 설명한다. 3장에서는 제안하는 특징 선택 알고리즘을 설계한다. 3장과 4장에서는 선택된 특징의 성능을 NSL_KDD 데이터가 제공하는 모든 특징들을 사용하는 경우와 다중 퍼셉트론을 이용해 비교하고 선택된 특징 조합과 가장 잘 어울리는 기계학습 방벙을 찾고가 세가지 학습방법간의 성능을 비교한다. 마지막으로 5장에서 결론 및 앞으로의 연구 방향 등에 대해 기술한다.

II. 데이터

제안한 특징 선택 알고리즘에 의해 얻어진 특징들의 성능 측정을 위해 NSL_KDD 데이터 집합 [6]을 사용하였다. 침입탐지 시스템의 성능 평가를 위해 기존에 널리 사용되어 오던 KDD'99 데이터는 [4] 약 500만개의 훈련 데이터와 30만개의 테스트 데이터로 구성되어있으며 4개의 범주 (Denial of service attack, User to root attack, Remote to local attack, Probing attack)로 공격 방법이 구별된다. 또한 41개로 이루어진 특징들은 기본 특징, 내용 특징, 트래픽 특징 등 3개의 범주로 나눠진다. 우선 기본 특징들로는 duration, protocol_type, service 등이 있으며 주로 TCP/IP 연결에서 추출할 수 있는 속성들이다. 내용 특징으로는 num_failed_logins, logged_in, num_compromised, su_attemped 등이 있으며 실패한 로그인 시도 등 의심스러운 행동을 탐지 할 수 있는 속성들과 관련이 있다. 마지막 트래픽 특성은 지난 2초간의 시간 윈도우를 이용해 연결들을 조사한 것으로 대상에 따라 same host 특징들과 same service 특징들로 구별된다. same host 특징에는 serror_rate, rerror_rate 등이 있고 same service 특징에는 srv_serror_rate, srv_rerror_rate 등이 있다.

하지만 KDD'99 데이터 집합은 지나치게 규모가 커서 모든 데이터를 사용하기 힘들다는 이유 때문에 연구자들의 주관적 판단에 따라 임의적으로 선별되어 사용됨으로써 연구 간의 객관적인 평가가 힘들었다. 또한 데이터 구성을 면밀히 분석한 연구들에 의해 데이터 내에 많은 중복된 데이터들이 존재함이 밝혀지면서 연구결과에 특정 데이터에 대한 편향되는 현상이 있을 수 있다는 지적들이 제기되었다.

NSL_KDD 데이터 집합은 KDD'99 데이터 집합이 가지고 있는 이러한 문제점들을 보완한 것으로 M. Tavallaee 등 [5]에 의해 제안 되었으며 기본적으로 KDD'99 데이터 집합에 기반하고 있지만 다음과 같은 점이 개선되었다. 우선 KDD 훈련 데이터나 테스트 데이터가 가지고 있는 중복성을 제거하여 시스템이 빈도가 높은 데이터에 편향되는 (bias) 현상을 방지하였다. 또한 각 부류의 데이터 집합을 난이도에 따라 크기를 정함으로써 성능 평가에 객관성을 높일 수 있다. 마지막으로 훈련 및 테스트 데이터의 수를 적정하게 제한함으로써 전체 집합을 대상으로 실험이 가능하도록 하여 부분 집합 선택시에 발생하는 자의성을 제거하여 실험 결과끼리의 공정한 비교를 가능하게 하였다.

Table 1. The compostion of data set

	Normal	Neptune	Teardrop	Smurf	Pod	Back	Land
Training data	67344	41214	892	2646	201	956	18
Test data	9711	4657	12	665	41	359	7

본 논문은 서비스 거부 공격 (DoS)만을 대상으로 필요최적의 특징 집합을 추출하는 것을 목적으로 하고 있기 때문에 전체 데이터에서 정상인 경우와 6가지 DoS 공격에 해당하는 데이터만을 따로 추출하여 사용하였다. 이렇게 추출된 데이터의 크기는, 훈련 데이터의 경우 113271개 이고 테스트 데이터는 15452개로 구성되어있다. 한편 NSL_KDD 데이터엔 DoS 공격의 유형이 총 6가지가 있으며 정상인 경우를 포함한 데이터 구성은 Table 1과 같다.

