

연관관계규칙을 이용한 트래픽 폭주 공격 탐지의 심층 분석

유재학, 강봉수, 이한성, 박준상, 김명섭, 박대희
고려대학교 컴퓨터정보학과

e-mail:{dbzzang, ares4you, mohan, runtoyou, tmskim, dhpark}@korea.ac.kr

An In-depth Analysis on Traffic Flooding Attacks Detection using Association Rule Mining

Jaehak Yu, Bongsu Kang, Hansung Lee, Jun-Sang Park, Myung-Sup Kim, Daihee Park
Dept. of Computer & Information Science, Korea University

요 약

본 논문에서는 데이터의 전처리과정으로 SNMP MIB 데이터에 대한 속성 부분집합의 선택 방법(attribute subset selection)을 사용하여 특징선택 및 축소(feature selection & reduction)를 실시하였다. 또한 데이터 마이닝의 대표적인 해석학적 분석 모델인 연관관계규칙기법(association rule mining)을 이용하여 트래픽 폭주 공격 및 공격유형별 SNMP MIB 데이터에 내재되어 있는 특징들을 규칙의 형태로 추출하여 분석하는 의미론적 심층해석을 실시하였다. 공격유형에 대한 패턴 규칙의 추출 및 분석은 공격이 발생한 프로토콜에 대해서만 서비스를 제한하고 관리할 수 있는 정책적 근거를 제공함으로써 보다 안정적인 네트워크 환경과 원활한 자원관리를 지원할 수 있다. 본 논문에서 제시한 트래픽 폭주 공격 및 공격유형별 데이터로부터의 자동적 특징의 규칙 추출 및 의미론적 해석방법은 침입탐지 시스템을 위한 새로운 방법론에 모멘텀을 제시할 수 있다는 긍정적인 가능성과 함께 침입탐지 및 대응시스템의 정책 수립을 지원할 수 있을 것으로 기대된다.

1. 서론

트래픽 폭주 공격 탐지에서의 전통적인 패킷 수집 방법들[1-3]은 공격에 대한 상세한 분석은 가능하나, 고가의 고성능 분석시스템이 요구될 뿐만 아니라 설치 및 운영상의 확장성이 부족하다는 단점을 가지고 있다. 따라서 이를 보완하기 위한 방법으로 최근 SNMP에서의 MIB 정보를 이용한 침입탐지 방법론[1-2]이 주목을 받고 있다. SNMP MIB 정보를 이용하는 DDoS 탐지 방법은 프로토콜별 추이분석, 일주 트래픽 추이분석, 그리고 MIB에서의 특정 속성과 속성 정보간의 상관관계를 이용하는 방법 등으로 분류된다[1-2]. 그러나 이러한 방법론들은 대부분 테스트에 사용된 공격들의 기능과 특성에 의존적으로 개발된 시스템으로, 새로운 공격 형태나 틀이 발견되면 그때마다 새롭게 알고리즘 전체를 수정해야하는 단점을 가지고 있다. 따라서 최근 학계에서는 단순한 문제점의 해결 방안을 데이터마이닝 및 기계학습 기법에서 찾고자 하는 시도가 활발히 진행 중이다.

최근의 연구문헌 조사에 의하면, 기계학습 기법과 SNMP MIB 정보를 이용한 매우 흥미로운 몇 개의 침입탐지 시스템이 발표되었다: Li 등[2]은 SNMP MIB-II 데이터를 probability density function으로 변환한 후, backpropagation 기반의 인공신경망을 이용하여 침입 여부를 결정하는 시스템을 제안하였다. Puttini 등[4]은 SNMP MIB 데이터를 Bayesian 분류기에 적용하여 Mobile Adhoc Networks(MANET)에서의 비정상 트래픽을 탐지하였다. 또한 Yu 등[1]은 Support Vector Machine(SVM)을 이용하여 트래픽 폭주공격을 탐지하고 공격유형별 분류를 수행하는 시스템을 제안하였다. 그러나

효율적인 시스템의 구축이라는 입장만을 견지하는 위의 기계 학습론적 방법론은 시스템 작동 원리의 역학적 해석을 간과하여, 핵심 동작원리를 black-box화 하였다. 따라서 휴리스틱한 방법론이긴 하나 전통적인 DDoS 탐지 방법론의 해석학적 장점도 고려할 수 있는 보다 포괄적인 시스템이 바람직해 보인다.

