NSRI

WISC2006

제 18 회 정보보호와 암호에 관한 학술대회

The 18th Workshop on Information Security and Cryptography

논 문 집

일 시 : 2006. 9. 7. (목) ~ 9. 8. (금)

장 소 : 천 안 상 록 리 조 트

주 최 : 한국전자통신연구원 부설 국가보안기술연구소

후 원 : 국가정보원
국방부
정보통신부

제 99 호

표 창 장

우 수 상

논문제목 : 스캐닝 웜의 트래픽 특성을 이용한
탐지 방법

성 명 : 김재현(아주대학교)
강신헌(아주대학교)

위 사람은 국가 정보보호 기술 발전을 위하여 개최하는 제 18 회 정보보호와 암호에 관한 학술대회에 두서와 같이 우수한 논문을 투고하여 본 학술대회 발전에 공헌한 바가 크므로 이에 표창장을 수여함.

2006 년 9 월 7 일

국가보안기술연구소장 박 춘 식

Session C1 : 침입탐지

C1.1 로그 연관 분석을 통한 웹 기반 공격 비정상 행위 탐지 ····· 53
장문수, 강정민, 김장하, 손기욱 (국가보안기술연구소)

C1.2 내부자 위협에 대한 계층적 침입탐지 모델에 관한 연구 ····· 63
엄정호*, 한영주**, 정태명** (* 공군, ** 성균관대학교)

C1.3 스캐닝 웜의 트래픽 특성을 이용한 탐지 방법 ····· 71
김재현, 강신헌 (아주대학교)

C1.4 데이터마이닝을 이용한 시스템 호출별 빈도 기준의 호스트 기반 침입탐지 모델 생성 체계 ····· 83
백승현, 오윤근, 오형근, 이도훈 (국가보안기술연구소)

Session A2 : 비공개 II

A2.1 음성용 암호장비의 누설스펙트럼 분석 ····· 101
김종규, 황인호 (국가보안기술연구소)

A2.2 순시주파수를 이용한 미지 주파수 도약 신호 검출 및 분석 ····· 102
이경훈, 주세훈, 양광모, 황인호 (국가보안기술연구소)

A2.3 높은 BER 환경 하에서 도래시간차 추정 방법 ····· 103
조상우, 왕진천, 황인호 (국가보안기술연구소)

Session B2 : 센서네트워크

B2.1 전자서명 유효성 연장을 위한 효율적인 인증서 상태검증 방법에 관한 연구 ··· 107
심재훈*, 강유경**, 김영진* (* 드림시큐리티, ** 한국무역정보통신)

B2.2 소규모 센서 네트워크를 위한 확률론적 그룹 키 관리기법 ····· 124
이영철, 이수진, 남길현 (국방대학교)

B2.3 무선 센서 망에서의 다층 구조 그리드를 이용한 위치정보 기반 키 관리 기법 ····· 139
이종협*, 권태경**, 송주석* (*연세대학교, ** 세종대학교)

스캐닝 웜의 트래픽 특성을 이용한 탐지 방법

김 재 현, 강 신 헌

Detection Algorithm of Scanning worms using network traffic characteristics

Jae-Hyun Kim, Shin-Hun Kang

요 약

스캐닝 웜은 네트워크 관리자가 미처 대응하기 전에 넓게 전파되므로 차단하기가 힘들고 그 피해가 상당히 크다. 따라서 자동으로 스캐닝 웜의 발생을 탐지하고 이에 대응할 수 있는 방법이 필요하다. 본 논문에서는 스캐닝 웜의 트래픽 특성을 분석하여 정상 트래픽과 이상 트래픽을 구분할 수 있는 탐지 알고리즘을 제안한다. 스캐닝 웜의 탐지를 위해 variance, VMR 및 correlation coefficient를 이용하는 방법을 제안하고, 시뮬레이션을 통해 기존의 방법과 성능을 비교하였다. 그 결과 기존의 방법에 비하여 간단한 계산을 통해 스캐닝 웜의 정확한 탐지가 가능함을 확인하였다.

Ⅰ. 서 론

최근 컴퓨터와 인터넷 기술의 급속한 발전으로 인터넷 사용 인구는 꾸준히 증가하고 있으며 네트워크는 고속화되어 가고 있다. 이에 따라 고속 네트워크를 통하여 멀티미디어 서비스를 비롯한 다양한 서비스가 가능하게 된 반면에, 네트워크를 통한 악의적인 공격에 의한 피해 또한 과거에 비해 많이 증가하였다. 2003년 1월 25일, 우리나라를 비롯한 전 세계의 네트워크를 한동안 마비시켜 인터넷 대란을 일으켰던 슬래머 웜[1]의 사례만 보아도 이러한 공격에 의한 피해가 심각하다는 것을 알 수 있다.

슬래머 웜과 같이 특정한 공격 대상 없이 임의의 호스트를 향해 계속 전파되는 웜을 스캐닝 웜이라고 하는데, 스캐닝 웜은 자기 스스로 복제가 가능하며 짧은 시간에 네트워크를 통해 아주 넓은 범위에 걸쳐서 전파되므로 그 피해가 크다. 특히 스캐닝 웜은 관리자가 미처 대응하기 전에 빠른 속도로 네트워크에 전파되므로 자동으로 스캐닝 웜을 탐지하고 대응할 수 있는 시스템이 필요하다. 이를 위해 많은 연구가 진행되고 있지만 대부분의 연구가 패킷 헤더 정보를 이용하는 방법에 중점을 두고 있으며 스캐닝 웜의 일반적인 트래픽 특성을 이용한 탐지에 관한 연구는 미비한 실정이다. 패킷 헤더 정보를 이용하는 방법은 트래픽 특성을 이용한 탐지 방법에 비하여 더 정확한 탐지가 가능하지만 탐지를 위해서 모든 패킷을 다 검사해야 하므로 탐지에 소요되는 시간이 길다. 반면에 트래픽 특성을 이용한 탐지 방법은 패킷 헤더 정보를 이용하는 방법에 비하여 정확도가 떨어지지만 모든 패킷을 다 검사할 필요가 없으므로 빠른 시간에 효율적인 탐지가 가능하다. 따라서 본 논문에서는 스캐닝 웜의 패킷 헤더 정보에 중점을 둔 기존의 탐지 방법대신 트래픽 특성을 이용하여 탐지하는 방법을 제안한다. 그리고 시뮬레이션을 통해 기존의 다른 탐지방법과 성능을 비교 분석하여 제안하는 탐지 알고리즘의 타당성을 검증한다.

Ⅱ. 관련 연구

웜을 탐지하기 위한 방법은 크게 패킷의 헤더 정보를 이용하는 방법과 전체 트래픽의 양을 이용하는 방법으로 나눌 수 있다. 일반적으로 스캐닝 웜은 가능한 한 많은 호스트를 감염시키기 위해 수

신 IP 주소가 랜덤하며, 특정 서비스의 취약점을 공격하기 때문에 수신 포트 번호가 고정되어 있다. 따라서 발신 IP 주소, 발신 포트 번호, 수신 IP 주소, 수신 포트 번호 및 프로토콜 등 패킷 헤더 정보를 이용하여 스캐닝 웜을 탐지할 수 있다. 기존의 연구 중 패킷 헤더 정보를 이용하는 방법을 살펴보면 패킷의 발신 IP 주소, 수신 IP 주소 및 수신 포트 번호를 3차원 그래프로 나타내 특정한 형태를 띠게 되면 스캐닝 공격이라고 판단하는 방법[2]이 있으며, IP 주소, 포트 번호에 대한 엔트로피를 측정하여 수신 IP 주소가 랜덤하게 분포되어 있거나 특정 포트에 트래픽이 집중되어 있는 것을 탐지하는 방법[3]이 있다. 그밖에 IP 주소, 포트 번호가 같은 패킷들을 플로우 단위로 묶어 플로우 헤더 정보와 트래픽 패턴 정보를 생성하여 공격 유형별로 정의된 기준에 따라 공격을 탐지하는 방법[4]과 각 네트워크 경계마다 설치된 스캔 모니터에서 수집한 데이터를 종합하여 칼만 필터를 통해 웜의 감염율을 추정하는 방법[5] 등이 있다. 패킷 헤더 정보를 이용하여 웜을 탐지하는 방법은 여러 가지 데이터를 이용하여 판단하기 때문에 정확도가 높지만, 모든 패킷의 헤더를 검사해야 하기 때문에 시스템이 복잡하고 탐지 시간이 오래 걸린다. 상대적으로, 패킷 헤더 정보를 이용하지 않고 단위 시간당 패킷의 수와 비트 수 정보만을 이용해 트래픽량에 대한 패킷 수의 비율을 구하여 웜을 탐지하는 방법[6]은 모든 패킷을 검사할 필요가 없어 빠르고 효율적인 탐지가 가능하다. 그러나 제한된 정보만을 이용하여 웜을 탐지하기 때문에 패킷 헤더 정보를 이용하여 탐지하는 방법에 비해 정확도가 떨어진다. 따라서 본 논문에서는 네트워크의 트래픽 정보를 이용하여 높은 정확도로 효율적인 탐지가 가능한 스캐닝 웜 탐지 방법을 제안한다.

