Journal of the Korea Academia-Industrial cooperation Society
Vol. 15, No. 8 pp. 5256-5262, 2014

http://dx.doi.org/10.5762/KAIS.2014.15.8.5256
ISSN 1975-4701 / eISSN 2288-4688

추세 모형 기반의 예측 모델을 이용한 비정상 트래픽 탐지 방법에 관한 연구

장상수[1*]
[1]한국인터넷진흥원

Study of The Abnormal Traffic Detection Technique Using Forecasting Model Based Trend Model

Sang-Soo Jang[1]
[1]**Korea Internet and Security Agency**

요 약 최근 국가기관, 언론사, 금융권 등에 대하여 분산 서비스 거부(Distributed Denial of Service, DDoS) 공격, 악성코드 유포 등 무차별 사이버테러가 발생하고 있다. DDoS 공격은 네트워크 계층에서의 대역폭 소모를 주된 공격 방법으로 정상적인 사용자와 크게 다르지 않는 패킷을 이용하여 공격을 하기 때문에 탐지 및 대응이 어렵다. 이러한 인터넷 비정상적인 트래픽이 증가하여 네트워크의 안전성 및 신뢰성을 위협하고 있어 비정상 트래픽에 대한 발생 징후를 사전에 탐지하여 대응할 수 있는 방안의 필요성이 대두되고 있다. 본 연구에서는 비정상 트래픽 탐지 기법에 대한 현황 및 문제점을 분석하고, 예측 방법인 추세 모형, 지수평활법, 웨이브렛 분석 방법 등을 비교 분석하여 인터넷 트래픽의 특성을 실시간으로 분석 및 예측이 가능한 가장 적합한 예측 모형을 이용한 탐지 방법을 제안하고자 한다.

Abstract Recently, Distributed Denial of Service (DDoS) attacks, such as spreading malicious code, cyber-terrorism, have occurred in government agencies, the press and the financial sector. DDoS attacks are the simplest Internet-based infringement attacks techniques that have fatal consequences. DDoS attacks have caused bandwidth consumption at the network layer. These attacks are difficult to detect defend against because the attack packets are not significantly different from normal traffic. Abnormal traffic is threatening the stability of the network. Therefore, the abnormal traffic by generating indications will need to be detected in advance. This study examined the abnormal traffic detection technique using a forecasting model-based trend model.

Key Words : Abnormal Traffic, DDoS, Malicious Code, Forecasting, Time Series, Trend Model

1. 서론

인터넷 트래픽은 복잡한 프랙탈(Fractal) 패턴 및 사용자 접속패턴에 따라 상이한 모습을 보이는 특성을 가지고 있기 때문에 정상과 비정상 트래픽을 판단하기 어렵다. 인터넷 트래픽 분석기법은 상세 분석보다 실시간 분석을 통한 판단이 요구된다. 이는 DDoS 공격과 같은 사이버 공격의 경우 과도한 트래픽은 네트워크 속도를 현저하게 지연시키고, 감염된 시스템은 과부하로 인하여 다운된다. 즉, 짧은 시간 내 분석이 선행되지 않으면 능동적인 대응기능과 그 효과측면에서 보안요구사항을 만족시키지 못한다. 본 연구에서는 비정상 트래픽의 유형과 기존의 통계적인 방법을 이용한 비정상 트래픽 탐지 기법에 대하여 알아보고 DDoS 공격 등 비정상 트래픽에 대해 시계열 예측 모형중 하나인 추세 모형 기반의 예측 모델을 이용한 탐지 방법을 제안하고자 한다.

*Corresponding Author : Sang-Soo Jang(KISA)
Tel: +82-10-8903-8953 email: ssjang0116@gmail.com
Received May 9, 2014 Revised (1st July 10, 2014, 2nd August 6, 2014) Accepted August 7, 2014

2. 관련 연구

2.1 비정상 트래픽 탐지 기법

비정상 트래픽 탐지 기법은 크게 2가지로 구분할 수 있다. 첫 번째는 시그너처 기반의 탐지 기법이고 두 번째는 비 시그너처 기반의 탐지 기법이다. 시그너처란 비정상을 탐지하는 룰 셋으로 정의할 수 있다. 비정상을 나타내는 행동에 대한 정규화된 표현식이다. 시그너처 기반의 탐지 기법에는 패턴 매칭 기법과 MIB 기반 모델, 프로토콜 행위에 대한 시그너처 모델 등이 있으며 비 시그너처 탐지 모델로는 플로우 기반 모델, 시계열 모델, 신호처리 모델 등이 있다[2].