최근 유재학 등[3]은 의사결정나무계통의 예측모형 알고리즘을 기반으로 트래픽 폭주공격 탐지의 동작원리를 의미론적으로 해석하는 매우 참신한 새로운 시도를 하였다. 하지만, 의사결정나무는 알고리즘의 성격상 예측모형 알고리즘으로써, 부분적인 의미론적 해석만을 제공한다.

본 논문에서는 데이터의 전처리과정으로 SNMP MIB 데이터에 대한 속성 부분집합의 선택 방법(attribute subset selection)을 사용하여 특징선택 및 축소(feature selection & reduction)를 실시한다. 또한 데이터 마이닝의 대표적인 해석학적 분석 모델인 연관관계규칙기법(association rule mining)을 이용하여 트래픽 폭주 공격 및 공격유형별 SNMP MIB 데이터에 내재되어 있는 특징들을 규칙의 형태로 추출하여 분석하는 의미론적 심층해석을 실시한다. 공격유형에 대한 패턴 규칙의 추출 및 분석은 공격이 발생한 프로토콜에 대해서만 서비스를 제한하고 관리할 수 있는 정책적 근거를 제공함으로써 보다 안정적인 네트워크 환경과 원활한 자원관리를 지원할 수 있다. 또한 트래픽 폭주 공격 및 공격유형별 데이터로부터의 자동적 특징의 규칙 추출 및 의미론적 해석방법은 침입탐지시스템을 위한 새로운 방법론의 개발을 위한 모멘텀을 제시할 수 있다는 긍정적인 가능성과 함께 침입탐지 및 대응시스템의 이론적 근거의 제시도 또한 기대된다.

본 논문의 구성은 다음과 같다. 2장에서는 속성 부분집합의 선택 방법과 연관관계규칙기법을 소개한다. 3장에서는 구축된 실험 환경의 소개 및 심층적인 의미론적 해석을 기술하며, 마지막으로 4장에서는 결론 및 향후 연구과제에 대해 논한다.

2. 연구배경

2.1 속성 부분집합의 선택

본 논문에서는 SNMP MIB 데이터에 대한 속성 부분집합의 선택 방법(attribute subset selection) 중 그 성능이 이미 검증된 Hall[6]의 방법을 사용하였다. 이는 최적우선 탐색(best first search) 방법과 속성 혹은 특징(attribute or feature) 값 Y 에 대한 엔트로피(entropy), 목표 클래스(target class)와 속성들 간의 피어슨 상관 계수(Pearson's correlation coefficient)를 이용한 조건부 확률(conditional probability)을 계산하여 전체 속성들의 확률 분포도를 가능한 가깝게 표현할 수 있는 최소 개수의 속성집합을 찾는 방법이다. 먼저 각 속성들에 대한 정보 이익(information gain)을 얻기 위해 임의의 속성 Y 에 대한 엔트로피를 식(1)로 계산한다.

$$H(Y) = - \sum_{y \in Y} p(y) \log_2(p(y)). \quad (1)$$

속성 X 와 Y 사이의 관계는 X 가 주어졌을 때 Y 가 발생하는 조건부 확률로써 식(2)와 같이 계산된다.

$$H(Y|X) = - \sum_{x \in X} p(x) \sum_{y \in Y} p(y|x) \log_2(p(y|x)). \quad (2)$$

각 특징에 대한 정보 이익은 식(1)과 식(2)를 이용하여 식(3)으로 정의된다.