Ⅲ. 스캐닝 웜의 트래픽 특성

스캐닝 웜이란 감염시킬 대상을 찾기 위해 순차적이거나 랜덤한 주소로 스스로 연결을 시도하는 웜으로서, 여러 종류의 웜 중에서 가장 심각한 피해를 일으킨다. 기존 스캐닝 웜에 의한 피해 사례를 살펴보면, 2001년 7월에 발생하여 14시간 만에 약 359,000 대의 컴퓨터를 감염시켰던 코드레드 웜[7], 2003년 1월에 발생하여 10분만에 약 75,000 대의 컴퓨터를 감염시켰던 슬래머 웜[1], 2003년 8월에 발생하여 수 시간 만에 500,000 대의 컴퓨터를 감염시켰던 블래스터 웜[8], 그리고 2004년 3월에 발생하여 45분만에 12,000 대의 컴퓨터를 감염시켰던 위티 웜[9] 등이 있다.

스캐닝 웜을 탐지하기 위해 그 특성을 분석하여 (표 1)에 정리하였다. 스캐닝 웜은 수신 포트 번호 또는 발신 포트 번호가 고정되어 있으며 수신 IP 주소는 랜덤하거나 부분적으로 랜덤한 특성을 보인다. 스캐닝 웜은 코드레드 웜, 블래스터 웜과 같이 패킷 크기가 비교적 크고 스캔 속도가 느린 것과 슬래머 웜처럼 패킷 크기가 작고 스캔 속도가 빠른 것이 있다. 따라서 패킷 크기와 스캔 속도를 개별적으로 고려해서는 스캐닝 웜의 특성을 파악할 수가 없다. 하지만 스캐닝 웜의 패킷 크기와 스캔 속도를 곱한 트래픽량은 항상 크다는 것을 알 수 있다. 따라서 본 논문에서는 트래픽량의 분석을 통해 스캐닝 웜의 트래픽 특성을 이용한 탐지 방법을 제안한다.

(표 1) 스캐닝 웜의 특성 분석

구 분	수신 IP 주소	수신 포트 번호	크기 (byte)	스캔 속도 (packets/sec)
코드레드 I v2	랜덤	80	3569	11
코드레드 II	12.5% : 랜덤 50% : 같은 클래스 B 네트워크 37.5% : 같은 클래스 C 네트워크	80	3818	.
슬래머	랜덤	1434	404	4000
블래스터	40% : 같은 클래스 C 네트워크 60% : 랜덤 (순차적 스캔)	135	6176	15
위티	랜덤	랜덤 (발신 포트 번호 : 4000)	796~1307	357

Ⅳ. 제안하는 트래픽 특성 탐지 알고리즘

본 논문에서는 Ⅲ장에서 살펴본 스캐닝 웜의 트래픽 특성을 분석하여 정상 트래픽과 구분되는 특성을 통해 스캐닝 웜을 탐지한다. (표 1)에서 알 수 있듯이 코드레드 웜은 약 4kbyte 크기의 패킷을 초당 11개씩 발생시킨다. 패킷의 크기가 크고, 패킷 간의 간격이 약 0.1초 정도로 비교적 길기 때문에 트래픽의 변화량이 크다. 따라서 트래픽의 변화량을 나타내는 variance의 값이 커진다. 반대로 404byte 크기의 패킷을 초당 4000개씩 발생시키는 슬래머 웜은 패킷의 크기가 작지만 초당 패킷 수가 많기 때문에 트래픽의 변화량이 커져 variance가 증가한다. 시뮬레이션을 통해 코드레드 웜, 슬래머 웜과 유사한 트래픽을 발생시킨 후 분석한 결과 이상 트래픽의 variance가 정상 트래픽에 비하여 커지는 것을 확인하였다. 본 논문에서는 이러한 시간에 따른 트래픽량의 variance를 측정하여 스캐닝 웜을 탐지하는 방법과 variance를 평균값으로 나눈 Variance to Mean Ratio(VMR)을 이용하는 방법을 제안한다. 그리고 두 구간 사이의 상관관계를 나타내는 correlation coefficient를 이용하는 방법을 제안한다.

4.1. Variance를 이용하는 방법

트래픽의 특성을 분석하기 위한 가장 간단한 방법으로 먼저 variance를 이용하는 방법을 고려하였다. $\mu = E(X)$가 랜덤 변수 X의 평균값이라고 할 때, X의 variance는 다음과 같이 정의된다.

$$var(X) = E[(X-\mu)^2]. \tag{1}$$

스캐닝 웜이 발생하면 네트워크의 트래픽량이 급증하므로 정상 트래픽에 비해 단위시간당 트래픽량의 variance가 증가한다. variance를 계산하기 위해 네트워크 링크에서 초당 트래픽량(bits/sec)을 측정한다. $X(t)$는 (t-1)~t초까지의 시간 동안 링크를 통과한 비트 수, W는 몇 초간의 variance를 계산할 것인지 정하는 윈도우 크기라고 했을 때 다음과 같은 방식으로 n초에 측정한 variance $v(n)$을 구할 수 있다.

$$v(n) = \frac{\sum_{k=n-W+1}^{n} [X(k)-\mu(n)]^2}{W}, \tag{2}$$

이 때,

$$\mu(n) = \frac{\sum_{k=n-W+1}^{n} X(k)}{W}$$

$n = 0, 1, 2, 3 \cdots$

$k \geq 0$.

예를 들어서, $W = 5, n = 5, 6$일 때의 variance를 구하는 과정은 다음과 같다.

$$v(5) = \frac{[X(1)-\mu(5)]^2 + [X(2)-\mu(5)]^2 + [X(3)-\mu(5)]^2 + [X(4)-\mu(5)]^2 + [X(5)-\mu(5)]^2}{5}$$

$$v(6) = \frac{[X(2)-\mu(6)]^2 + [X(3)-\mu(6)]^2 + [X(4)-\mu(6)]^2 + [X(5)-\mu(6)]^2 + [X(6)-\mu(6)]^2}{5}.$$

4.2. Variance to Mean Ratio(VMR)를 이용하는 방법

Variance는 단순히 트래픽의 변화량을 나타내므로 정상 트래픽도 경우에 따라 variance가 크게 나타날 수 있다. 또한 variance는 트래픽의 전체 크기에 따라 같은 양의 변화에 대해서 다른 값을 갖게 되므로 정확한 탐지기준으로 삼기에는 부족하다. 따라서 variance를 평균값으로 나눈 값인 VMR을 이용한다면 시간에 따른 전체 트래픽량의 크기 변화에 상관없이 스캐닝 웜의 탐지가 가능하다. $\sigma^2 = var(X)$가 랜덤 변수 X의 variance라고 할 때, X의 VMR은 다음과 같이 정의된다.

$$VMR = \frac{\sigma^2}{\mu}. \tag{3}$$

VMR은 다음과 같은 방식으로 측정할 수 있다.

$$VMR(n) = \frac{v(n)}{\mu(n)}. \tag{4}$$

4.3. Correlation Coefficient를 이용하는 방법

정상적인 트래픽도 변화량이 커질 수 있으므로 단순히 트래픽의 변화량만 측정해서는 스캐닝 웜의 정확한 탐지가 어렵다. 따라서 두 구간 X와 Y 사이의 상관관계를 알 수 있는 correlation coefficient를 이용하는 방법을 고려하였다. correlation coefficient는 -1과 1사이의 값을 가지며, X와 Y가 independent하면 0의 값을 갖고, X가 증가할 때 Y가 linear하게 증가하면 1, linear하게 감소하면 -1의 값을 갖는다. correlation coefficient가 일정시간 이상 계속 1에 가까운 큰 값을 보이면 네트워크의 트래픽량이 계속해서 늘어나고 있다고 볼 수 있으며, 이는 이상 트래픽의 징후라고 생각할 수 있다. σ_X, σ_Y가 각각 랜덤 변수 X, Y의 표준편차이며, $COV(X, Y)$는 X, Y의 covariance라고 할 때, X와 Y의 correlation coefficient $\rho_{X,Y}$는 다음과 같이 정의된다.

$$\rho_{X,Y} = \frac{COV(X, Y)}{\sigma_X \sigma_Y}, \tag{5}$$

이 때, $COV(X, Y) = E[(X-\mu_X)(Y-\mu_Y)]$.