2.2 시계열 예측 모형

통계를 이용한 시계열(time-series) 분석의 가장 큰 목적은 미래 예측에 있다. 시계열 분석이란 어떠한 자료가 과거 시점에서 얻어진 관찰 값 등의 변화 패턴에 따라 앞으로도 변화할 것이라는 전제하에 향후 시점의 값을 예측하는 것이다. 시계열 분석의 기본적인 방법은 시계열 데이터의 변동을 몇 개의 구성 요소로 나누고, 그 중 가장 기본이 되는 구성 요소의 변동에 대하여 이미 알려진 곡선 가운데 가장 적합한 것을 적용한다. 시계열 분석에서 다루고자 하는 자료들은 시간의 흐름에 따라 일정한 간격으로 관측된 자료들로써, 시계열 자료라 한다. 시계열 자료의 특징은 한 시점에서 관측된 자료가 그 이전의 관측된 자료들에 의존한다는 것이다[3-5].

2.2.1 예측 모형

시계열 예측 모형은 시계열 자료가 생성된 시스템 또는 확률 과정을 모형화하여 시스템 또는 과정을 이해 또는 제어할 수 있도록 하는 것이다. 시계열자료의 분석 목적은 예측(forecasting)과 시스템의 제어이며 시계열 분석 과정은 모형(model)의 식별, 추정, 진단 3단계로 이루어진다. 관측된 과거자료에 포함된 정보를 이용하여 예측에 필요한 경험적 법칙을 추정하여 예측하는 방법에는 전제 조건으로 과거에 대한 정보가 존재해야하며 이 정보는 양적인 자료로 표현 가능하며 또한 과거의 현상이 미래에도 지속된다는 가정이 필요하다. 이러한 예측 방법들로는 대표적인 것이 추세 모형, 지수평활법, 웨이브렛 분석 방법 등을 이용하는 방법이 있다[6-8]. 본 연구에서는 이 3가지 예측 방법들을 비교 분석하여 대용량 인터넷 트래픽에 대한 비정상 트래픽 분석 기법으로 가장 적합한 모델을 제시하고자 한다.

2.4 추세 모형(Trend Model)

추세 모형을 이용한 분석에서는 시계열의 주요 구성요소를 모형에 얼마나 잘 흡수해 주느냐에 따라 분석 결과가 달라지며 시간의 변화에 따라 변동이 거의 없는 시계열 자료에 주로 사용하게 된다. 종류로는 다항추세모형, 비선형 추세 모형, 자기회귀오차 모형 등이 있다.

다항추세모형은 전통적으로 시계열 자료를 분석하기 위해 많이 사용되어 온 방법 중의 하나로 관찰값 Z_t를 시간의 함수로 표현하는 방법이다. 즉, $Z_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \cdots + \beta_k t^k + \epsilon_t$ 와 같이 관찰값 Z_t를 시간 t의 다항식으로 나타내는 것으로, 위와 같은 식을 다항추세모형이라 한다.

비선형 추세 모형은 시간의 흐름에 따라 초기에는 완만히 증가하다가 어느 시점부터 갑자기 급격히 증가하여 서서히 증가 추세가 완화되어 어떤 포화상태에 도달하는 양상을 나타내는 경우 사용되는 방법으로 일반적으로 선형추세 모형으로 선형변환을 통해 시계열회귀모형을 적합 시킨 후 다시 원래대로 변환함으로써 미래를 예측할 수 있다. 자기회귀오차모형은 회귀모형을 이용하여 자료를 분석할 때 일반적으로 오차항이 서로 독립이라는 가정을 전제로 하여, 시계열 자료를 회귀모형에 적합 시킬 때 오차들이 시간에 따른 자기상관관계를 갖는 경우에 적용하는 방법이다[4-6].

