WDM 망에서 인공면역체계 기반의 네트워크 공격 탐지 제어 모델 및 대응 기법 설계

정회원 유 경 민*, 양 원 혁*, 종신회원 김 영 천**°

Design of Network Attack Detection and Response Scheme based on Artificial Immune System in WDM Networks

Kyung-Min Yoo*, Won-Hyuk Yang* *Regular Members,* Young-Chon Kim**° *Lifelong Member*

요 약

동적인 네트워크 공격에 대응하기 위하여 인공 신경망, 유전 알고리즘, 면역 알고리즘과 같은 지능적 기술들이 공격 탐지에 적용되어 왔으며 최근에는 인공 면역 체계를 이용한 공격 탐지가 활발히 연구되고 있다. 기존의 인공면역체계 기반의 공격 탐지 기법들은 주로 자기 세포 집합과 비교를 통하여 항원을 인지하고 제거하는 부정 선택 원리만을 이용하였다. 그러나 실제 네트워크에서는 정상 상태와 비정상 상태가 거의 유사한 상태를 보이는 경우가 발생하므로 오탐지가 빈번히 발생하는 문제점이 있다. 이러한 문제점을 해결하기 위하여 본 논문에서는 새로운 인공면역체계 기반의 공격 탐지 및 대응 기법을 제안하고 그 성능을 평가한다. 제안하는 기법에서는 인간면역체계에서 발생하는 수지상 세포와 T 세포의 면역 상호 작용을 적용하여 버퍼 점유율 변화를 이용한 검출기 집합을 발생시키고 공격 탐지 모듈과 대응 모듈을 다음과 같이 설계하였다. 첫째, self/non-self 구별을 위한 부정 선택 원리를 이용하여 검출기 집합을 발생시킨다. 둘째, 공격 탐지 모듈에서는 발생된 검출기 집합을 이용하여 네트워크 이상 상태를 탐지하고 경고 신호를 발생시킨다. 이때 오탐지를 줄이기 위하여 위험이론을 적용하며 위험도를 추측하기 위해 퍼지 이론을 이용한다. 마지막으로 공격 대응 모듈에서는 역추적된 공격 노드에 제어 신호를 전송하여 공격 트래픽을 제한하도록 한다.

Key Words : AIS, Attack Detection, Immune System, WDM

ABSTRACT

In recent, artificial immune system has become an important research direction in the anomaly detection of networks. The conventional artificial immune systems are usually based on the negative selection that is one of the computational models of self/nonself discrimination. A main problem with self and non-self discrimination is the determination of the frontier between self and non-self. It causes false positive and false negative which are wrong detections. Therefore, additional functions are needed in order to detect potential anomaly while identifying abnormal behavior from analogous symptoms. In this paper, we design novel network attack detection and response schemes based on artificial immune system, and evaluate the performance of the proposed schemes. We firstly generate detector set and design detection and response modules through adopting the interaction between dendritic cells and T-cells. With the sequence of buffer occupancy, a set of detectors is generated by negative selection. The detection module detects the network

※ 본 논문은 2009년도 교육과학기술부의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(2009-0077301)
* 전북대학교 컴퓨터공학과, ** 전북대학교 IT정보공학부 영상정보신기술연구소(° : 교신저자)
논문번호 : KICS2009-10-494, 접수일자 : 2009년 10월 29일, 최종논문접수일자 : 2010년 3월 24일

anomaly with a set of detectors and generates alarm signal to the response module. In order to reduce wrong detections, we also utilize the fuzzy number theory that infers the degree of threat. The degree of threat is calculated by monitoring the number of alarm signals and the intensity of alarm occurrence. The response module sends the control signal to attackers to limit the attack traffic.