따라서 correlation coefficient는 다음과 같이 측정한다.

$$\rho_{X,Y} = \frac{E\left[(X-\mu_X)(Y-\mu_Y)^T\right]}{\sqrt{E(X^2)-E^2(X)}\ \sqrt{E(Y^2)-E^2(Y)}}, \tag{6}$$

이 때,

A^T는 행렬 A의 전치행렬을 의미하며

$X = [X(n-1.5W+1)\ X(n-1.5W+2) \cdots\ X(n-0.5W)]$

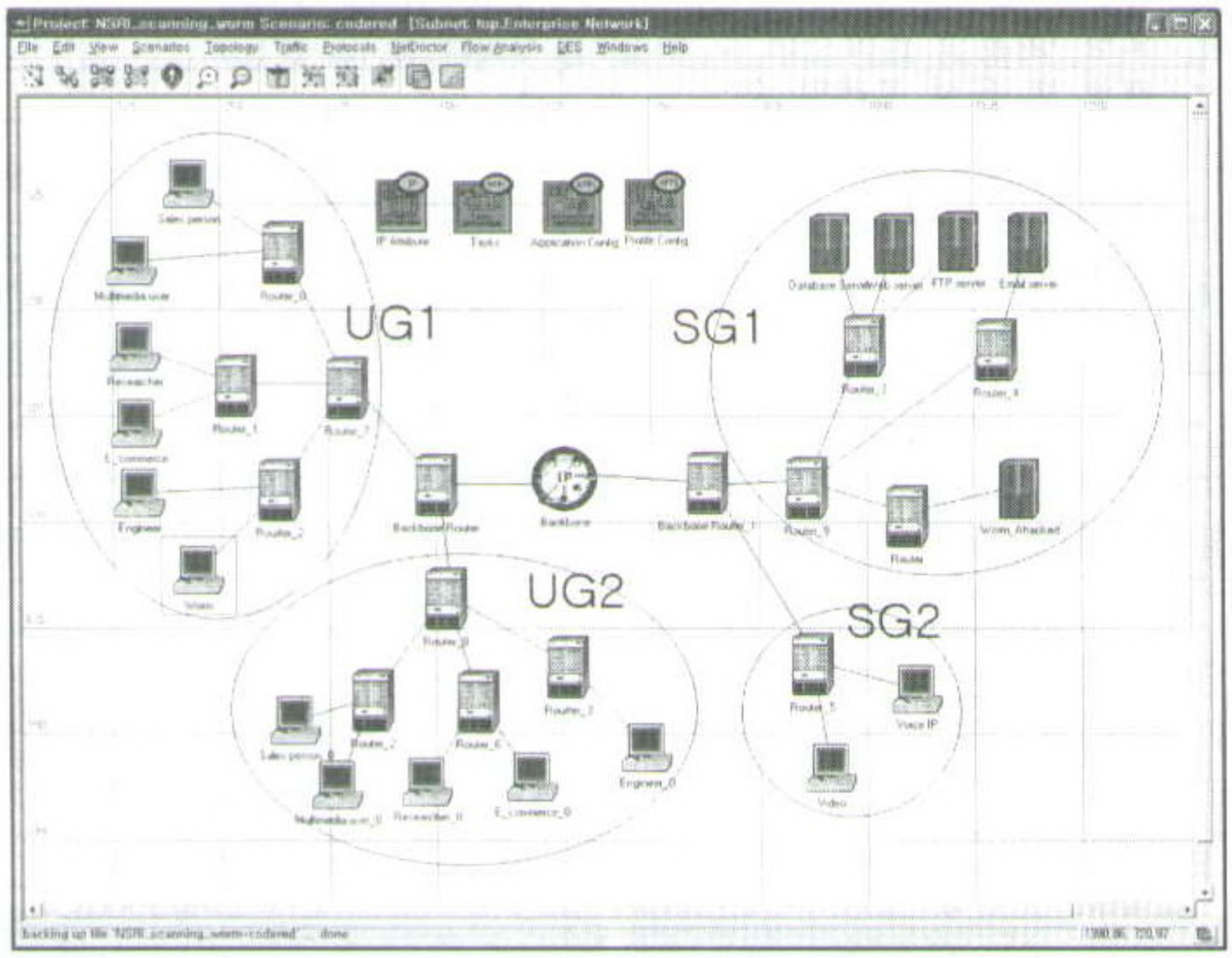
$Y = [X(n-W+1)\ X(n-W+2) \cdots\ X(n)]$,

여기서 W는 계산의 편의상 짝수이다.

V. 성능 평가

본 논문에서는 제안한 알고리즘의 성능을 평가하기 위해 OPNET을 이용하여 정상 트래픽과 이상 트래픽을 발생시키는 시뮬레이터를 제작하였다. 이상 트래픽으로는 코드레드 웜과 슬래머 웜을 고려한다. 그 이유는 스캐닝 웜 중 가장 상반된 특성을 가진 두 웜을 탐지할 수 있다면 다른 웜의 탐지 역시 가능할 것이기 때문이다. 따라서 스캐닝 웜 중 가장 패킷 크기가 크고 스캔 속도가 낮은 코드레드 웜[7]과 가장 패킷 크기가 작고 스캔 속도가 높은 슬래머 웜[1]을 발생시켜 트래픽 특성을 분석하였다.

본 연구에서 고려한 참조망 구조는 (그림 1)과 같다. 전체적인 구조를 살펴보면 일반적인 사용자 5~6명이 모여 있는 사용자 그룹 UG1, UG2와 해당 서비스를 제공하는 서버 그룹 SG1, SG2로 구성되어 있다. 시뮬레이션 시작 후 100초부터 웹 브라우징, 이메일 및 FTP 서비스 등 정상적인 트래픽을 발생시키고 300초부터 순차적으로 코드레드 웜, 슬래머 웜과 같은 특성을 갖는 트래픽을 발생시켰을 때 네트워크 백본망에서 트래픽량을 측정하여 Ⅳ장에서 제안한 방법으로 분석하였다.



(그림 1) 시뮬레이션 참조망 구조

5.1. 환경 설정

주로 사용하는 프로그램에 따라 사용자 별 트래픽 특성이 다르므로 실제 네트워크 트래픽과 유사한 트래픽을 발생시키기 위해 트래픽 유형에 따라 각 사용자를 Sales Person, Multimedia User, E-commerce User, Researcher, Engineer 및 Worm 이렇게 6가지 유형으로 나누었다. 각 사용자 별 트래픽 유형은 (표 2)와 같다.

(표 2) 사용자 별 트래픽 유형

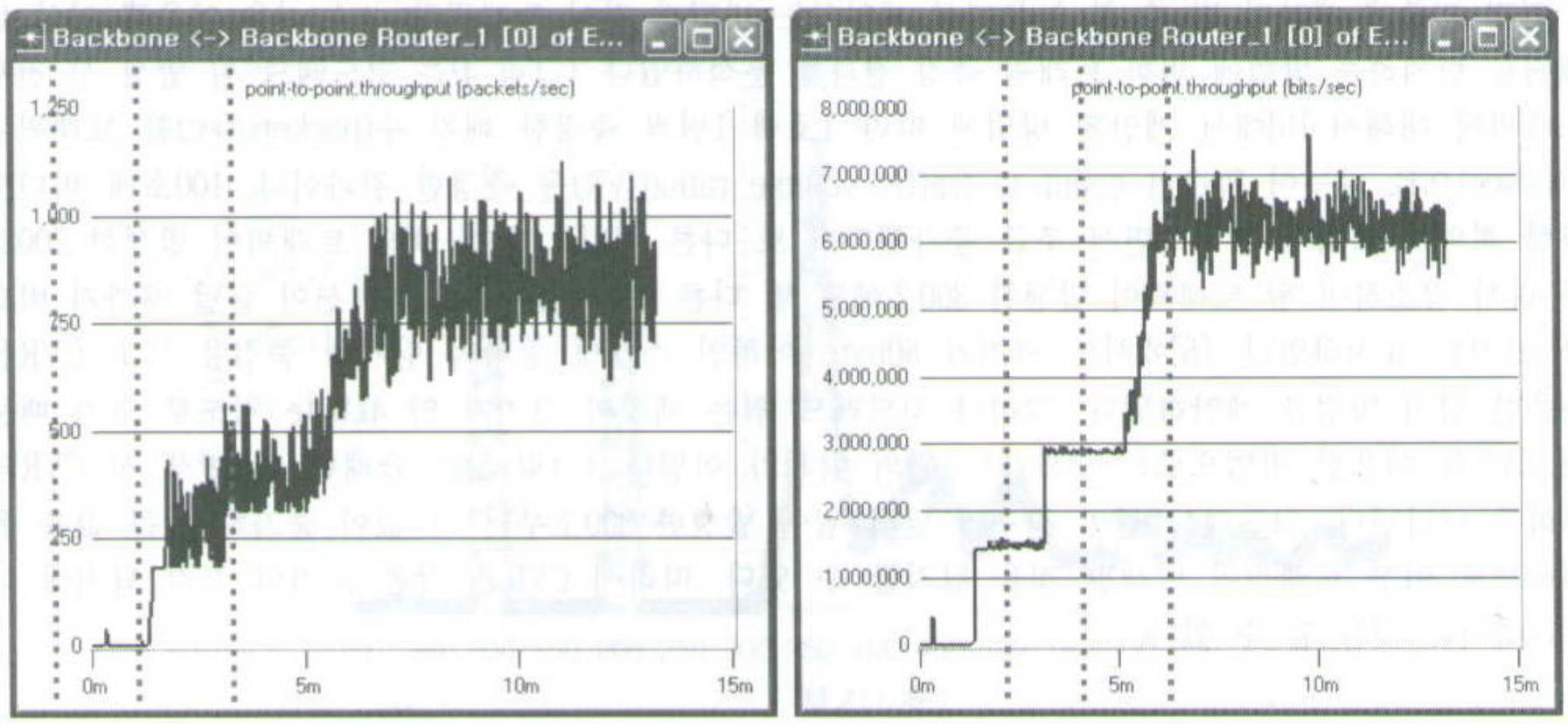
구 분	트 래 픽 유 형
Sales person	데이터베이스 접근, E-mail, 웹 브라우징
Multimedia user	비디오 화상 회의, 데이터베이스 접근
E-commerce user	웹 브라우징, VoIP, E-mail
Researcher	웹 브라우징, E-mail, FTP
Engineer	웹 브라우징, E-mail
Worm	코드레드 웜 I v2 또는 슬래머 웜

트래픽 유형별 세부 설정사항은 (표 3)에 정리하였다.