DoS 공격에 비해 정상 트래픽에 대한 데이터가 많으며 DoS 공격 유형간의 데이터 편차도 크다는 사실을 알 수 있다. 특히 Teardrop의 경우 훈련집합의 데이터 수에 비해 평가집합의 데이터 수가 지나치게 작은데, 이는 NSL_KDD 데이터 집합에서 수정되어야 할 부분이다.

III. 특징선택 알고리즘

본 장에서는 정규화를 위한 데이터의 전처리를 포함해 제안하는 특징 선택 알고리즘에 대해 자세히 기술한다.

A. 데이터 전처리

NSL_KDD 데이터에는 다양한 형태의 특징들이 존재한다. 따라서 이들을 기계학습의 입력으로 사용하기 위해서는 사용 전에 일정 범위의 수치적 특징으로 정규화해야 할 필요가 있다. protocol_type, servie, flag와 같이 숫자가 아닌 특징이 존재하는 한편 단위가 10억이 넘는 src_bytes, dst_bytes와 같은 특징이 존재한다. 이러한 특징들은 숫자 데이터로 변환이 필요하고 특정 특징에 의해 분류기가 왜곡되지 않도록 하기 위해 정규화가 필요하다. 정규화는 [2][10]의 논문을 참조하여 다음과 같은 방법으로 하였다.

1) 심볼릭 특징 – 각 특징이 가지고 있는 종류에 0부터 양의 정수를 부여하고 이를 [0, 1]로 선형 정규화 : 예) protoco_type의 경우 tcp, udp, icmp 세 종류가 있는데, 각각에 대해 0, 1, 2 값을 부여하고 이를 0과 1사이의 값으로 선형 정규화

2) src_bytes, dst_bytes와 같이 아주 큰 값을 갖는 특징 : 10을 베이스로하는 로그값으로 정규화

3) 이진 값을 갖는 특징 : 0과 1을 그대로 사용

4) 수치를 값으로 갖는 나머지 특징들 : [0, 1] 사이 값으로 선형 정규화

41개의 특징 요소로 구성된 특정 개체의 i 번째 특징의 특징 값, s에 대한 선형 정규화($N(s)$)는 훈련 데이터와 테스트 데이터 모두를 상대로 각 특징들의 최소값과 최대값을 구한 후 다음 식에 의해 정규화하였다.

$$\frac{s-\min(f_i)}{\max(f_i)-\min(f_i)} \quad (1)$$

여기서 $min(f_i)$와 $max(f_i)$는 각각 해당 특징의 최소값과 최대값을 나타낸다.

B. 특징 선택 알고리즘

[10]은 기존에 제안된 특징 선택 알고리즘과 달리 특징 선택을 최적화 문제로 정의하고 이를 해결하기 위한 방법을 제시한다. 41개의 특징들로부터 가능한 특징 조합은 2^{41}-1개나 된다. 이는 개별 특징조합을 대상으로 각각의 성능을 측정하고 평가하는 일이 사실상 불가능함을 의미한다. 이러한 이유로 기존에 제시된 대부분의 알고리즘들은 개별 특징들을 대상으로 데이터와의 연관성을 측정하고 등급을 매겨 저순위의 특징들을 제거하는 방식에 의존하였다. 최적 특징 조합 선택 문제는 다음과 같이 정의 된다 [10].

정의 1. *최적 특징 조합 선택 문제*

주어진 특징 집합 $f = \{f_1, f_2, f_3, \ldots, f_n\}$과 비용함수 $C:f \rightarrow q$ $(0 \le q)$ 로부터 최소 비용을 보장하는 특징 부분 집합 f'을 찾아라.

최적 특징 조합 선택 문제를 해결하기 위한 방법으로 조합 최적화 문제에 자주 사용되는 다중시작 지역탐색 알고리즘을 설계하였다. 우선 지역탐색 알고리즘에 사용되는 해는 아래 Eq. 2와 같이 길이가 41인 2진 벡터 v로 표현된다. 선택된 특징은 1로 그렇지 않는 특징은 0 값을 갖는다.

$$v = < v_1, v_2, v_3, \ldots, v_{41}>, \quad \text{where } v_i \in \{0, 1\}, 0 \le i \le 41 \quad (2)$$

예를 들어 41개의 특징을 사용하는 특징 부분 집합은 1이 41개인 특징 벡터, v =<1, 1, 1, ..., 1> 로 대표된다.