Ⅰ. 서 론

분산 서비스 거부 공격을 탐지하기 위하여 다양한 기법들이 연구되어 왔으며 이러한 기법들은 그 분석 방법에 따라 오용 탐지 기법과 이상 상태 탐지 기법으로 분류될 수 있다. 오용 탐지 기법은 시스템을 침입할 때 이용된다고 알려진 침입 패턴이나 증거들을 참조하여 탐지하는 것으로 새로운 형태의 침입이 발생하면 탐지하지 못하는 단점이 있다. 반면에 이상 상태 탐지 기법은 시스템의 정상적인 동작들을 정의하고 분류하여 정상적인 동작에서 벗어나는 현상이 발생하는 경우 침입으로 간주한다. 즉, 알려지지 않은 침입 패턴이 발생하더라도 탐지가 가능하다[1]. 그러나 CPU나 메모리 사용률 등의 임계값 기반으로 이상 상태 탐지를 수행하기 때문에 자원 사용의 동적 변화 상태에 따라 오탐지가 발생하는 문제점이 있다. 따라서 정확한 이상 상태 탐지를 위해서는 다양한 공격 형태의 변화를 고려하여 새로운 형태의 침입이나 이상 상태를 탐지할 수 있는 탐지 기법이 요구되어지고 있다[2].

이를 해결하기 위한 방안으로 인간면역체계의 학습 및 기억, 자기 조절, 새로운 항원에 대한 적응력 등을 이용하여 침입 탐지, 네트워크 고장 탐지, 컴퓨터 보안 등에 적용하는 인공면역체계(Artificial Immune System: AIS) 기반의 연구가 활발히 진행되고 있다[2,5]. 그러나 기존의 AIS 기반의 공격 탐지 기법들은 주로 부정 선택 원리만을 이용하여 여전히 오탐지가 높게 발생하는 문제점이 있다.

Ⅱ. 관련 연구

2.1 인간면역체계

인간면역체계는 특정 조직, 기관, 세포나 화학 물질 등의 복잡한 네트워크로 구성되며 외부 병원체를 인식하고 이를 무력화시키거나 제거하는 기능을 가지고 있다.

일반적으로 특별한 면역반응을 일으키는 물질을 항원이라고 하며 효과적인 방어 체계를 구축하기 위해서 면역 시스템은 오직 외부 항원에만 반응해야 한다[3]. 즉, 자기(self) 세포에 해당하지 않는 항원의 구별이 필수적인 특성인데 그림 1은 항원이 인식되고 면역 세포가 활성화되어 항원이 제거되는 과정을 순차적으로 보이고 있다. (II)단계는 (I)단계에서 침입한 항원에 대해 수지상 세포 같은 항원 제시 세포(Antigen Presenting Cell: APC)가 항원을 섭취하는 과정을 보여준다. 그림 1과 같이 항원은 펩타이드 조각들로 이루어져 있고 이러한 펩타이드 조각은 MHC(Major Histo-compatibility Complex)의 분자와 반응하여 APC의 표면에 MHC/펩타이드 복합체로 나타난다. (III)단계에서는 APC가 제시한 복합체를 인식할 수 있는 수용체를 가진 T 세포가 복합체를 인식하고 (IV)단계에서는 활성화된 T세포가 림포카인을 분비함으로써 면역 시스템을 작동시키며 B세포도 활성화시킨다. B세포는 MHC 분자의 도움 없이도 항원을 인식할 수 있는 수용체를 가지고 있기 때문에 (V)단계에서와 같이 수용체를 이용하여 특정 항원에 직접 반응하기도 한다. 항원 결합 반응이 일어난 B세포는 활성화되어 급격히 증식하게 되고 증식된 다량의 개체는 항체를

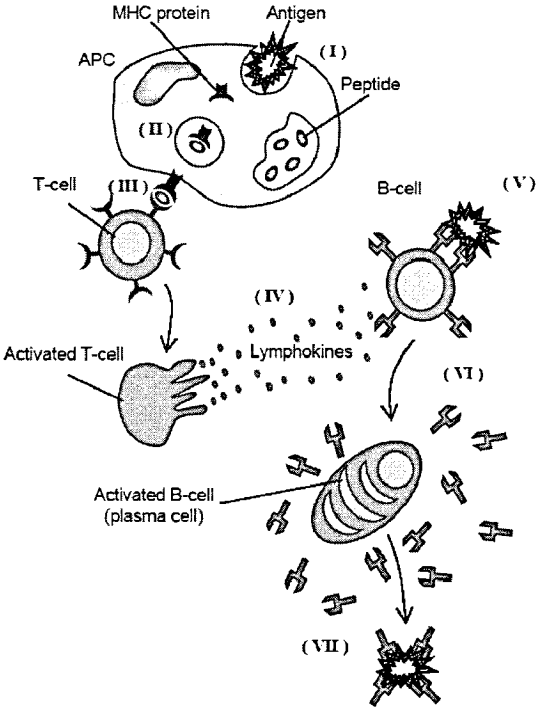

그림 1. 항원 인식 및 면역 세포 활성화
Fig. 1. Antigen recognition & immune response

생성한다. 결국 (VI)단계에서 분비된 항체들이 (VII)에서 항원을 파괴하게 된다.