(표 3) 트래픽 유형 별 설정

구 분			설 정 값	
VoIP	무음 구간 (초)		Exponential (0.65)	
	유음 구간 (초)		Exponential (0.352)	
	코딩 방식		G.711	
	패킷 별 음성 프레임 수		1	
	시작 시간 (초)		Exponential (100)	
비디오 화상회의	초당 프레임 수		10 frames/sec	
	프레임 크기		Uniform (min : 20380, max : 20480) Uniform (min : 17800, max 17920)	
	시작 시간 (초)		Exponential (100)	
FTP	요청 시간 간격 (초)		Exponential (3600)	
	파일 크기 (bytes)		Constant(1000)	
	시작 시간 (초)		Exponential (100)	
웹 브라우징	HTTP 버전		HTTP 1.1	
	페이지 도착 시간 간격 (초)		Exponential (720)	
	페이지 속성		개체 크기 (bytes)	개체 개수
			Constant(500)	Constant(1)
			Uniform(10,400)	Constant(5)
	시작 시간 (초)		Exponential (100)	
웜	코드레드	요청 시간 간격 (초)	Exponential (0.1)	
		요청 패킷 크기 (bytes)	Constant(4000)	
		요청 개수	Constant (500000)	
		시작 시간 (초)	Constant (300, 310, 318, 325, 331, 336, 340, 343, 345, 346)	
	슬래머	요청 시간 간격 (초)	Exponential (0.00025)	
		요청 패킷 크기 (bytes)	Constant(404)	
		요청 개수	Constant (1000000)	
		시작 시간 (초)	Constant (300, 320, 335, 347, 357, 365, 372, 378, 383, 387, 390)	

(표 3)에서 VoIP, 비디오 화상회의, FTP, 웹 브라우징 및 데이터베이스 접근, E-mail 트래픽 모델은 OPNET에서 제공하는 표준 모델을 참조하여 설정하였고, 슬래머 웜과 코드레드 웜 트래픽은 각각 [1]과 [7]을 참조하여 설정하였다.



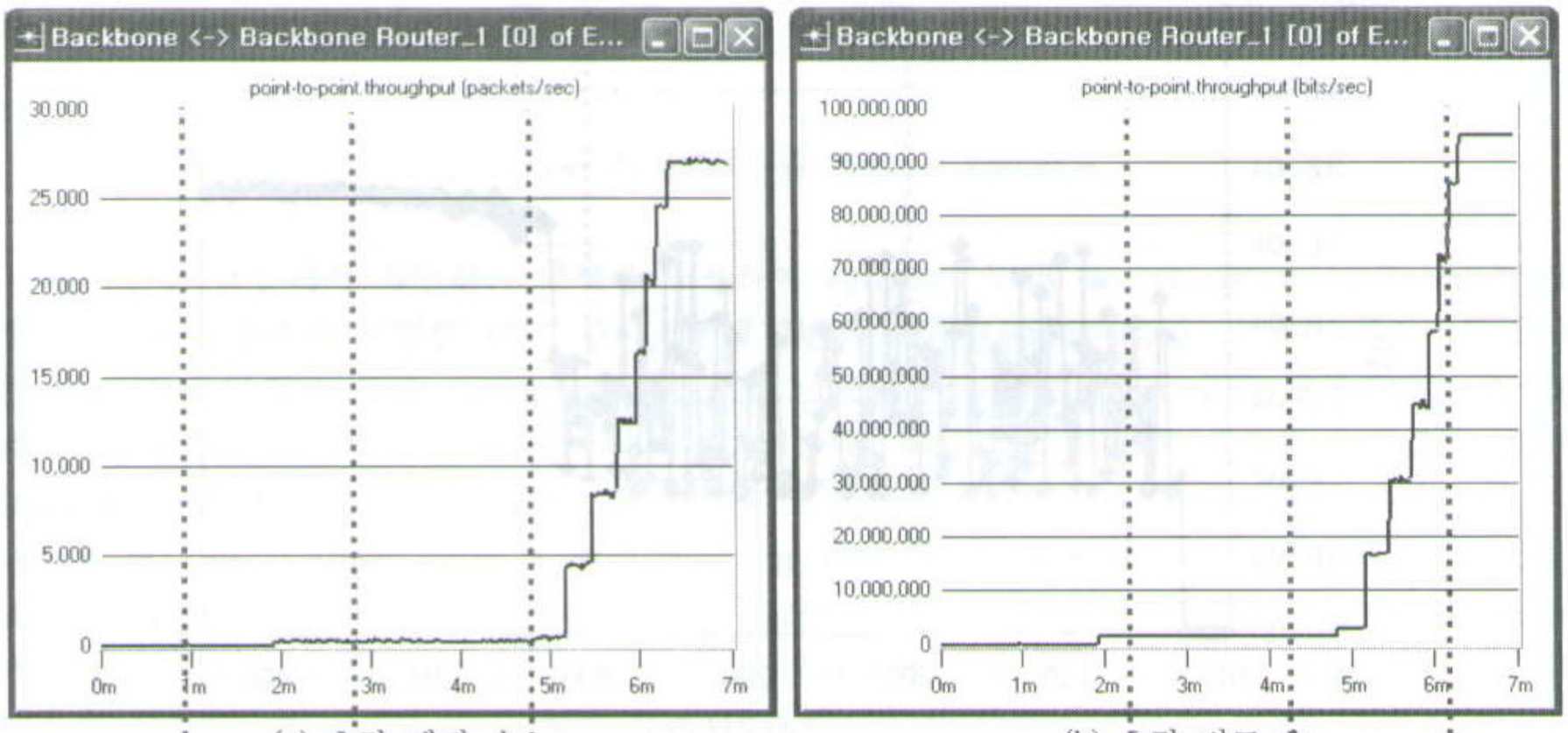
(a) 초당 패킷 수 (b) 초당 비트 수

(그림 2) 코드레드 웜 발생 시 네트워크 트래픽

5.2. 트래픽 측정 및 분석

(그림 2)는 정상 트래픽과 함께 코드레드 웜 트래픽을 발생시켰을 때 백본 링크에서 초당 패킷 수(packets/sec)와 초당 비트 수(bits/sec)를 측정한 결과를 나타낸다. 100초와 200초에 각각 발생시킨 비디오 화상 회의 트래픽에 의해 전체 트래픽량이 급증하였고(첫 번째, 두 번째 점선), , 300초부터 발생시킨 웜 트래픽에 의해 전체 트래픽량이 다시 급증하는 것을 볼 수 있다(세 번째 점선).