다중시작 지역탐색 알고리즘과 같은 최적화 휴리스틱 알고리즘의 성능을 크게 좌우하는 요소 중 하나는 개별 해를 평가하기 위한 비용함수이다. 즉, 비용함수를 어떻게 정의하느냐에 따라 알고리즘의 성능이 크게 영향을 받는다. 본 논문이 제안하는 개별 해에 대한 비용함수는 [10]이 제시한 방법과 동일하다. 비용함수의 기본 아이디어는 주어진 해, 즉 선택된 특징들만을 이용해서 군집화를 시도했을 때의 분류 정확성을 이용한다는 데 있다. 개별 해가 제시하는 특징 조합만을 가지고 학습 데이터를 군집화 알고리즘을 사용해 군집화한 후 학습 데이터를 얼마나 정확하게 분류했는지를 개별해의 비용으로 사용한다.

이때 사용하는 군집화 알고리즘은 k-평균 군집화 알고리즘이다. k-평균 군집화 알고리즘은 주어진 데이터를 사전에 정해진 k 개의 군집으로 분할하는 알고리즘으로 다음과 같은 군집들의 분산 합(V)을 최소화하는 k 개의 C_i $(1 \le i \le k)$들을 찾는 것이 목적이다.

$$\mathrm{V} = \sum_{i=1}^{k} \sum_{j \in C_i} |x_j - u_i|^2 \quad (3)$$

Table 2. Pseudo code for feature selection algorithm

```
Algorithm: Feature selection algorithm
Input : Training data set, the number of initial solutions: R
Output : Combination of features: v_f
1.   v_f ← Null;    // final solution
2.   c_f ← Max Integer;

3.   for i = 0 to R:
4.     Generate an initial solution, v_i;
5.     v_b ← v_i;
6.     Calculate the cost of initial solution, C(v_b);

7.     while(1) begin:
8.       Find neighbor solutions of v_b which have one bit
                 different from v_b;
9.       Calculate the cost of each neighbor solution;
10.      Select the best one, v_n, from neighbor solutions;
11.      if C(v_b)  >  C(v_n):
12.        v_b ← v_n;
13.      else:
14.        break;
15.    end // for while loop

16.    If c_f > C(v_b):
17.      v_f ← v_b;
18.      c_f  ← C(v_f);
19.  end // for for loop
```

여기서 C_i는 i 번째 군집을 나타내고 u_i는 i 번째 군집의 중심을 가리킨다. k-평균 군집화 알고리즘은 초기에 k 개의 군집 중심을 임의로 선정한 후 대상 데이터를 유클리디안 거리를 이용해 가까운 중심으로 집단화 한다. 그다음 분류된 군집 각각으로부터 새로운 군집 중심을 계산한다. 이러한 과정을 군집에 더 이상의 변화가 없을 때까지 반복한다. 논문이 다루고 있는 특징 선택 문제는 정상인 경우와 DoS 공격 6가지 경우를 분류하는 문제에 해당하므로 k = 7이다. 그리고 주어진 해의 비용에 해당하는 군집의 정확성은 훈련 집합을 대상으로 다음과 같이 계산한다.

$$\varphi(x) = \begin{cases} 1 & if\ \ p(x) = q(x) \\ 0 & otherwise \end{cases} \tag{4}$$

$\varphi(x)$는 입력 데이터 x에 대한 비용함수로 군집화에 의한 분류값 $p(x)$와 원 소속 $q(x)$ 값이 같으면 1이고 그렇지 않으면 0을 갖는다. 주어진 해 v 에 대한 비용, $C(v)$는 N 개인 전체 훈련 집합을 대상으로 다음 식에 의해 구한다.

$$C(\upsilon) = \frac{1}{\sum_{i=1}^{N} \varphi(x_i)} \tag{5}$$

한편, 다중시작 지역탐색 알고리즘에서 다음 번 탐색의 대상이 되는 이웃 해는 해당 해와 1 비트 다른 해로 정의한다. 따라서 임의의 해에 대한 이웃 해는 총 41개가 가능하며, 탐색은 이러한 41개의 해 중 $C(v)$값 가장 큰 해를 다음 탐색의 대상으로 삼고 동일한 과정을 반복한다. 특정 반복에서 어떠한 이웃해도 현재 해의 정확성 보다 높은 정확성을 갖지 못한다면 해당 해를 알고리즘의 최적해로 삼고 더 이상의 탐색은 중단한다. 하지만 [10]에서 제시한 지역탐색 알고리즘은 무작위로 선택된 초기해에의 특성에 따라 지역 최적에 빠질 염려가 크다. 따라서 본 논문은 이를 극복하기 위한 대표적인 방법으로 다중시작 지역탐색 알고리즘을 설계하였다. 다중시작 지역탐색 알고리즘이 지역탐색 알고리즘과 다른 점은 사전에 정의된 수만큼의 다양한 초기해를 무작위로 선택하고 지역탐색 알고리즘을 실시한 다음 가장 좋은 해를 최종 해로 한다는 점이다.