$$Gain = H(Y) + H(X) - H(X, Y). \quad (3)$$

식(3)에서 얻은 정보 이익을 기반으로 식(4)에서와 같이 symmetrical uncertainty를 이용하여 임의의 두 속성 X 와 Y 의 분포와 상관관계를 계산한다. 이때 속성 X 를 기준으로 Y 가 높은 분포와 상관관계를 보이면 전체 속성들을 효율적으로 표현할 수 있는 부분집합에 속성 X 는 포함되지만 Y 는 포함되지 않는다. 마찬가지로 목표 클래스와 속성들 간의 분포와 상관관계를 계산하여 부분집합을 구성한다.

$$\text{symmetrical uncertainty coefficient} = 2.0 \times \left[\frac{Gain}{H(Y) + H(X)} \right] \quad (4)$$

각각의 부분집합 $F_S \subset F$ 가 전체 속성들을 얼마나 효율적으로 표현하는지를 평가하기 위하여 메리트 함수(merit function)(식(5))를 사용한다. 메리트 함수의 값이 가장 큰 부분집합이 전체 속성들을 최적으로 표현할 수 있는 부분집합으로 결정된다[6].

$$Merit(F_S) = \frac{\overline{kr_{cf}}}{\sqrt{k + k(k-1)r_{ff}}} \quad (5)$$

여기서, k 는 부분집합 F_S 에서의 속성들의 개수를 의미하

며, $\overline{r_{cf}}$ 는 F_S 에 포함된 속성의 평균 분포(contribution), $\overline{r_{ff}}$ 는 속성의 평균 상관관계 값이다.

2.2 연관관계규칙 마이닝

연관관계규칙 마이닝이란 데이터 안에 존재하는 각 객체들 간의 의미 있는 연관관계를 찾아내는 방법론으로, 연관관계규칙은 $A \& B \Rightarrow C$ 와 같이 조건 명제의 형태로 표현된다. 먼저 항목들의 집합 $I = \{I_1, I_2, \dots, I_m\}$ 와 각각의 트랜잭션 T 는 $T \subseteq I$ 의 관계를 가진 항목들의 집합이 있을 때, 각각의 트랜잭션 T 는 고유한 트랜잭션 구분자(transaction identifier)를 갖는다. A 를 항목들의 집합이라고 하면, 트랜잭션 T 가 필요충분조건으로 $A \subseteq T$ 를 만족하는 경우에만 트랜잭션 T 가 항목 A 를 포함한다고 한다. 여기서 $A \subseteq I, B \subseteq I, A \cap B = \emptyset$ 을 만족하는 경우 연관규칙은 $A \Rightarrow B$ 의 형식으로 표현된다. 규칙 $A \Rightarrow B$ 는 트랜잭션 집합 D 에서 집합 A 와 B 를 동시에 포함하는 트랜잭션의 백분율이 s 인 경우 지지도(support) s 를 갖는다고 한다. 이는 확률 $P(A \cup B)$ 를 계산함으로써 얻을 수 있다. 집합 A 를 포함하는 트랜잭션 중에서 집합 B 도 포함하고 있는 트랜잭션의 백분율이 c 인 경우, 규칙 $A \Rightarrow B$ 는 신뢰도(confidence) c 를 갖는다고 한다. 신뢰도는 조건부확률 $P(B|A)$ 를 계산함으로써 얻을 수 있다[7-8].

$$\begin{aligned} \text{support}(A \Rightarrow B) &= P(A \cup B) \\ \text{confidence}(A \Rightarrow B) &= P(B|A) \end{aligned} \quad (6)$$