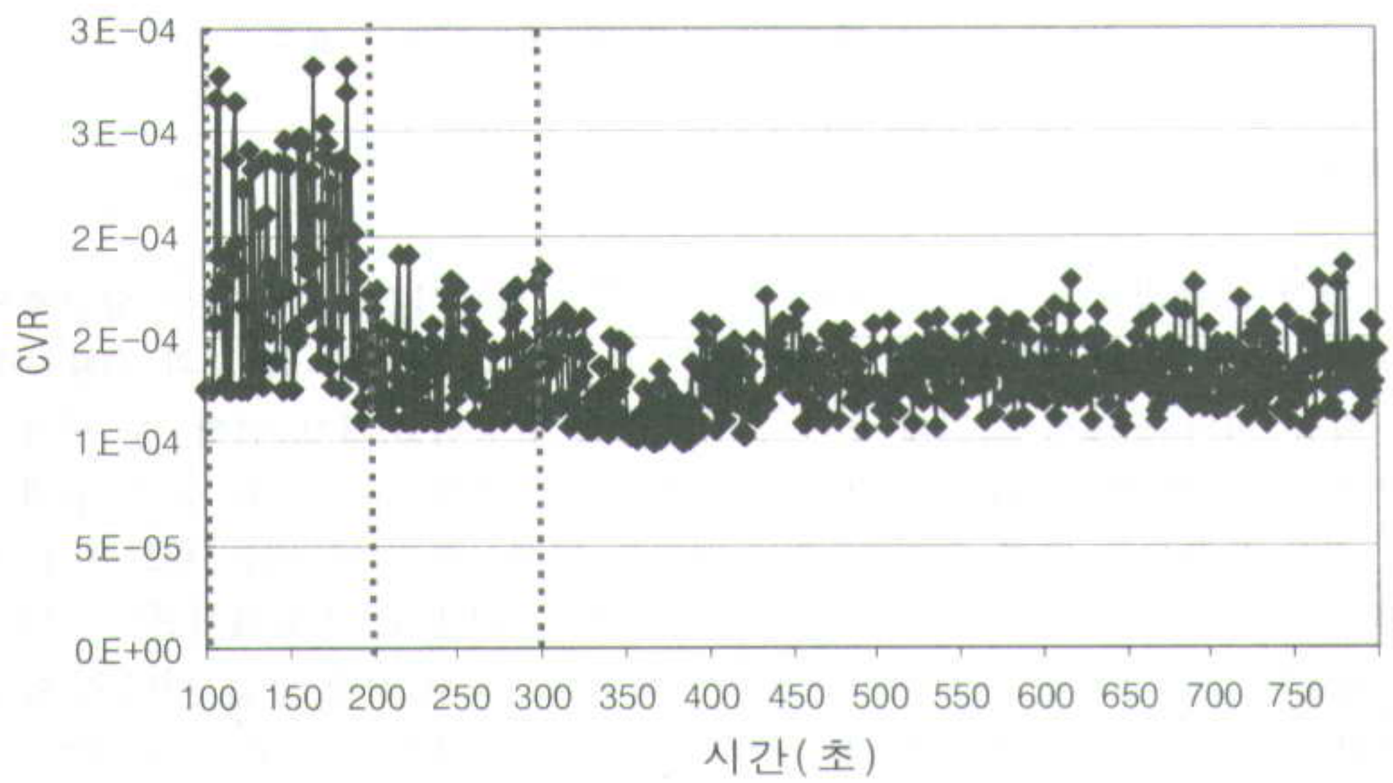
(그림 3)은 정상 트래픽과 함께 슬래머 웜 트래픽을 발생시켰을 때 백본 링크에서 초당 패킷 수(packets/sec)와 초당 비트 수(bits/sec)를 측정한 결과를 나타낸다. 시뮬레이션에 사용한 컴퓨터의 메모리 제한 때문에 7분 동안만 시뮬레이션을 실시하였다. (그림 2)와 달리 비디오 화상 회의 트래픽에 의한 변화는 거의 보이지 않으며 300초부터 발생시킨 슬래머 웜으로 인해 패킷 수와 비트 수가 기하급수적으로 증가하는 것을 볼 수 있다.



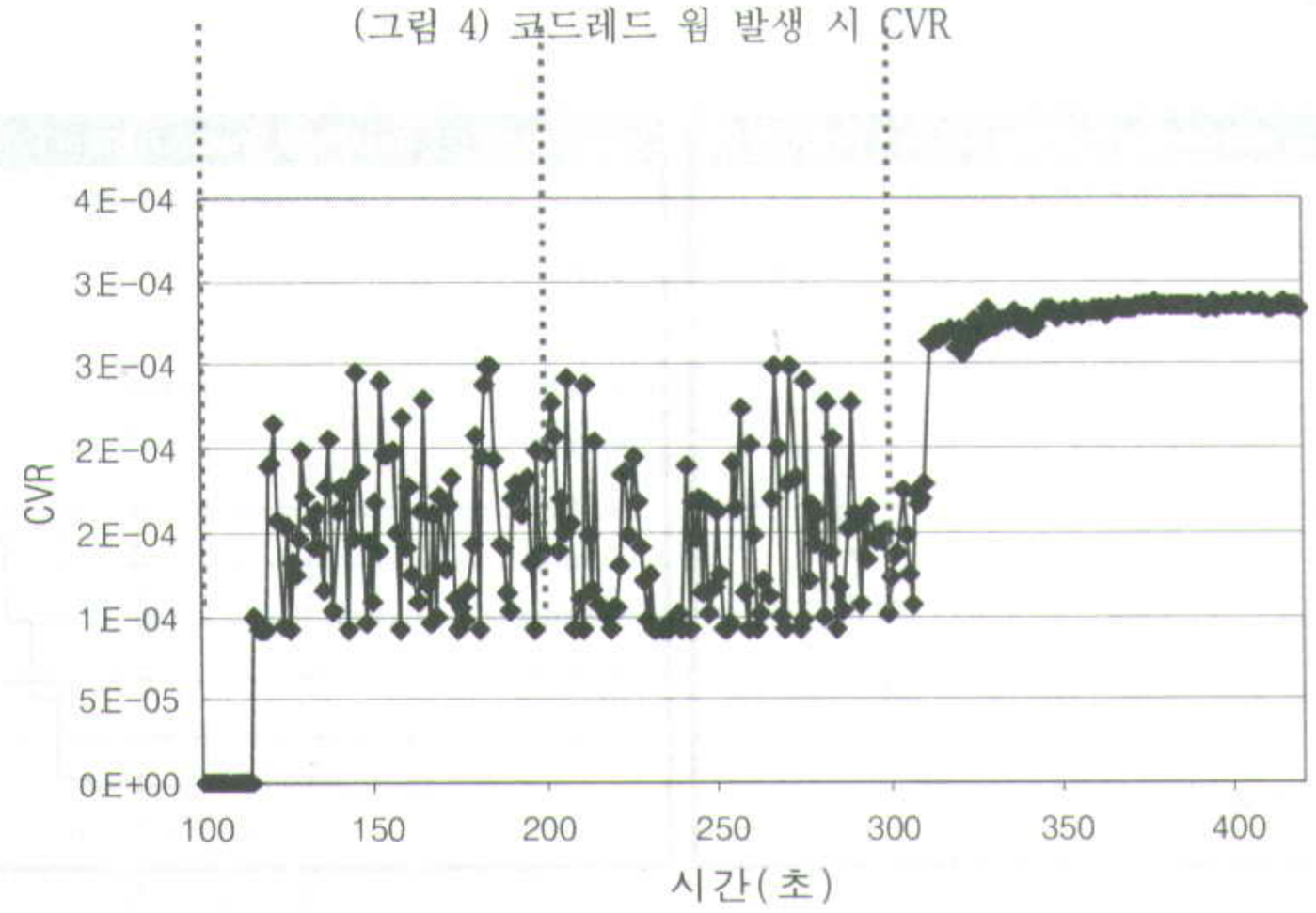
(a) 초당 패킷 수 (b) 초당 비트 수

(그림 3) 슬래머 웜 발생 시 네트워크 트래픽

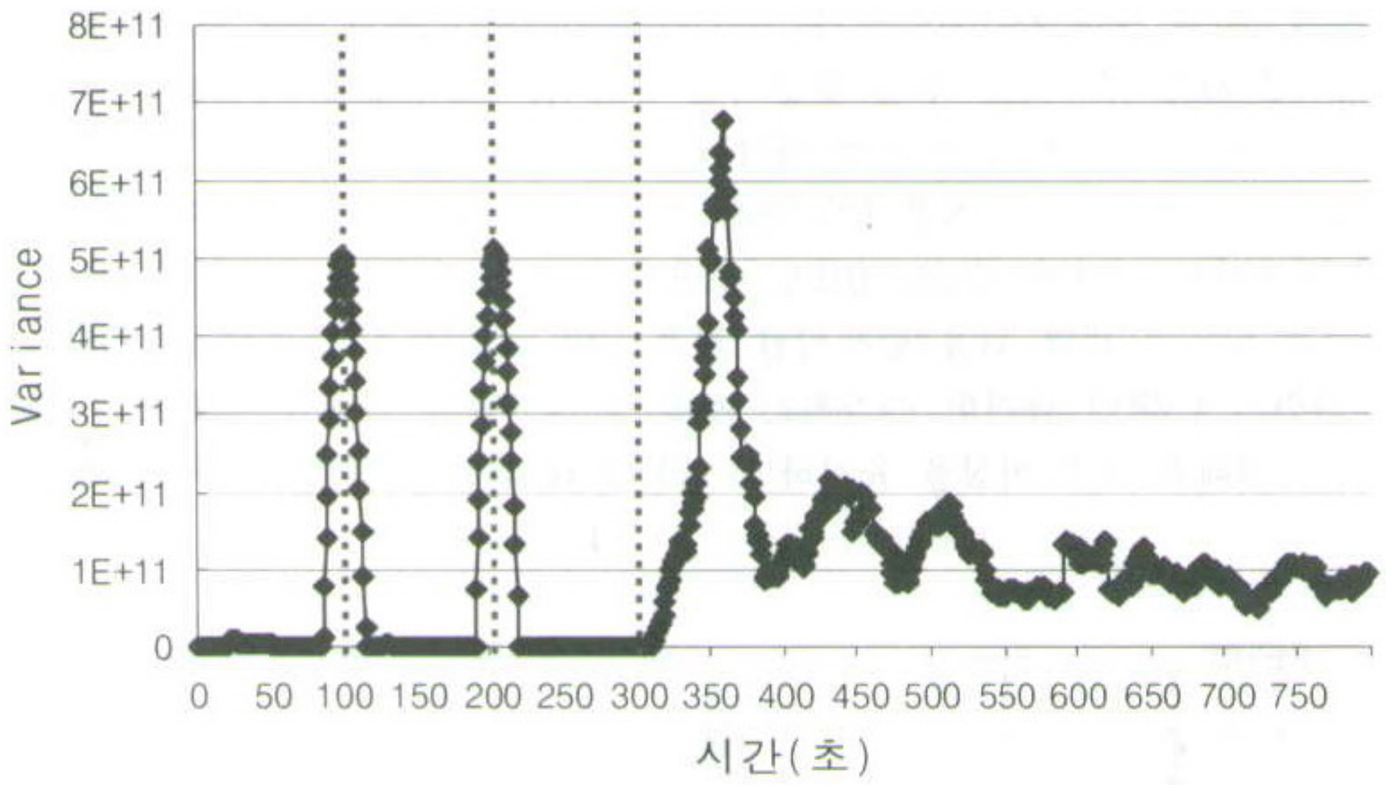
먼저 기존에 제안된 법 중 본 논문에서 제안하는 방법과 같이 트래픽량(bits/sec)을 이용해 이상 트래픽을 탐지하는 방법에 의한 트래픽 측정 결과를 분석하였다. (그림 4)는 코드레드 웜 발생 시 전체 트래픽에 