Table 2에 다중시작 지역탐색 알고리즘에 기반한 특징 선택 알고리즘에 대한 의사코드를 제시한다.

IV. 실험 및 결과 분석

A. 특징선택 알고리즘의 성능 평가

제안한 특징 선택 알고리즘의 성능 평가를 위하여 20개의 해를 생성하였다. 이어서 각 해의 정확성 분석을 위해 해가 가진 특징들(1값을 가진)만을 이용해 훈련 데이터와 테스트 데이터를 원본 데이터로부터 수정하여 생성하였다.

선택된 특징 조합의 성능은 알려진 기계학습 방법 중 다층 퍼셉트론을 사용하여 평가하였다. 다층 퍼셉트론은 입력층, 은익층, 출력층의 일반적인 형태인 3층 구조로 설계하였다. 입력층의 노드 개수는 사용하는 특징 조합의 크기와 같고 출력층의 노드 개수는 7이다. 은익층의 노드 개수는 여러 번의 실험을 통해 가장 높은 성능을 보여주는 개수로 설정하였다. 출력층 노드들의 목표치는 정상 트래픽의 경우 <0.9, 0.1, 0.1, 0.1, 0.1, 0.1, 0.1> 를 갖도록 하고 DoS 각 공격 패턴에 따라 상응하는 출력 노드의 목표치를 0.9 나머지 노드의 목표치들은 0.1을 갖도록 하였다. 예를 들어 neptune인 경우 출력층의 목표치는 <0.1, 0.9, 0.1, 0.1, 0.1, 0.1, 0.1> 이다. 은익층과 출력층의 각 노드는 수식 Eq. 6의 시그모이드 함수를 사용하였고 다양한 학습률과 모멘텀을 가지고 오류 역전파 알고리즘을 사용하여 학습하였다.

$$S(v) = \frac{1}{1+e^{-v}} \tag{6}$$

훈련집합 전체를 대상으로 한 번의 학습이 이루어진 과정을 에포크 (epoch)라 하는데 종료조건은 연속한 두 에포크 간의 정확성 사이에 개선이 0.1% 이하인 경우 학습을 종료하였다.

선택 알고리즘에 의해 얻어진 특성 조합의 성능 분석은 침입탐지 시스템에 대한 연구 논문들에서 일반적으로 사용되는 정확성을 사용하였다. 정확성은 테스트 집합을 대상으로 정상 트래픽과 6가지 DoS 공격 패턴을 정확히 판별한 비율을 말한다. 한편 실시간 침입탐지/방지 방법을 평가하기 위한 지표중 정확성 못지 않게 중요한 것들은 판별 시간과 관련있는 것으로 모델 생성 시간과 테스트 시간, 특징 조합의 크기 등이다.

Table 3에 생성된 20개 특징 조합해을 대상으로 멀티 퍼셉트론을 이용했을 때의 정확성 및 학습시간, 훈련 데이터를 대상으로한 테스트 시간, 특징 조합의 크기를 평균하여 표준편차와 함께 나타내었다.

Table 3으로부터 제안한 특징 선택 알고리즘을 이용해 얻은 특징 조합의 평균 정확성은 96.77%로 41개 모

Table 3. The results of experiments, which shows accuracy, taken time to training, taken time to testing and feature size

	Accuracy	Time taken to train (sec)	Time taken to test (sec)	Size of feature set
Selected feature set	96.77±1.52	315.98±49.31	0.71±0.11	19.05±3.47
41 features	96.98	799.65	1.48	41