2.5 지수평활법(Exponential Smoothing Method)

시계열자료들이 변동폭이 큰 계절요인이나 추세만으로는 설명하기 힘든 불규칙요인을 포함하고 있는 경우에는 추세를 정확하게 파악하기 어렵다. 이러한 계절요인 또는 불규칙요인들을 자료들의 평활에 의해 제거하여 전반적인 추세를 뚜렷하게 파악할 수 있도록 해 주거나, 예측을 쉽게 할 수 있도록 해 주는 방법이 평활법이다. 지수평활법은 가장 최근 데이터에 가장 큰 가중치가 주어지고 시간이 지남에 따라 가중치가 기하학적으로 감소되는 가중치 이동 평균 예측 기법의 하나다[3,6].

2.6 웨이브렛(Wavelet) 분석

Wavelet은 디지털 신호처리 및 이미지 압축에 사용되

는 유용한 수학 함수이다. Wavelet의 근본 원리는 푸리에(Fourier) 분석과 비슷하며, 19세기 초반에 처음 개발되었다. 신호처리를 위해 Wavelet을 이용하면 잡음 속에 섞인 약한 신호를 복원할 수 있으며 이런 방법으로 처리된 이미지는 세부적인 내용에 흐릿함이 없이 깨끗하게 처리될 수 있다. Wavelet 분석 기법은 비교적 최근에 개발된 시계열 및 신호분석 도구이다. 기존의 Fourier 분석이 시간 영역에서 주파수 영역으로 변환 될 때 시간에 대한 정보가 사라진다는 단점을 보완하고자 제안되었다[2,9].

3. 추세 모형 기반의 예측모델 제안

3.1 예측 모형

3.1.1 추세 모형 구현

본 연구에서는 기존의 예측 방법들의 문제점을 해소하기 위하여 인터넷 트래픽의 특성을 실시간으로 분석하여 반영하고 예측 가능한 통계 추론 기법 중의 하나인 추세 모형 기반의 예측 모델을 이용한 비정상 트래픽 탐지 기법을 제시하고자 한다. 추세 모형의 구현은 측정점이

$$(x_1, f(x_1)), (x_2, f(x_2)), (x_3, f(x_3)) \ldots (x_n, f(x_n))$$

일 때 $p(x_i) = a_0 + a_1 x_i + a_1 x_i^2 + \ldots + a_1 x_i^m$ 된다. 이때

$$S = \sum_{i=1}^{n} \{p(x_i) - f(x_i)\}^2 = \sum_{i=1}^{n} \{a_0 + a_1 x_i + \ldots + a_m x_i^m - f(x_i)\}^2$$

도출되는데, 여기서 S가 최소가 되도록 하기 위한 편도함수를 0으로 둔다.

$$\frac{\partial S}{\partial a_j} \equiv 0 \; (j = 0, 1, \ldots, m)$$

이것을 행렬로 표현하고, 연립방정식 해법을 통해 근사식을 구한다.

$$\begin{bmatrix} n & \sum x_i & \cdots & \sum x_i^m \\ \sum x_i & \sum x_i^2 & \cdots & \sum x_i^{m+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum x_i^m & \sum x_i^{m+1} & \cdots & \sum x_i^{2m} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} \sum f(x_i) \\ \sum f(x_i) x_i \\ \vdots \\ \sum f(x_i) x_i^m \end{bmatrix}$$

3.1.2 가중치 부여

가장 최근 과거 데이터에 대한 가중치를 적용하는데 가중평균 w_i 산출방식은 다음과 같다.

$w_i =$ (요일 가중치) * (결정계수 가중치) * (주별(날짜) 가중치) * (특정요일 가중치) * (최근 패턴데이터 가중치)

가중평균을 구한다음 아래의 식을 이용하여 지수평활법을 적용한다.

$$f(t_i) = \sum w_i (a_{0,i} + a_{1,i} t^1 + a_{2,i} t^2 + \ldots + a_{m,i} t^m)$$

3.1.3 주요 트래픽 형식 구성

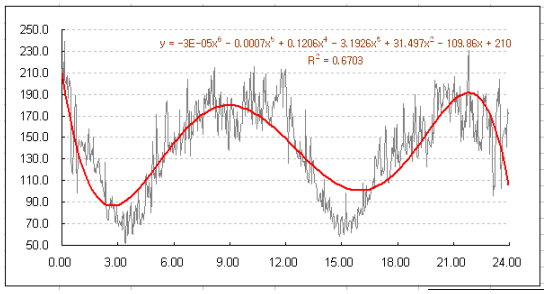
IF(intermediate frequency), BPS(bits per second) /PPS(packets per second), 주요AS(autonomous system) BPS/ PPS 등의 일별 대표 트래픽에 적용하였다.