본 논문에서는 선천적 면역 특성을 가지는 수지상 세포와 후천적 면역 특성을 가지는 T 세포의 상호 작용을 공격 탐지에 적용한다. 수지상 세포들은 항원으로 인한 손상의 증거를 잡기 위해 주요 조직을 감시하는 선천적인 침입 탐지 관리자이다. 따라서 T 세포는 수지상 세포가 탐지한 항원에 대해 T 세포 표면에 있는 항원 수용체를 통하여 항원을 인식하며, 이러한 항원 수용체가 T 세포 수용체이다. 림프 절에 있는 성숙된 T 세포가 수지상 세포가 제시한 MHC와 항원의 복합체를 인식하게 되면, 그 T 세포는 증식되어 특정한 항원 수용체를 가진 T 세포의 수가 늘어난다. 복제된 T 세포는 기억 T 세포와 작용 T세포로 구별된다. 작용 T세포는 면역 반응 조절에 관여하는 보조 T세포(Helper T-cell)와 조직 세포에 침입한 세포내의 병원균을 죽이는 독성 T 세포(Cytotoxic T-cell: CTL)로 나누어진다. 활성화된 T 세포가 보조 T 세포이면 B세포를 자극하여 항체를 생성할 수 있도록 하거나 대식세포 활성화 그리고 CTL의 활성화 등 면역 반응을 촉진시키는 역할을 수행한다[4]. 이처럼 체내에서는 선천적 면역에 관여하는 수지상 세포와 후천적 면역에 관여하는 T 세포 사이의 상호 작용을 통하여 인간면역 시스템의 주요 기능이 수행된다.

2.2 위험 이론

부정 선택과 같이 이미 알려진 면역 이론들은 면역 시스템이 체내의 자기와 비자기(non-self)를 구분하고 비자기로 밝혀진 세포들은 모두 제거하는 방식이다. 그러나 체내에서는 기존의 면역 이론으로는 설명할 수 없는 현상이 있다. 예를 들어 소화 기관에 있는 외부 박테리아나 우리가 먹는 음식이 외부 물질이라고 하더라도 이것들에 대하여 면역 반응이 일어나지 않는다. 위험이론은 이러한 현상을 규명하기 위해 1994년 Matzinger에 의하여 제안되었다[5,6]. 위험 이론은 면역 시스템이 외부 물질에 대하여 모두 반응하지는 않는다. 중요한 특징은 위험 또는 경고 신호는 건강한 세포가 정상적으로 생리적 죽음에 이르는 세포에서는 발생하지 않는다. 즉 병원균의 외부성이 면역 반응을 일으키는 가장 중요한 특징이 아닌 점이다.

Matzinger는 세포들의 비 자연적인 죽음은 그림 2와 같이 그 세포 주변의 작은 지역(Danger Zone)에 경고 신호를 발생시킨다고 가정하였다. 발생된 신호는 오직 위험 지역안의 항원 제시 세포만이 위험 신호를 받을 수 있다. 이때 항원 제시 세포는 면역 세포를 활성화시킨다. 그림 2와 같이 위험 지역 내에서 면역세포의 활성화가 이루어지며 위험 지역 밖에서는 항원을 인식하더라도 면역세포는 활성화되지 않는다.

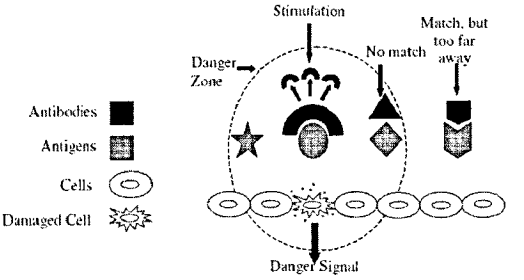

그림 2. 위험 이론 모델
Fig. 2. Danger theory model

Ⅲ. 인공면역체계 기반 공격 탐지 및 대응 모듈

3.1 공격 탐지 및 대응 모듈 설계

인간면역체계의 선천적 면역과 후천적 면역 원리를 이용한 공격 탐지 및 대응 모듈을 제안한다. 제안하는 모듈에서 면역 원리 적용 과정은 다음과 같다.