최소 지지도 임계값(minimum support threshold)과 최소 신뢰도 임계값(minimum confidence threshold)을 동시에 만족하는 규칙을 강한(strong) 규칙이라고 한다. 이때 최소 지지도 값 이상을 갖는 항목집합(itemset)을 빈발항목집합(frequent itemset)이라 하고 k 개의 항목들로 이루어진 빈발항목 집합을 k -빈발항목집합이라고 한다. 이진 연관규칙에 대한 빈발항목집합을 찾는데 유용한 Apriori 알고리즘[7]은 k 번째 항목 집합이 $(k+1)$ 번째 항목집합을 발견하기 위해 사용되는 반복적 접근방법을 사용하는데, 이는 수준별(level-wise) 방법으로 알려져 있다. 먼저, 빈발 1-항목집합을 L_1 로 나타내며, L_1 은 2-항목집합인 L_2 를 찾는데 사용되고 이것은 다시 L_3 를 찾는데 이용된다. 이러한 방법은 더 이상 빈발 k -항목집합이 없을 때까지 진행된다.

3. 실험 및 의미론적 해석

3.1 실험 환경

본 논문에서 트래픽 폭주 공격 실험을 위하여 그림 1과 같은 실험 환경을 구축하였다. 하나의 L2 스위치 장비에 타깃 시스템(victim)을 연결하고, DDoS 공격 환경을 만들기 위하여 2대의 attack agent와 1대의 attack handler를 L2 스위치 장비에 연결하였다. 또한 타깃 시스템으로부터 MIB 정보의 수집 및 탐지를 위해 1대의 탐지 시스템을 L2 스위치에 연결하였다. L2 스위치는 학내 네트워크를 통하여 인터넷과 연결되어 있기에 타깃 시스템에서는 다양한 인터넷 트래픽이 생성된다. 실제 타깃 시스템에 Apache Web Server, VNC Server, FTP Server, SSH Server, Samba Server 등의 다양한 서버를 운영함으로써 다양한 종류의 정상 트래픽을 발생하였다. 본 실험에서 사

용된 시스템은 모두 Linux Fedora 7 또는 8이며, 타깃 시스템의 SNMP agent는 Net-SNMP v5.4.1이 사용되었다.

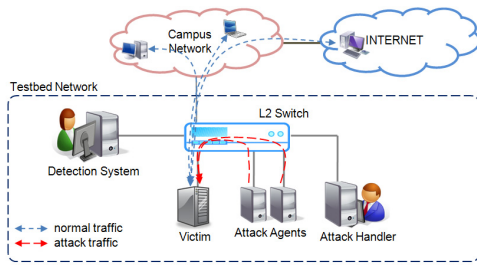


그림 1. 실험 환경 구성도

3.2 실험데이터

본 논문에서는 java 기반의 machine learning tool인 weka[9]를 사용하였으며, RFC1213[10]에서 정의한 mib-2 그룹의 MIB 속성들 중, 실제 트래픽 폭주 공격에 반응하는 13개의 MIB 속성들만을 선정하였다. 또한 트래픽 폭주 공격의 대표적 공격 톨인 Stacheldraht[11]를 이용하여 TCP-SYN flooding 공격, UDP flooding 공격, ICMP flooding 공격 등을 타깃 시스템에서 실시하였다. 본 논문의 실험에서 사용된 MIB 속성들을 [표 1]에 정리하였다.

[표 1] 탐지 시스템에서 사용된 MIB 속성들[10]

mib-2 group	SNMP MIB objects	MIB object description
ip	ip.ipInReceives	인터페이스로부터 받은 ip 데이터그램의 총 개수
	ip.ipInDelivers	수신된 ip 데이터그램 중 상위계층으로 전달된 데이터그램의 총 개수
	ip.ipOutRequests	상위계층에서 송신 요청한 ip 데이터그램의 개수
	ip.ipOutDiscards	정상적으로 송신 요청된 ip 데이터그램 중 buffer overflow 등에 의해 drop 되는 ip 데이터그램의 개수
tcp	tcp.tcpAttemptFails	비정상적으로 tcp connection이 종료된 횟수
	tcp.tcpOutRsts	RST flag를 포함하여 송신된 tcp 세그먼트의 개수
udp	udp.udplnErrors	udp 패킷 구성오류에 의해 전달되지 못한 데이터그램의 개수
icmp	icmp.icmplnMsgs	수신된 icmp 메시지의 총 개수
	icmp.icmplnDestUnreachs	수신된 icmp destination unreachable 메시지의 총 개수
	icmp.icmplnEchos	수신된 icmp echo 메시지의 총 개수
	icmp.icmpOutDestUnreachs	송신된 icmp destination unreachable 메시지의 개수
	icmp.icmpOutEchoReps	송신된 icmp echo response 메시지의 개수
	icmp.icmpOutMsgs	송신된 icmp 메시지의 총 개수