[Fig. 1] Major Daily Traffic Type Sample

Fig. 1은 일별 대표 트래픽 형식의 예를 보이고 있다. 본 연구에서 제안하고자 하는 추세 모형의 구현모델은 현재 다중회귀방정식인 8차 다항식으로 계산하고 있으며, 수식은 다음과 같다.

$$y(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_m x^m$$
$$r_i = y(x_i) - f(i)$$
$$s = \sum_{i=1}^{m} \{y(x_i)^2 - f(i)\}^2$$

추세 모형을 네트워크의 수집 자료를 기반으로 적용한 예측모델의 시험 결과가 아래 Fig. 2와 같이 나타났다.



[Fig. 2] The Result of forecast model Measurement Based Trend model

3.1.4 비정상트래픽의 탐지

이상 트래픽 분석은 잘 알려진 네트워크 공격에 대해 임계값을 설정하고 이를 감지하는 방식의 오용 탐지(Misuse Detection), 정규 패턴과 비교하여 이상 패턴이 발생하면 이를 감지, 보고하는 프로파일 탐지(Profile Detection), 트래픽 볼륨과 관계없이 변화(Fluctuation) 비율에 따라 급격한 변동이 있을 경우 이를 감지하는 방식의 비정규적 패턴 탐지(Unusual Pattern Detection)로 구성된다. 시험 데이터에서 감지된 이상 트래픽은 요약 정보 별로 유형(Type)을 구분하였다[2]. Table 1은 이상 트래픽의 유형을 나타내고 있다.

[Table 1] Anomaly Traffic Type

Anomaly Type	Anomaly Description
Misuse (HostScan)	Too many connecting to target host
Misuse (TCPFlood)	Send many TCP packets to target host
Misuse (UDPFlood)	Send many UDP packets to target host
Misuse (ICMPFlood)	Too many sending ICMP echo request to broadcast
Misuse (TCPSYNFlood)	Too many sending sync packets
Misuse (NetworkScan)	Network Scanning
Profile (HighBps)	bps for pattern case
Profile (HighPps)	pps for pattern case
Profile (Port)	Port Profile case
Profile (As)	As Profile case
Jump (TrafficPort)	Application Port case
Jump (TrafficBps)	bps case
Jump (TrafficPps)	pps case

3.1.5 비정상 탐지(Anomaly Detection) 적용 방법

비정상 트래픽 탐지 적용 방안으로 오탐율(False Positive) 제한 방법과 스레시홀드(Threshold) 설정 방법을 적용 하였다.

1) 오탐율(False-Positive) 제한

스레시홀드 값(Threshold Value) 설정 시 오탐율 요소(False Positive Factor)를 설정하며 오탐율 요소는 5단계로 구성하였다.

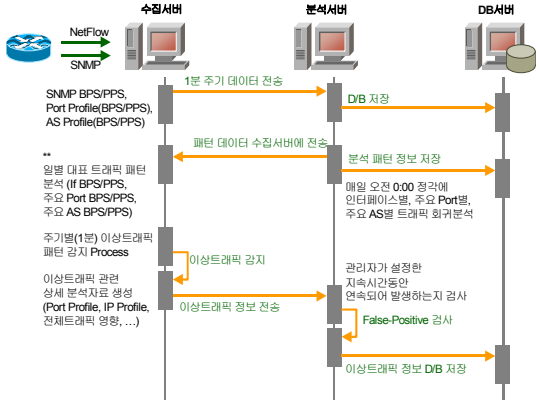
-Factor 1: 연속하여 1분간 지속
-Factor 2: 연속하여 2분간 지속
-Factor 3: 연속하여 3분간 지속
-Factor 4: 연속하여 4분간 지속
-Factor 5: 연속하여 5분간 지속

2) 스레시홀드(Threshold) 설정

각각의 인터페이스 별로 임계값을 설정한다. 요일별로 AS / Port에 대한 Traffic Pattern을 회귀식을 통해 구성한다. 회귀식은 다중회귀식으로 구성하며, 특정 시간대에 대한 트래픽 기준 값에 대해 Critical/Major/Minor의 비율로 구성하였다.