첫째, 선천적 면역 과정에서 세포들의 위험 신호들을 인지하고 위험 신호의 심각성이 결정되었을 때 T 세포의 활성화를 유도하는 위험 이론을 네트워크의 공격 징후 인지와 최종 공격 상태 결정에 이용한다. 이와 같은 원리의 적용은 동적인 네트워크 상태 변화 때문에 정상 상태와 공격 상태 동안 시스템의 상태가 유사하게 나타날 수 있으므로 네트워크에 위험을 미친다고 판단될 때만 공격 상태로 결정하도록 하기 위함이다.

둘째, 인체의 T 세포는 골수에서 생성되고 흉선에서 성숙과정을 거치게 된다. 이때 자기 세포와는 결합하지 않고 항원에만 결합할 수 있는 수용체를 가진 T 세포만이 살아남게 된다. 이 면역 원리를 부정 선택 원리라 한다. 이를 이용하여 네트워크의 비정상적인 상태 데이터로 구성된 검출기 집합을 생성한다.

셋째, 특정 항원에 대한 수용체를 가진 T 세포는 항원을 인식하면 독성 T 세포를 활성화시켜 항원을 제거하거나 보조 T 세포를 이용하여 항체를 가진 B 세포를 활성화시키는 후천적 면역 활동을 수행한다. 이러한 원리를 적용하여 공격 상태가 결정되면 공격의 근원지를 역추적하고 공격 트래픽을 제어할 수 있도록 한다.

그림 3은 앞서 제시한 원리들을 적용한 공격 탐지 모듈, 공격 대응 모듈, 검출기 집합으로 구성된 인공면역체계 기반의 공격 탐지 및 대응 모듈을 보인다.

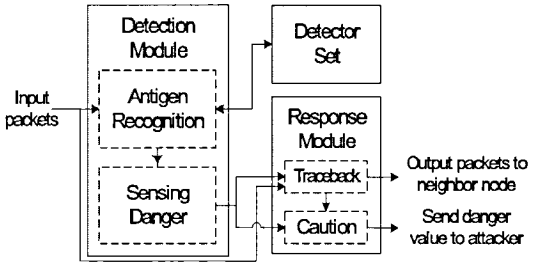

그림 3. 인공면역체계 기반의 공격 탐지 및 대응 모듈
Fig. 3. Attack Detection and Response Modules based on AIS

제안된 모듈의 기능에 따라 공격을 탐지하고 공격에 대응하는 과정을 그림 4에 보였다. 각 노드에서는 먼저 공격 상태 후보 데이터로 구성된 검출기 집합을 생성하여 유지하고 그림 4의 순서도에 따라 공격 탐지 및 대응을 수행한다. 검출기 집합은 시스템의 공격 상태 후보 데이터들의 집합으로써 주기적으로 수집한 시스템의 상태와 비교된다. 만일 현재 시스템 상태가 검출기 데이터와 동일한 상태로 결정되면 공격 상태로 판단되므로 특정 항원을 인식하는 T 세포와 같은 기능을 수행한다. 공격 탐지 모듈은 주기적으로 수집한 시스템의 상태를 수집하고 검출기 집합과의 비교 결과를 분석하여 시스템의 위험도를 계산함으로써 최

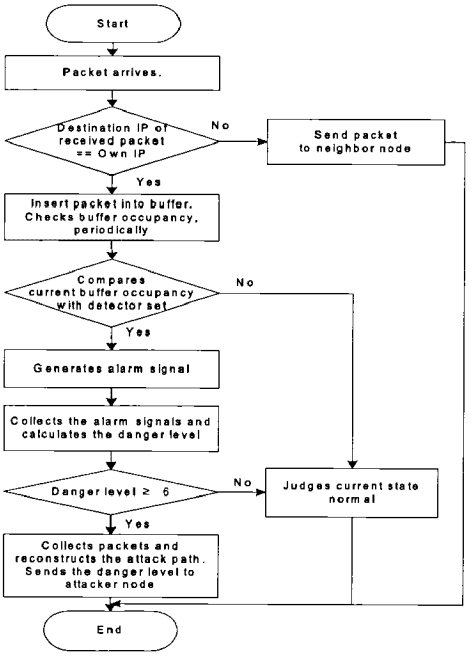

그림 4. DDoS 공격 탐지 및 대응 순서도
Fig. 4. Flowchart of Attack Detection and Response

종적으로 공격 여부를 판단한다. 공격 대응 모듈은 공격이 판단되었을 때 공격의 근원지를 찾아 공격 트래픽을 제어하도록 하는 기능을 수행한다.