3.3 연관관계규칙의 심층적 분석

첫 번째 실험은 정상트래픽과 공격트래픽 데이터에 내재되어 있는 유용한 패턴들의 발견 및 이에 대한 심층적 분석 실험으로써, 정상트래픽 1000개와 공격트래픽은 유형별로 500개씩 랜덤하게 추출하여 실험하였다. weka에서의 CFS(Correlation Feature Selection, 식(5))로 선택된 3개의 최적 속성 부분집합은 {ipInReceives, udpInErrors, icmpOutMsgs}이며, 최소 지지도는 10%, 최소 신뢰도는 80%로 하였다. 실험결과 추출된 특징들 간의 주요 연관관

계 규칙들을 [표 2]에 정리하였다.

[표 2] 정상/공격트래픽 분류를 위한 연관규칙들

규칙 번호	규칙 내용	신뢰도
1	udpInErrors \geq 1545 이면, 공격에 속하는 비율이 높다.	100%
2	ipInReceives \geq 75945 이면, 공격에 속하는 비율이 높다.	99%
3	icmpOutMsgs \geq 85162 이면, 공격에 속하는 비율이 높다.	100%
4	ipInReceives \geq 75945 이고 udpInErrors \geq 1545 이면, 공격에 속하는 비율이 높다.	100%
5	ipInReceives \geq 75945 이고 icmpOutMsgs \geq 85162 이면, 공격에 속하는 비율이 높다.	100%
6	ipInReceives $<$ 75945 이고 icmpOutMsgs $<$ 12166 이고 udpInErrors $<$ 1545 이면, 정상에 속하는 비율이 높다.	95%

[표 2]의 규칙들을 종합적으로 분석해보면, 공격 트래픽에 관한 규칙(규칙번호 1-5)의 경우, 트래픽 폭주 공격의 성격상 대량의 패킷을 전송하기 때문에 ipInReceives와 icmpOutMsgs 및 udpInErrors 중에 하나라도 큰 값을 가질 때가 공격트래픽의 전형적 패턴임을 보여준다. 특히 전송받은 데이터그램의 개수, 송신된 icmp 메시지의 개수가 많거나 또는 udp 패킷 구성오류로 전달되지 못하는 데이터그램의 개수가 클 때가 트래픽 폭주 공격의 전형적 패턴임을 보여준다.

정상트래픽(규칙번호 6)의 경우, ipInReceives와 icmpOutMsgs 및 udpInErrors의 MIB 값들이 모두 작은 값을 갖는 패턴을 보이고 있다. 이는 정상적인 트래픽일 때에는 트래픽 폭주공격에 비해 상대적으로 전송받은 데이터그램의 개수가 적고, 특히 수신된 icmp 메시지의 개수와 udp 패킷 구성오류로 전달되지 못하는 데이터그램의 개수가 적다는 정상트래픽의 패턴을 보여준다. 특히 [3]의 실험결과와 비교했을 때, 본 연구에서는 [3]에서 제공하는 규칙들(규칙번호 1, 2, 6)을 포함하는 보다 종합적인 규칙들(규칙번호 3-5)을 제공함을 확인하였다.