3.1.6 탐지 절차

추세 모형을 적용하여 탐지하는 절차는 다음 Fig. 3과 같다.



[Fig. 3] Anomaly Traffic Detection Procedure

4. 실증 분석

4.1 제안된 추세 모형의 검증

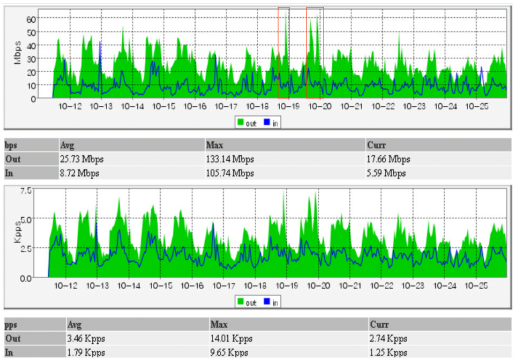
4.1.1 검증 요건

본 연구에서 제안한 추세 모형에 대하여 다른 시계열 분석법인 지수평활법 및 웨이브렛(Wavelet) 기법에 대비하여 우위를 검증하기 위해서는 다음과 같은 항목으로 비교하여 검증하였다[9].

-실시간 탐지 요건
-주기별로 반복되는 변동 요인 반영(계절요인)
-특정일에 대한 처리 기능
-이상트래픽 감지시 상세 분석 기능 제공
-이상트래픽에 의한 모델 왜곡 방지
-자가학습을 통한 모델 정규화
-오탐율(False-Positive) 최소화 방안

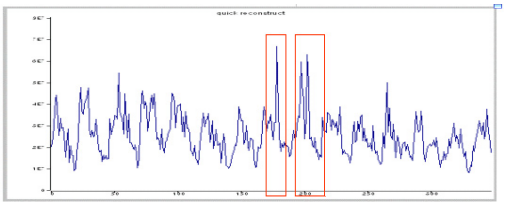
4.1.2 실측 데이터를 이용 추세 모형을 적용한 사례

Fig. 4은 추세 모형 분석에 적용한 실측 데이터를 나타내며 적색 박스로 표시한 부분이 비정상트래픽 부분이다.



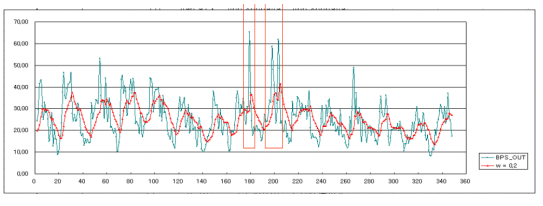
[Fig. 4] The case of trend model analysis using Experimental data

추세 모형 분석에 적용한 실측 데이터를 Wavelet 기법에 적용하여 정상트래픽을 탐지한 사례는 Fig. 5와 같다.



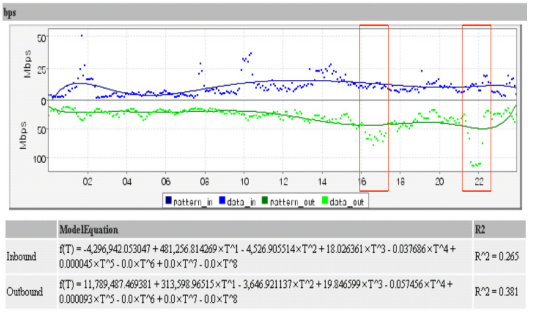
[Fig. 5] The case of Detection using Wavelet Analysis Methods

단순지수평활(w=0.2)법을 적용하여 비정상트래픽을 탐지한 사례는 Fig. 6과 같다.