3.2 검출기 집합 생성

3.2.1 유전자 결정 및 표현 방법

본 논문에서는 네트워크의 정상 상태를 자기로 정의하고 공격 상태를 비자기로 정의하였고 네트워크 상태를 반영할 수 있는 단위 유전자로 버퍼 점유율을 선택하였다. 정상 상태와 공격 발생 시의 버퍼 상태 변화를 0.1초 간격으로 측정하여 버퍼 점유율을 계산하고 10개의 유전자를 모아 시스템 상태를 나타내는 인자로 이용하였다. 주기 내에서 각 버퍼 점유율은 버퍼를 8개의 구역으로 나누어 해당 구역에 해당하는 인덱스 값을 이용하였다. 인자의 자기/비자기 여부는 10개의 유전자의 조합으로 결정되고 결정 주기는 1초 단위로 하였다. 버퍼 점유율의 유전자는 3비트로 표현되고 시스템 상태 인자는 10개의 유전자의 조합으로 그림 5와 같이 총 30 비트 길이의 이진수로 표현된다.

자기 집합 S는 식(1)과 같이 정의된다.

$$S=\{S_1,S_2,\cdots,S_n\},\quad S_i=s_{29}s_{28}\cdots s_0 \tag{1}$$

이에 따라 버퍼 점유율 시퀀스는 그림 5와 같은 형태의 이진수로 표현되고 자기 데이터 수집을 위하여 1,000초씩 반복 실험을 수행하였다. 이와 함께 네트워크의 공격 상태로 정의된 비자기 집합 Ag는 식(2)과 같이 정의된다.

$$\begin{aligned} Ag&=\{Ag_1,Ag_2,\cdots,Ag_m\},\\ Ag_i&=a_{29}a_{28}\cdots a_0 \end{aligned} \tag{2}$$

이러한 데이터 수집과 이진 표현 과정은 항원 인식 모듈의 분석기 모듈에 의해 이루어지며, 생성된 이진 데이터는 매칭 모듈에 전달된다.

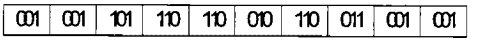

그림 5. self/non-self 데이터 표현 방법
Fig. 5. Data representation of self/non-self

3.2.2 매칭 규칙 결정 및 검출기 집합 생성

본 논문에서는 참고문헌 [7]에서 사용된 두 이진 데이터 사이의 r-contiguous 매칭 규칙을 이용한다. r-contiguous 매칭 방법은 그림 6과 같이 정의된다.

$x = x_1 x_2 \ldots x_n$ and detector $d = d_1 d_2 \ldots d_n$

d *matches* $x \equiv \exists_i \leq n-r+1$ *such that* $x_j = d_j$ for $j = 1, \ldots, i+r-1$

그림 6. r-contiguous 매칭
Fig. 6. r-contiguous matching method

본 논문에서는 하나의 유전자를 3비트로 표현하므로 두 데이터의 일치 여부 검사 방법은 유전자 단위로 이루어져야 한다. 3개의 연속적인 유전자가 일치하는 경우에 두 데이터가 일치하는 것으로 간주한다. 그림 7은 30 비트로 구성된 두 개의 버퍼 점유율 시퀀스 데이터를 보였다. 이때 9-contiguous 매칭 규칙을 수행하면 두 데이터는 3개의 유전자가 연속적으로 같으므로 일치하는 것으로 간주된다. 검출기 집합의 각 데이터는 공격 상태 후보 데이터들로 이루어지고 self 데이터들과 동일한 형식을 가지므로 검출기 집합 D는 식(3)과 같이 정의한다.

$$D = \{D_1, D_2, \ldots, D_m\}, \quad D_i = d_{29} d_{28} \ldots d_0 \qquad (3)$$