든 특징 요소를 사용하는 경우의 정확성 96.98%에 비해 약간 낮지만 가장 높은 경우는 98.8%를 보이는 등 정확성에서 의미있는 차이가 있다고 보기는 어려운 결과다. 하지만 제안한 방법에 의해 선택된 특징 조합의 평균 크기는 19.05로 NSL_KDD 데이터가 예비한 특징 수의 1/2 이하였다. 또한 모델 생성을 위한 학습 시간이나 훈련집합을 대상으로 한 테스트 시간도 제안한 특징 선택 알고리즘을 이용해 얻은 특징 조합의 것이 두배 이상 작다는 사실을 확인할 수 있다. 이는 제안한 알고리즘에 의해 구해진 특징조합을 사용할 경우 모든 특징조합을 사용했을 때와 비슷한 정확성을 보장하면서도 실시간 침입탐지/방지 시스템에 보다 적합하다는 사실을 확인할 수 있다. 이러한 실험 결과로부터 지역탐색 알고리즘 기반의 특징선택 알고리즘이 DoS 공격 유무만을 판별했던 이전 연구 결과 [10]와 비슷한 결과를 가져다 준다는 사실을 확인하였다. Table 4는 98.8%의 정확성을 보여준 특징 조합의 특징들을 보여준다.

B. 기계학습 비교

다음으로 기계학습 기반의 침입탐지/방지 시스템의 높은 성능 보장을 위해 본 특징 선택 알고리즘과 가장 잘 어울리는 기계학습 방법을 찾고자 잘 알려지 세가지 기계학습 방법, 인공 신경망, 베이지 분류기, 서포트 벡터 머신 (SVM)을 대상으로 정확성을 비교하였다.

베이즈 분류기 [11]는 우도 (likelihood)와 사전확률로부터 사후확률을 계산하는 베이즈 정리에 기반을 둔 지도학습 기반의 대표적인 기계학습 방법이다. 그리고 SVM은 Vapnik [12]에 의해 제안된 지도학습 기반의 기계학습 방법으로, 분류를 위한 결정 초평면의 결정에 서포트 벡터의 간격 (margin)을 고려함으로써 다른 기계학습 방법에 비해 일반성을 갖는다는 평가를 받고 있다. 비선형 SVM을 사용하기 위해 커널함수는 $K(x,y)=(x\cdot y+1)^p$를 사용하였다. 다음 Table 5는 제안한 특징선택 알고리즘을 이용해 생성한 20개 특징조합을 대상으로 세가지 기계학습을 사용해 정확성을 평균한 것이다.

Table 5에서 알 수 있듯이 평균 정확성은 SVM이 인공 신경망이나 베이즈 분류기에 비해 높다는 사실을 알 수 있다. 하지만 표준편차를 고려할 경우 인공 신경망의 정확성이 가장 안정적이며 평균치에서도 SVM의 평균치와 무시할 수 있을 정도의 차이를 보이고 있다. 한편 시스템 구현에 있어서 인공 신경망의 구현이 SVM에 비해 용이하다는 점에서 실제 사용에 있어서는 SVM 이외에 충분히 고려해 볼 만한 기계학습 방법이 될 수 있음을 알 수 있다.

Table 4. A feature set which shows accuracy of 98.8%

num	feature name	type	min value	max value
1	duration	numeric	0	54451
2	protocol_type	symbolic	0	2
4	flag	symbolic	0	10
5	src_bytes	numeric	0	89581520
8	wrong_fragment	numeric	0	3
10	hot	numeric	0	101
13	num_compromised	numeric	0	7479
14	root_shell	numeric	0	1
15	su_attempted	numeric	0	2
16	num_root	numeric	0	7468
17	num_file_creations	numeric	0	100
19	num_access_files	numeric	0	9
23	count	numeric	0	511
26	srv_serror_rate	numeric	0.0	1.0
28	srv_rerror_rate	numeric	0.0	1.0
29	same_srv_rate	numeric	0.0	1.0
33	dst_host_srv_count	numeric	0	255
34	dst_host_same_srv_rate	numeric	0.0	1.0
35	dst_host_diff_srv_rate	numeric	0.0	1.0
36	dst_host_same_src_port_rate	numeric	0.0	1.0
38	dst_host_serror_rate	numeric	0.0	1.0
39	dst_host_srv_serror_rate	numeric	0.0	1.0
40	dst_host_rerror_rate	numeric	0.0	1.0
41	dst_host_srv_rerror_rate	numeric	0.0	1.0

Table 5. The performance comparisons between 3 machine learning methods

	Multi-layer perceptron	Bayesian classifier	SVM
Accuracy	96.77%(±1.52)	81.30%(±19.12)	97.17%(±2.13)