[Fig. 6] The case of Detection using Exponential Smoothing Methods

Fig. 7은 상기 실측 데이터를 추세 모형에 적용하여 비정상트래픽을 탐지한 사례이다.



[Fig. 7] The case of Detection using trend model Methods

4.2 요건별 비교 분석

4.2.1 실시간 탐지 요건

시스템의 취약점이 알려진 후 취약점을 공격하는 DDOS, Worm 등의 출현이 점차 그 개발 주기가 짧아져 Zero-Day를 향하고 있다. 따라서 비정상트래픽을 탐지하기 위해서는 실시간 탐지가 기본 요건이 된다[16]. 웨이브렛 방법은 적용되는 통계기법 자체가 예측보다는 분석에 있으므로 저주파, 중저주파에서 감지되는 이상트래픽은 최소 2시간이 넘어야 감지가 가능하다. 따라서 현실상 실시간 분석이 불가능 하다. 하지만 지수평활법과 추세 모형법은 실시간 탐지에 웨이브렛 방법보다 효과적이다.

4.2.2 주기별로 반복되는 변동요인 반영

웨이브렛 방법은 주파수 대역별 분석이 이루어지므로 가장 정확한 분석을 수행할 수 있다. 지수평활법은 Holt-Winters 방법에서 계절요인 항목이 있어 지원이 가능하지만 각 요인별 가중계수(α, β, ɣ) 설정에 많은 계산량이 소요된다. 또한 모델 구성 이전에 많은 데이터가 축적되어 있어야 하는 제약점이 있다. 추세 모형은 회귀 분석식만으로는 제공이 불가능한데, 제안한 모델에서는 일별 회귀분석식들에 대해 대표 트래픽 패턴 분석 시에 가중평균을 적용하는바, 이때 계절요인이 적용되도록 하였다.

4.2.3 특정일에 대한 처리 기능

웨이브렛 방법은 주파수 대역의 저주파에서 이에 대한 보정이 이루어지므로 특정일에 대한 처리가 가능한 반면, 지수평활법은 연속적인 시계열 분석에 기반을 둔 모델이므로 이에 대한 처리가 불가능하다. 또한, 이로 이한 데이터 오류가 단시간 내에 상쇄될 수가 없다. 그러나 추세 모

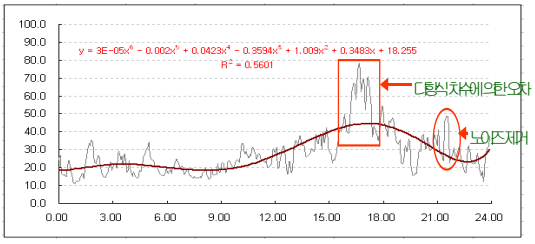
형은 일별로 회귀식을 만들고 수집서버에서 이들에 대한 조합을 수행하기 때문에 관리자가 모델을 자유롭게 관리할 수 있으며, 특정일에 대해 관리자가 미리 설정할 수 있는 기능을 제공함으로써 타 모델에 비해 유리하다.

4.2.4 이상트래픽 감지 시 상세분석 기능 제공

웨이브렛 방법은 실시간 분석이 되지 않으므로, 상세분석을 위해서는 상당한 시간동안 원시데이터 저장이 필요하다. 따라서 실시간 분석을 통한 감지 및 제어 기능을 제공하지 못한다. 지수평활법은 가장 적은 시스템 자원을 이용하여 실시간 분석이 가능하고, 실시간적으로 모델이 보정될 수 있다. 그러나 Port, AS와 같이 인터페이스별 트래픽에 비해 상대적으로 변동 폭이 큰 경우 이 방법을 적용하기가 어렵다. 따라서 회선이용량 분석을 통한 이상트래픽 감지 및 이에 대한 정확한 원인 규명에 있어 새로운 모델이 고안되어야 한다. 추세 모형은 모델 구성에서부터 인터페이스, 주요 Port, 주요 AS별 분석이 이루어지므로, 이상트래픽 원인 규명이 용이하다.