대해서 [6]에서 제안된 방법에 따라 1초에 1번씩 측정한 패킷 수(packets/sec)를 트래픽량(bits/sec)으로 나누어 Packet count to traffic volume ratio(CVR)을 측정한 결과이다. 100초에 비디오 화상 회의 트래픽이 발생하면서 조금 증가했다가 또 다른 비디오 화상 회의 트래픽이 발생한 200초에 다시 감소하여 웜 트래픽이 발생한 300초에도 별 다른 변화 없이 시뮬레이션이 끝날 때까지 비슷한 크기를 유지하였다. [6]에서는 크기가 80byte 이하인 스캐닝 공격의 탐지에 중점을 두고 CVR을 이용한 탐지 방법을 제안하였다. 그러나 코드레드 웜은 패킷의 크기가 약 4kbyte 정도로 크기 때문에 CVR을 이용한 방법으로는 코드레드 웜의 탐지가 어렵다. (그림 5)는 슬래머 웜 발생 시 CVR의 변화를 나타낸다. 코드레드 웜의 경우와 달리 웜이 발생한 300초부터 그 값이 증가하여 큰 값을 유지하므로 이상 트래픽이 발생한 것을 탐지할 수 있다. 따라서 CVR은 작은 크기의 공격 탐지에 더 유리하다는 것을 알 수 있다.