두 번째 실험은 DDoS의 공격유형에 대한 패턴의 추출 및 의미론적 해석을 위한 실험으로써, 공격유형별로 500개씩 랜덤하게 추출하여 실험하였다. weka에서의 CFS(식(5))로 선택된 5개의 속성 부분집합은 {ipInDelivers, icmpInMsgs, icmpInEchos, tcpOutRsts, udpInErrors}이며, 최소 지지도는 10%, 최소 신뢰도는 80%로 하였다. 실험결과 추출된 속성들 간의 주요 연관관계규칙들을 [표 3]에 정리하였다.

[표 3] 공격유형별 분류를 위한 연관관계규칙들

규칙 번호	규칙 내용	신뢰도
1	tcpOutRsts \geq 14418 이면, TCP-SYN flooding 공격에 속하는 비율이 높다.	100%
2	tcpOutRsts \geq 14418 이고 ipInDelivers \geq 32541 이고 icmpInMsgs $<$ 17027 이고 icmpInEchos $<$ 16974 이고 udpInErrors $<$ 2207 이면, TCP-SYN flooding 공격에 속하는 비율이 높다.	100%
3	udpInErrors \geq 2207 이고 icmpInMsgs $<$ 17027 이고 icmpInEchos $<$ 16974 이고 tcpOutRsts $<$ 14418 이고 ipInDelivers $<$ 32541 이면, UDP flooding 공격에 속하는 비율이 높다.	100%
4	icmpInEchos \geq 67896 또는 icmpInMsgs \geq 85135 이면, ICMP flooding 공격에 속하는 비율이 높다.	100%
5	ipInDelivers \geq 32541 이고 icmpInMsgs $<$ 17027 이고	99%

	udpInErrors < 2207 이면, TCP-SYN flooding 공격에 속하는 비율이 높다.	
6	ipInDelivers < 32541 이고 icmpInMsgs < 17027 이고 tcpOutRsts < 14418 이면, UDP flooding 공격에 속하는 비율이 높다.	81%
7	ipInDelivers < 32541 이고 tcpOutRsts < 14418 이고 udpInErrors < 2207 이면, ICMP flooding 공격에 속하는 비율이 높다.	81%

TCP-SYN flooding 공격은 IP를 속여 서버에 메시지를 전송하고 서버는 클라이언트에 SYN/ACK를 보내 클라이언트가 ACK를 받기 위해 무기한 대기시키는 공격 방법이다. TCP-SYN flooding에 대한 규칙들(규칙번호 1-2)을 분석해보면, tcpOutRsts 값이 크거나 tcpOutRsts와 ipInDelivers 값이 동시에 클 때는 TCP-SYN flooding 공격이라는 지식을 얻었다. 특히 RST flag를 포함하여 송신된 tcp 세그먼트의 개수가 큰 값을 갖고, 수신된 ip 데이터그램 중 상위계층으로 전달된 데이터그램의 총 개수가 큰 값을 가질 때 TCP-SYN flooding 공격임을 확인하였다.

UDP flooding 공격은 공격대상에 연속적으로 패킷을 보냄으로써, 공격 대상 네트워크의 대역폭을 소모시켜 정상적인 서비스를 마비시키는 공격 방법이다. UDP flooding 공격(규칙번호 3)은 udpInErrors 값을 제외한 나머지 MIB들이 비교적 작은 값을 가질 때 UDP flooding 공격에 해당하는 패턴임을 확인하였다. 또한 udpInErrors는 udp 패킷 구성오류에 의해 증가되는 값으로 UDP flooding 공격의 특징을 설명하는 유용한 패턴임을 알 수 있었다.