4.2.5 이상트래픽에 의한 모델 왜곡 방지

웨이브렛 방법은 실시간 분석 방법이 아니므로 이에 대한 고려 사항이 아니다. 지수평활법은 이상트래픽이 감지될 경우 이에 대해 데이터 보정을 생략할 수는 있지만 본질적으로 그 단점을 보완하지는 못한다. 또한 이상트래픽의 끝점을 분석할 수 없기 때문에 모델 왜곡을 미연에 방지할 수 없다. 추세 모형은 회귀식 자체가 최소자승법을 이용하기 때문에 회귀식 구성에서 부터 이상트래픽에 대한 데이터 왜곡을 방지할 수 있다.



[Fig. 8] Preventing distortion due to abnormal traffic

4.2.6 자기학습을 통한 모델 정규화

지수평활법은 모델 구성에 있어서부터 기존 자료를 이용하여 가중계수를 결정해야 하기 때문에 자가 학습이 어려우며, 모델 운용 도중에 가중계수를 변경할 경우 기존 데이터와의 호환성에 문제가 발생한다. 추세 모형은 일별로 회귀식이 구성되고, 최종적으로 사용할 때 기존 모델의 가중평균을 구하기 때문에 자가학습이 가능하다. 또한, 최초 모델 구성 시 관리자의 관리가 필요 없고, 시간이 지나면 지날수록 모델이 정규화 되는 장점을 지닌다.

4.2.7 오탐율(False-Positive) 최소화 방안

웨이브렛 방법은 여러 주파수 대역에 대한 검토가 되기 때문에 오탐 확률이 가장 낮다. 그러나 여러 가지 트래픽 형태에 대한 규칙을 새워야 하기 때문에 고도의 관리자 이해력을 요구한다.

지수평활법은 실시간적으로 추정 값과 실제 값을 비교해야 하며, 이전 데이터 누락 시 다음 데이터 추정이 불가능하여 지속시간에 대한 고려를 이용할 수 없다. 회선별 임계값 설정이 가장 중요하다. 추세 모형은 관리자가 설정한 임계값 이외에, 이상트래픽의 지속시간에 대한 고려를 할 수 있기 때문에 일시적인 트래픽 증가에 대한 감지 여부를 관리자가 설정할 수 있다[14].

4.2.7 웨이브렛, 지수평활법, 추세 모형 비교

비정상 네트워크 트래픽 탐지를 위한 통계적 분석 방법인 시계열 모형의 웨이브렛, 지수평활법, 추세 모형 예측 모델을 실증 분석한 결과 7개 평가 항목 중 5개 항목에서 추세 모형 분석 기법이 가장 높은 점수를 받아 적합한 것으로 검증 되었다. 추세 모형은 실시간 탐지 뿐만 아니라 특정일 트래픽 처리, 이상트래픽 감지시 상세 분석, 이상 트래픽에 대한 모델 왜곡 방지, 자가학습을 통한 모델 정규화 등 많은 부분에서 우수함이 검증되었다.

[Table 2] Wavelet, Exponential Smoothing, Trend Model Comparison

Comparison item	Wavelet	Exponential Smoothing	Trend Model
real time detection	low	high	high
Factor per cycle reflects changes in repetitive	high	high	middle
Traffic on a particular day can be processed	middle	low	high
Provides detailed analysis of abnormal traffic detection	high	middle	high
Distortion caused by abnormal traffic protection model	-	low	high
Normalized through self-learning model	-	low	high
Minimize false positives	middle	low	middle

Table 2은 웨이브렛, 지수평활법, 추세 모형에 대한 비교를 한 것이다.

5. 결론

네트워크 침해 또는 악성 트래픽으로 불리는 DDoS나 Worm과 같은 비정상 트래픽을 찾아내고 막는 것은 물론 네트워크 자원의 효율적 활용에 악영향을 끼치는 비정상 트래픽에 대한 정확한 탐지가 필요하다. 이에 따라 본 연구에서는 비정상 트래픽을 탐지하는 방법인 지수평활법, Wavelet 분석법, 추세 모형 등에 대하여 분석하여 가장 적합한 모델인 추세 모형 기반의 예측 모델을 제안하였다. 통계학적 추세 모형 기반의 예측 모델 실시간 탐지기법을 제시하고, 이에 대한 S/W를 구현・검증하였다. 지수평활법은 많은 자료의 저장이 필요 없고 계산이 쉽다는 장점이 있으나, 평활상수 설정에 따라 상이한 측정치를 갖는다는 것이 매우 큰 단점으로 분석되었다. Wavelet 분석기법은 적용되는 통계기법 자체가 예측보다는 분석에 있으며, 저주파・중주파에서 감지되는 이상트래픽은 최소 2시간이 넘어야 감지가능 하므로 현실상 실시간 분석에는 부적합한 기법으로 분석되었다. 추세모형은 실시간 탐지 뿐만 아니라 특정일 트래픽 처리, 이상트래픽 감지시 상세 분석, 이상 트래픽에 대한 모델 왜곡 방지, 자가학습을 통한 모델 정규화 등 많은 부분에서 우수함이 검증되었다.