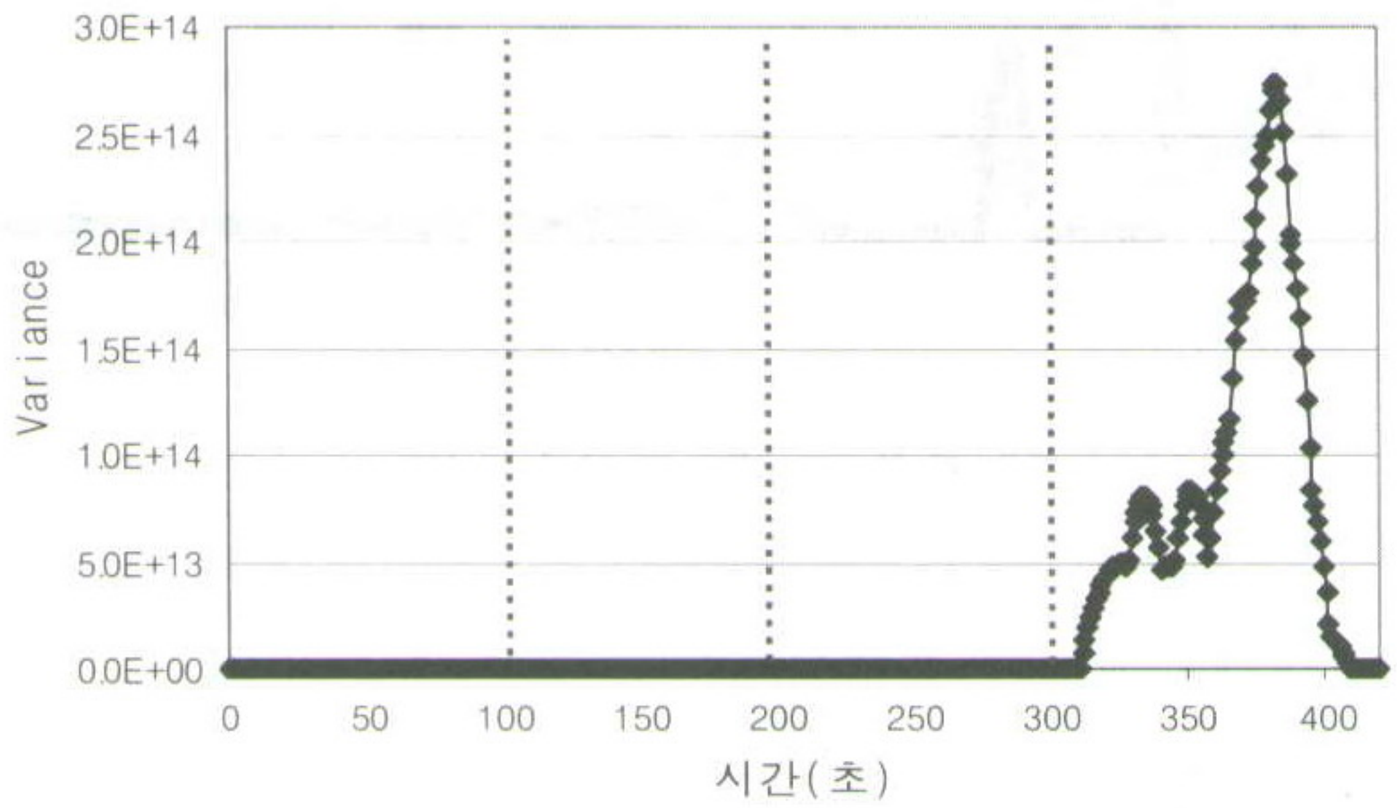
(그림 4) 코드레드 웜 발생 시 CVR



(그림 5) 슬래머 웜 발생 시 CVR



(그림 6) 코드레드 웜 발생 시 variance

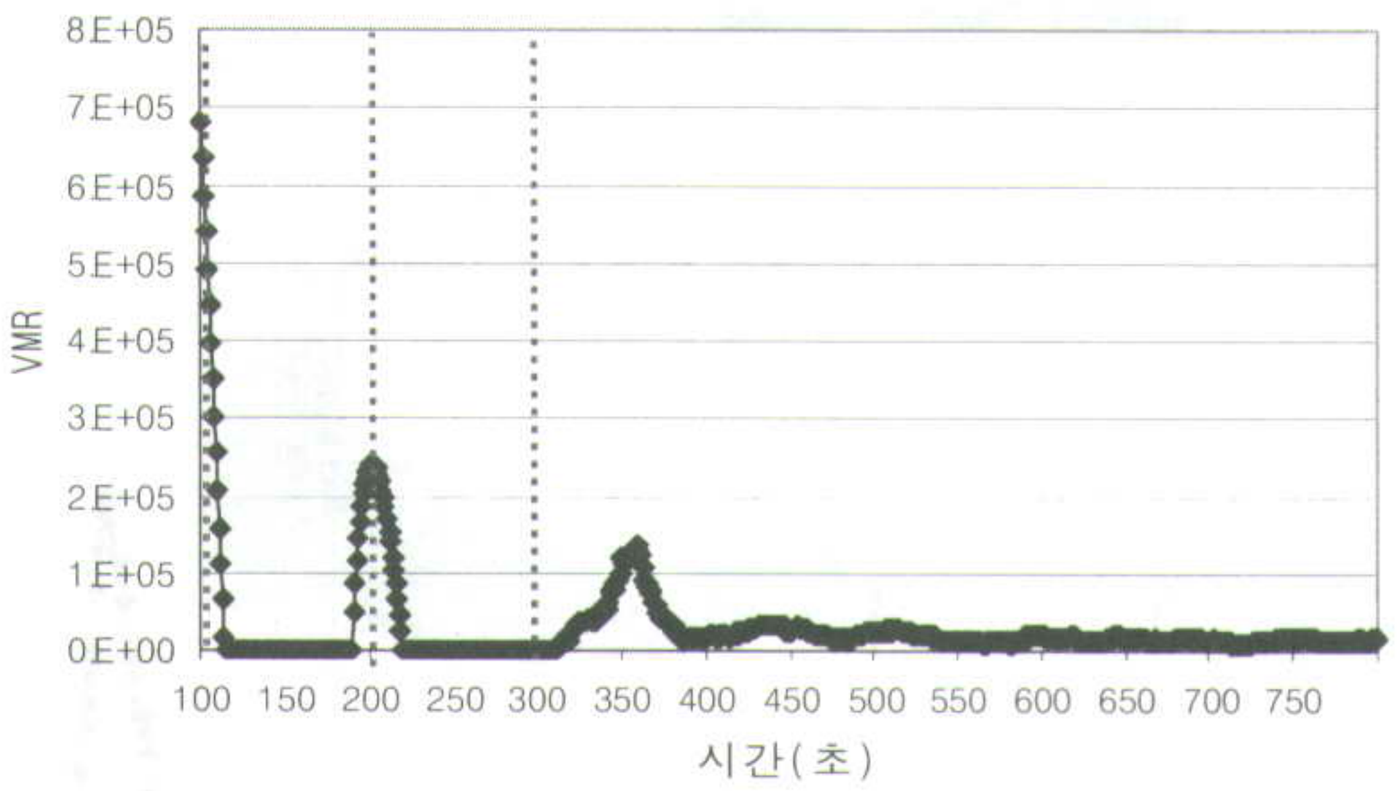


(그림 7) 슬래머 웜 발생 시 variance

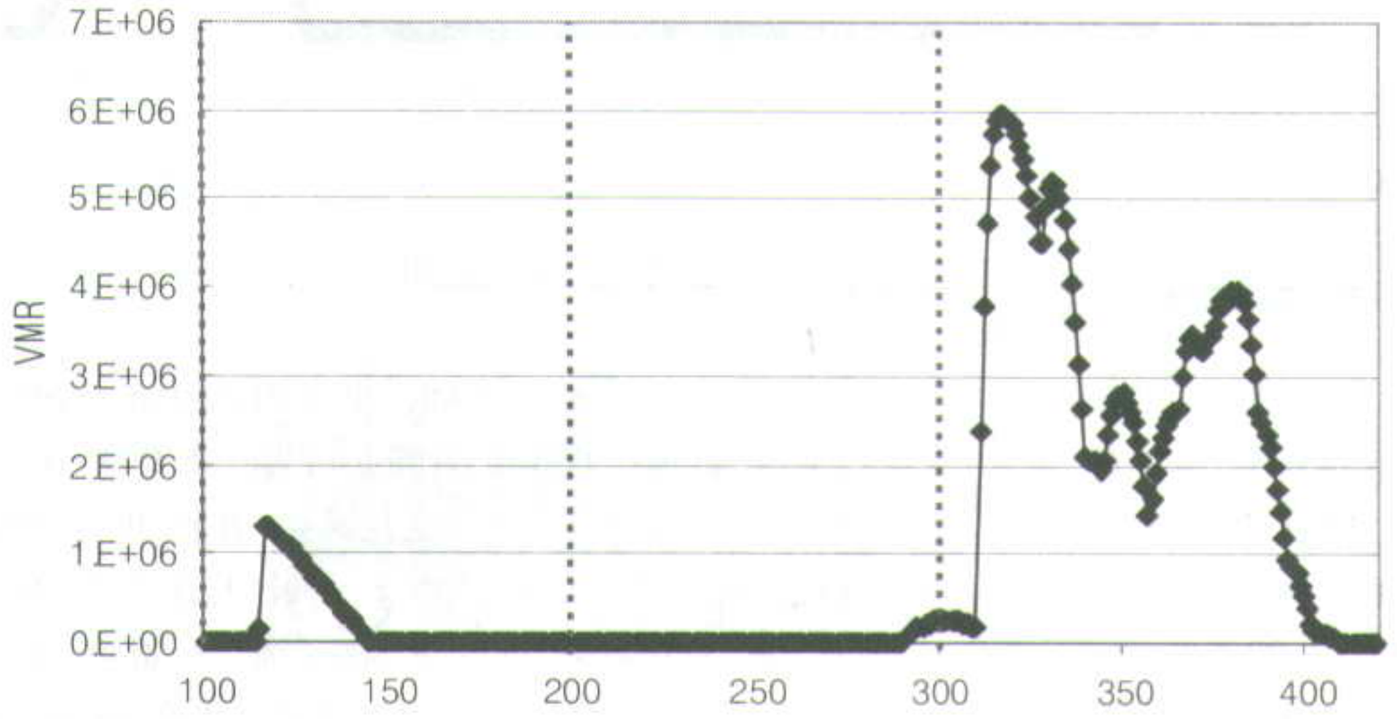
이어서 본 논문에서 제안하는 방법을 이용하여 variance, VMR 및 correlation coefficient를 계산한다. 먼저 Ⅳ장에서 설명한 바와 같이 1초에 1번씩 트래픽량(bits/sec)을 측정한 데이터를 Sliding time window 단위로 묶어 variance를 계산하였다. (그림 6)은 코드레드 웜을 발생시키고 W=30일 때, 시간에 따른 variance의 변화를 나타내고 있는데, 비디오 화상 회의 트래픽이 발생한 100초와 200초에 각각 약 30초간 높은 값을 보이고, 웜 트래픽이 발생한 300초에 그 값이 증가한 후 지속적으로 큰 값을 유지하는 것을 볼 수 있다. (그림 7)는 슬래머 웜을 발생시켰을 때의 variance의 변화를 나타낸다. 이렇게 정상 트래픽과 이상 트래픽의 variance가 큰 차이를 보이므로 variance를 측정하여 이상 트래픽을 탐지할 수 있다. 정상 트래픽에 비해 이상 트래픽의 variance가 증가하는 이유를 생각해보면 비디오 화상 회의 트래픽은 짧은 시간 단위로 일정한 크기의 패킷이 발생하여 트래픽의 변화량이 작은데 비해서 코드레드 웜 트래픽은 패킷과 패킷간의 간격이 비교적 길어서 트래픽의 변화량이 크기 때문이다. 그리고 슬래머 웜의 경우는 짧은 시간에 트래픽량이 급격히 증

가하므로 variance가 증가한다.

다음으로 1초에 1번씩 트래픽량(bits/sec)을 측정한 데이터를 Sliding time window 단위로 묶어 VMR을 구하였다. (그림 8)은 코드레드 웜을 발생시킨 트래픽 측정결과에 대해 W=30일 때, 100초에서 800초 사이의 시간에 따른 VMR의 변화를 나타낸다. variance와 마찬가지로 비디오 화상 회의 트래픽이 발생할 때마다 잠시 크기가 커졌다가 다시 0으로 수렴하며 웜 트래픽이 발생한 300초부터 그 값이 증가하기 시작하여 시간이 지나도 0으로 수렴하지 않고 일정 크기 이상을 계속 유지하는 것을 볼 수 있다. (그림 9)은 슬래머 웜을 발생시킨 경우 VMR을 계산한 결과를 나타낸다. 300초에 웜이 발생하기 시작하면서 VMR이 급격히 증가하는 것을 볼 수 있다. 이렇게 트래픽의 VMR을 측정했을 때 일정시간 이상 정해진 크기 이상을 유지하면 이상 트래픽이라고 판단할 수 있다.