마지막으로 ICMP flooding 공격은 icmp 핑 메시지를 수신한 네트워크에 있는 모든 시스템들이 응답하면, 핑 메시지의 내용을 변조시켜 네트워크의 대역폭을 소모시키는 공격 방법이다. ICMP flooding 공격패턴(규칙번호 4)은 icmpInMsgs 또는 icmpInEchos 값이 클 때이며, 송신된 icmp 메시지의 총 개수를 의미하는 icmpInMsgs나 수신된 icmp 메시지의 개수를 의미하는 icmpInEchos 값은 TCP-SYN flooding 및 UDP flooding 공격에 영향을 받지 않기 때문에 ICMP flooding 공격을 의미하는 유용한 패턴으로 활용될 수 있음을 확인하였다.

특히 [3]의 실험결과와 비교했을 때, 본 연구에서는 [3]에서 제공하는 규칙들(규칙번호 3-5)을 포함한 보다 종합적인 규칙들(규칙번호 1-2, 6-7)을 제공함을 확인하였다.

4. 결론

본 논문에서는 데이터마이닝의 대표적 분석 모델인 연관관계규칙기법(association rule mining)을 사용하여 트래픽 폭주공격과 공격유형별 데이터 속에 내재되어 있는 유용한 지식의 발견과 심층적 분석을 수행하였다. 트래픽 폭주 공격 및 공격유형별 데이터로부터의 자동적 특징의 규칙 추출 및 의미론적 해석방법은 침입탐지시스템을 위한 새로운 방법론의 개발을 위한 모멘텀을 제시할 수 있다는 긍정적인 가능성과 함께 침입탐지 및 대응시스템의 이론적 근거의 제시도 또한 기대된다. 본 연구에서 시도한 연관관계규칙을 기반으로 한 공격 및 공격유형에 대한 의미론적 심층 해석방법은 이제까지 시도되지 않았던 참신한 연구방향으로 향후 기대되는 공헌 가능성이 크다고 할 수 있다.

향후 연구과제로는 침입대응시스템 및 침입예방시스템

의 정책 수립 등을 지원할 수 있는 보다 구체적인 연구를 수행하고자 한다.

참고문헌

- [1] J. Yu, H. Lee, M. Kim, and D. Park, "Traffic flooding attack detection with SNMP MIB using SVM", Computer Communications, In Press.
- [2] J. Li and C. Manikopoulos, "Early statistical anomaly intrusion detection of DOS attacks using MIB traffic parameters", Information Assurance Workshop, IEEE, pp. 53-59, 2003.
- [3] 유재학, 오승근, 이한성, 박준상, 김명섭, 박대희, "트래픽 폭주 공격 탐지 시스템의 의미론적 해석", 한국정보처리학회 추계학술대회(심사중), 2008.
- [4] R. Puttini, M. Hanashiro, F. Miziara, R. Sousa, L. García-Villalba, and C. Barencó, "On the anomaly intrusion-detection in mobile adhoc network environments", Proc. of PWC 2006, LNCS 4217, pp. 182-193, 2006.
- [5] Y. Wu, and A. Zhang, "Feature selection for classifying high-dimensional numerical data", IEEE Conference on Computer Society, CVPR 2004, Vol. 2, pp. 251-258, 2004.
- [6] M. Hall, "Correlation-based Feature Selection for Machine Learning", PhD Diss. Department of Computer Science, Waikato University, Hamilton, NZ, 1998.
- [7] J. Han and M. Kamber, Data Mining: Concept and Techniques, Morgan Kaufmann Publishers, 2nd Ed., pp. 227-242, 2007.
- [8] B. Wu, W. Zhou, W. Zhang, "The applications of data mining technologies in dynamic traffic prediction", Intelligent Transportation Systems, IEEE, Vol. 1, pp. 396-401, 2003.
- [9] Machine Learning Lab in The University of Waikato, <http://www.cs.waikato.ac.nz/ml>.
- [10] IETF RFC 1213, "Management information base for network management of TCP/IP-based internets: MIB-II", <http://www.rfc-editor.org/rfc/rfc1213.txt>.
- [11] "Distributed denial of service (DDoS) attacks/tools", <http://staff.washington.edu/dittrich/misc/ddos/>.