제안하는 추세 모형기반의 예측 모델을 이용한 비정상 트래픽 탐지 기법의 특징은 첫째, 백본 네트워크와 같이 네트워크의 크기가 클수록 트래픽의 일주기성이 뚜렷하게 나타나므로 고속 대용량의 네트워크에서 실시간으로 트래픽탐지가 가능하다. 둘째, 변화하는 네트워크 트래픽 특성의 추세를 반영할 수 있으므로 초기의 학습기간 이후에 관리가 매우 용이하다. 셋째, 인터넷 백본 네트워크는 물론 트래픽의 일주기성 경향을 띠고 있는 모든 네트워크에서 다차회귀방정식을 통해 네트워크 관리자의 입장에서 적절한 트래픽 유형을 정의하여 탐지가 가능하다. 넷째, 많은 수의 인터페이스를 분석할 수 있으며, 대략의 추세를 모델링하기 때문에 DDoS나 웜과 같이 갑자기 증가하는 비정상 트래픽 탐지에 적합하다.

향후 연구되어야 할 과제로는 현재 비정상 트래픽 탐지와 검증을 실측 데이타를 이용하여 추세 모형에 대한 검증과 또한 실시간으로 유해 트래픽 탐지를 위해 네트워크 트래픽 데이터의 수집과 분석을 실시간 처리 하기 위한 해결방안에 대해 연구가 이루어져야 할 것이다. 또한 비정상트래픽 제어에 있어서도 네트워크 관리자의 판단에 의해서 이루어지는 것에서 비정상 트래픽 탐지 시스템에서의 능동적 제어에 대한 연구가 이루어져야 할 것이다.

References

[1] P. Barford and D. Plonka, "Characteristics of Network Traffic Flow Anomalies", In Proceedings of the ACM SIGCOMM Internet Measurement Workshop, 2001. DOI: http://dx.doi.org/10.1145/505202.505211

[2] Seung-pyo Hong, "A study on abnormal network traffic detection method using fisher linear discriminant". korea university the graduate school of engineering, 2013.

[3] Min seok Lee, "Research of HTTP GET flooding attack detection methods using ARIMA time series forecasting model". yonsei university the graduate school of engineering, 2013.

[4] Myeung hee Hoe, Yoo sung Park, 'Time Series Analysis', Liberty Academy, 1994.

[5] Sung-Min Park, "Comparisions of forecasting Methods Using Time Series Model". Graduate School, Chungbuk National University, 2013.

[6] Shin sup Cho, Young sook Son, 'A Time Series Analysis', Yulgok Publishing Company, 1999.

[7] B. Abraham, and J. Ledolter, "Statistical Methods for Forecasting", Wiley, 1983. DOI: http://dx.doi.org/10.1002/9780470316610

[8] P. Brockwell and R. Davis, "Introduction to Time Series and Forecasting, Springer, 1996.

[9] J. Lewalle, "Tutorial on Continuous Wavelet Analysis of Experimental Data", Syracuse University, 1995.

장 상 수(Sang-Soo Jang) **[정회원]**

- 1989년 2월 : 한국항공대학교 항공통신정보공학과 (학사)
- 2003년 2월 : 동국대학교 정보보호학과 (공학석사)
- 2011년 8월 : 조선대학교 정보보호학과 (정보보호학박사)
- 2000년 5월 ~ 현재 : 한국인터넷진흥원 수석연구위원

<관심분야>
ISMS/PIMS, 위험관리, 정보보호 성과 측정 등