V. 결론

본 논문은 NSL_KDD 데이터를 대상으로 정상 트래픽과 6가지 서비스 거부 공격을 탐지하기 위해 최적의 특징 조합 선택을 위한 알고리즘을 제시하였다. 제시한 알고리즘은 특징 조합 선택 문제를 비용 최소화를 목적으로 한 조합 최적화 문제로 정의하고 최적 해를 찾기 위해 다중시작 지역 탐색 알고리즘을 응용하였다. 이때 사용하는 비용함수는 조합 해가 제시한 특징들만을 가지고 *k*-평균 군집화 알고리즘을 실행하여 얻은 군집해의 분류 정확성을 응용해 정의하였다. 한편 제시한 알고리즘에 의해 선출된 특징 조합들의 정확성과 효율성을 확인하고자 인공 신경망을 설계하고 41개로 구성된 모든 특징을 사용했을 때와 비교하였다. 실험 결과로부터 선출된 특징들의 평균 크기는 약 19였고 모든 특징들을 사용했을 때와 비슷한 정확성을 보여 주었으며 학습 시간과 테스트 시간에서는 2배 이상의 효율성을 확인할 수 있었다. 한편 제안한 특징 선택 알고리즘과 가정 적합한 기계학습 방법을 찾고자 인공 신경망, 베이즈 분류기, SVM을 사용하여 20개의 특징 조합을 대상으로 정확성을 측정하였다. 정확성의 순서는 SVM, 인공 신경망, 베이즈 분류기 순서였으나 SVM과 인공 신경망의 차이는 크지 않았고 표준편차를 고려했을 때 인공 신경망이 보다 안정적이었다. 또한 구현상의 용이성을 고려하면 SVM 외에 인공 신경망도 제안한 특징 선택 알고리즘과 함께 침입탐지/방지 시스템에 적합함을 확인하였다.

ACKNOWLEDGMENT

이 논문은 한국전력공사의 재원으로 기초전력연구원의 2015년 선정 기초연구개발과제의 지원을 받아 수행된 것임. [과제번호 : R15XA03-63]

REFERENCES

[1] S. Paliwal and R. Gupta, "Denial-of-Service, Probing & Remote to User (R2L) Attack Detection using Genetic Algorithm," International Journal of Computer Application, 60(19), 2012.

[2] M. Sabhnani and G. Serpen, "Application of Machine Learning Algorithms to KDD Intrusion Detection Dataset within Misuse Detection Context," Proc. of International Conference on Machine Learning: Models, Technologies, and Applications, 23-26 June 2003, Las Vegas, Nevada, USA, pp. 209-215, 2003.

[3] C. F. Tsai, Y. F. Hsu, C. Y. Lin, and W. Y. Lin, "Intrusion detection by machine learning: a review," Expert System with Applications 36(10), 2009.

[4] KDD Cup 1999. Available on:http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html, 2007.

[5] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A Detailed Analysis of the KDD CUP 99 Data Set," Proc. 2009 IEEE Int. Conf. Comput. Intell. Security Defense Appl. CISDA 2009, pp. 53-58.

[6] NSL_KDD data set. Avalilable on: http://nsl.cs.unb.ca/NSL-KDD/

[7] H. G. Kayacik, A. N. Zincir-Heywood, and M. I. Heywood, "Selecting Features for Intrusion Detection: A Feature Relevance Analysis on KDD 99 Intrusion Detection Datasets," in Thrid Annual Conference on Privacy, Security and Trust, St. Andrews, New Brunswick, Canada, 2005.

[8] A. A. Olusola, A. S. Oladele, and D. O. Abosede, "Analysis of KDD '99 Intrusion Detection Dataset for Selection of Relevance Features," in Proc. of the World Congress on Engineering and Computer Science, Vol. 1, 2010.

[9] S. Parazad, E. Saboori, and A. Allahyar, "Fast Feature Reduction in Intrusion Detection Datasets," in MIPRO, Proceedings of the 35th International Convention, pp.1023-1029, 2012.

[10] S. H. Kang, and K. J. Kim, "A feature selection approach to find optimal feature subsets for the network intrusion detection system," Cluster Computing, 2015. DOI 10.1007/s10586-015-0527-8

[11] G. H. John, and P. Langley, "Estimating continuous distributions in Bayesian classifier," Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence, Montreal, QU, Canada, 1995.

[12] C. J. C. Burges, "A tutorial on support vector machine for pattern recognition," Data Mining and Knowledge Discovery 2, pp.121-167, 1998.