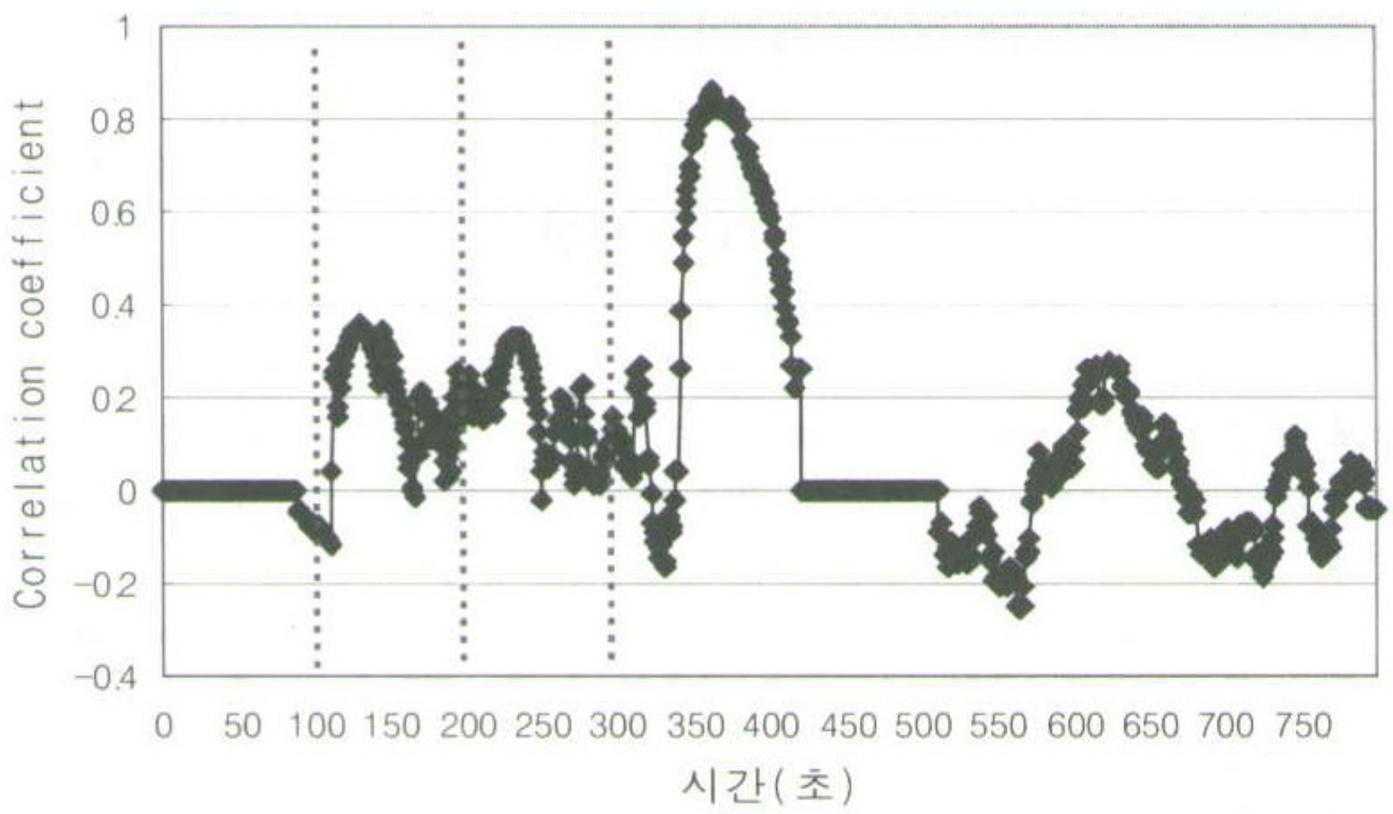


(그림 8) 코드레드 웜 발생 시 VMR

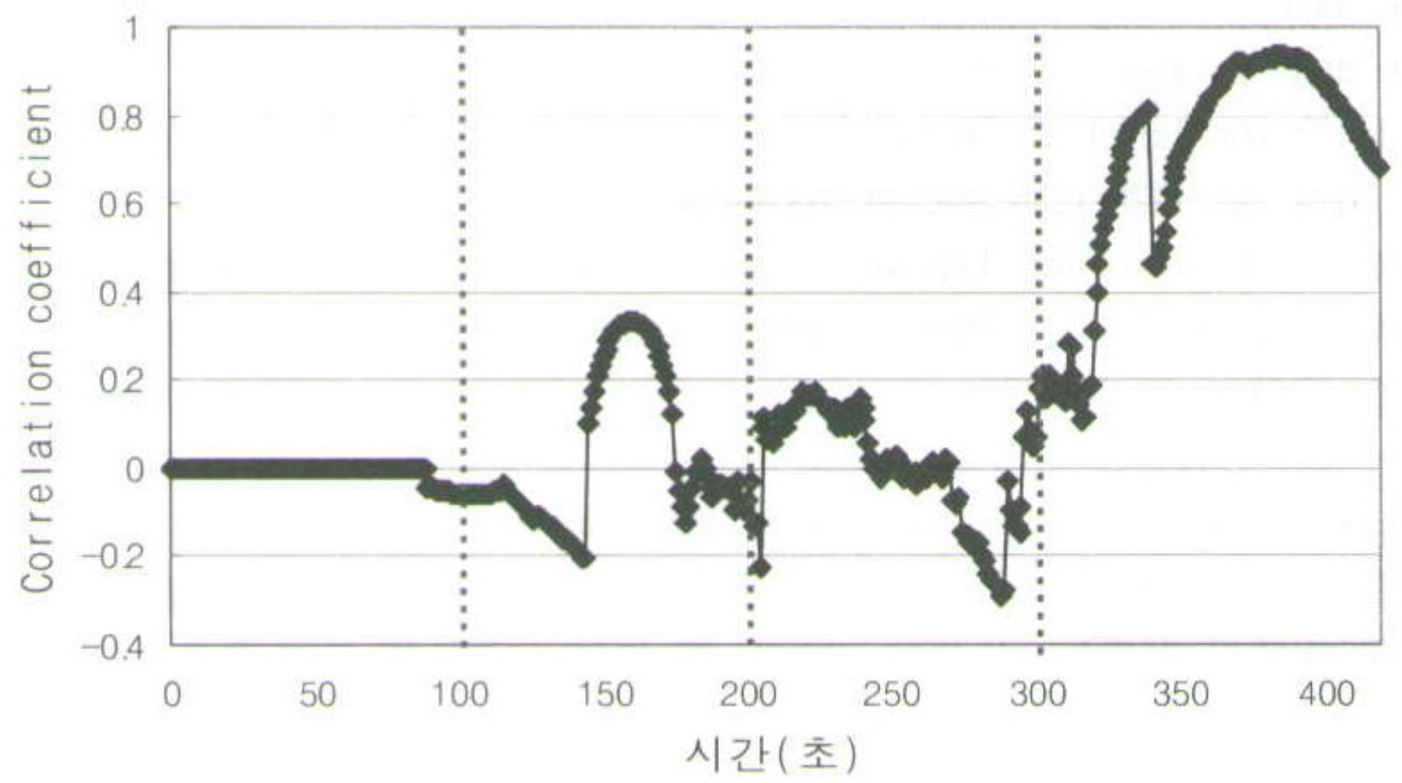


(그림 9) 슬래머 웜 발생 시 VMR

마지막으로 1초에 1번씩 트래픽량(bits/sec)을 측정한 데이터를 Sliding time window 단위로 묶어 30초간의 데이터와 그 후 30초간의 데이터 간의 correlation coefficient를 계산하였다. (그림 10)는 코드레드 웜을 발생시켰을 때의 결과를 나타내고 있으며 100초와 200초에 발생시킨 비디오 화상 회의 트래픽에 의해서 그 크기가 약 0.4까지 증가하는 데 비하여 300초에 발생시킨 웜 트래픽에 의해서는 그 크기가 0.8 이상으로 비디오 화상 회의 트래픽을 발생시켰을 때보다 두 배 이상 커지는 것을 볼 수 있다. (그림 11)은 슬래머 웜을 발생시켰을 때의 결과를 나타내고 있으며 역시 정상 트래픽에 대해서는 0.4 이하의 값을 가지며 300초부터 웜이 발생함에 따라 약 0.9 이상까지 계속 증가하는 것을 볼 수 있다. 이렇게 트래픽의 correlation coefficient를 측정했을 때 정해진 크기 이상의 값이 발생하면 이상 트래픽이라고 판단할 수 있다.



(그림 10) 코드레드 웜 발생 시 correlation coefficient



(그림 11) 슬래머 웜 발생 시 Correlation coefficient

Ⅵ. 결 론

본 논문에서는 스캐닝 웜의 일반적인 트래픽 특성을 분석하여, 네트워크의 트래픽량에 대한 정보만으로 스캐닝 웜을 탐지해냄으로써 탐지 시간을 줄이면서 높은 정확도를 유지할 수 있는 탐지방법을 제안하였다. 그리고 기존에 제안된 방법과 본 논문에서 제안하는 방법에 대해서 시뮬레이션을 통해 성능을 비교 분석하였다. 그 결과 기존의 CVR을 이용하여 스캐닝 웜을 탐지하는 방법은 본 논문에 비해 패킷 수에 대한 정보가 추가적으로 필요하고, 코드레드 웜의 탐지가 어려웠다. 반면에 본 논문에서 제안한 세 가지 방법은 모두 코드레드 웜과 슬래머 웜의 발생을 정확하게 탐지하여 높은 정확도를 보였다. 하지만 트래픽 특성을 이용한 탐지 방법은 빠르고 효율적인 탐지가 가능하지만 웜이 발생한 근원지와 패킷의 특성에 대한 정보 등을 알 수 없기 때문에 웜에 대한 차단 및 대응이 어렵다. 향후에는 패킷 헤더 정보를 이용하는 방법을 같이 고려하여 탐지의 정확도를 높이고, 탐지된 웜에 대한 대응까지 가능한 시스템을 연구할 예정이다.

참 고 문 헌

[1] D. Moore, V. Paxson, S. Savage, C. Shannon, S. Staniford, and N. Weaver, "Inside the Slammer Worm," *IEEE Security & Privacy*, pp. 33-39, July/August 2003.

[2] H. Kim, I. Kang, and S. Bahk, "Real-Time Visualization of Network Attacks on High-Speed Links," *IEEE Network*, pp. 30-39, Sept/Oct 2004.

[3] S. Noh, C. Lee, K. Ryu, K. Choi, and G. Jung, "Detecting Worm Propagation Using Traffic Concentration Analysis and Inductive Learning" *Lecture Notes in Computer Science*, 3177(1), pp. 402-408, 2004.

[4] M. Kim, H. Kang, S. Hong, S. Chung, and W. Hong, "A Flow-based Method for Abnormal Network Traffic Detection," in *Proc. IEEE/IFIP NOMS 2004*, pp. 599-612, April 2004.

[5] C. Zou, W. Gong, D. Towsley, and L. Gao, "The Monitoring and Early Detection of Internet Worms," in *Proc. 10th ACM conference on Computer and communication security*, pp. 190-199, 2003.

[6] B. Roh and S. Yoo, "A Novel Detection Methodology of Network Attack Symptoms at Aggregate Traffic Level on Highspeed Internet Backbone Links," *Lecture Notes in Computer Science*, 3124, pp. 1226-1235, August, 2004.

[7] D. Moore, C. Shannon, and J.Brown, "Code-Red: a case study on the spread and victims of an Internet worm," in *Proc. Second Internet Measurement Workshop*, pp. 273-284, November, 2002.

[8] W32.Blaster.Worm, http://securityresponse.symantec.com/avcenter/venc/data/w32.blaster.worm.html

[9] C. Shannon and D. Moore, "The Spread of the Witty Worm," *IEEE Security & Privacy*, pp. 46-50, July/August 2004.