선택된 데이터 표현 방법과 매칭 방법을 이용한 부정 선택 기반의 검출기 집합 생성 과정은 초기 검출기 집합 생성 과정과 학습 과정을 거치게 된다. 초기 검출기 생성 과정은 그림 8과 같이 self 데이터 집합과 임의로 발생된 검출기 후보 데이터와의 매칭 검사를 수행한다. 매칭 검사 결과 self 데이터와 검출기 후보 데이터가 매칭되는 경우 발생된 검출기 후보 데이터

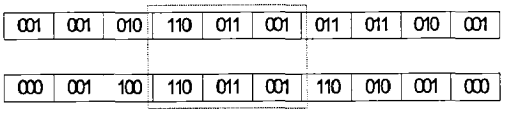

그림 7. 두 상태 파라미터의 9-contiguous 매칭 예
Fig. 7. Example of 9-contiguous matching

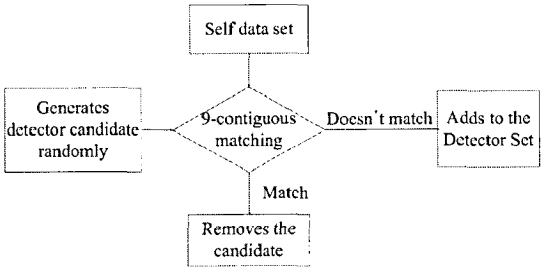

그림 8. 검출기 집합 생성 메커니즘
Fig. 8. Detector Set Generation

는 정상 상태와 동일한 것으로 간주되기 때문에 검출기 집합에 추가하지 않는다. 반면 매칭되지 않는 경우는 비정상 상태를 의미하므로 검출기 집합에 추가한다. 본 논문에서는 1,000개의 데이터를 가진 자기 데이터 집합을 이용하여 그림 8의 과정을 반복 수행함으로써 300개의 비정상 상태 후보 데이터를 가지는 검출기 집합을 생성하였다. 또한 제한된 자기 데이터 집합으로 생성된 초기 검출기 집합은 정확성이 낮으므로 새로운 정상 데이터에 반복 적용하는 학습 과정을 거침으로써 보다 정확한 검출기 집합이 생성되도록 하였다. 이렇게 생성된 검출기 집합은 인간면역체계에서 각기 다른 항원에 대한 수용체를 가지는 T 세포들의 모임과 같다.

3.3 공격 탐지 모듈 설계

공격 탐지 모듈은 시스템 상태 인자를 수집하여 공격 징후를 인지하는 항원 인식모듈과 경고 신호들을 수집하여 위험도를 분석하는 위험 감지 모듈로 구성된다.

3.3.1 항원 인식 모듈

항원 인식 모듈은 네트워크 상태를 감시하여 경고 신호를 발생시키고 위험 감지 모듈에게 전송함으로써 네트워크 상태의 위험도를 계산할 수 있는 기초 데이터를 제시하는 기능을 수행하며, 그림 9와 같이 분석기 모듈과 매칭 모듈로 구성된다.

분석기 모듈은 시스템 상태 인자를 주기적으로 추출하여 매칭 모듈에 전송하는 기능을 수행하는데 본 논문에서는 버퍼 점유율 시퀀스를 주기적으로 전송한다. 시스템 상태 인자를 수신한 매칭 모듈은 미리 생성되어 있던 검출기 집합과 비교하여 일치하는 데이터가 있는 경우 공격 징후로 인식하여 경고 신호를 발생시키다. 경고 신호 발생 메커니즘은 그림 10과 같다. 각 라우터는 버퍼 점유율 시퀀스를 주기적으로 계산하고 검출기 집합과 비교한다. 계산된 인자 값이 검출기 집합과 일치하면 비정상 상태이므로 경고 신호를 발생한다. 만일 검출기 집합 중 일치하는 값이 없

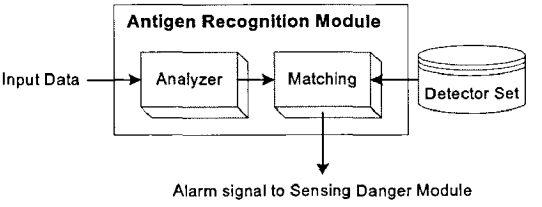

그림 9. 항원 인식 모듈의 구성도
Fig. 9. Block diagram of Antigen Recognition Module

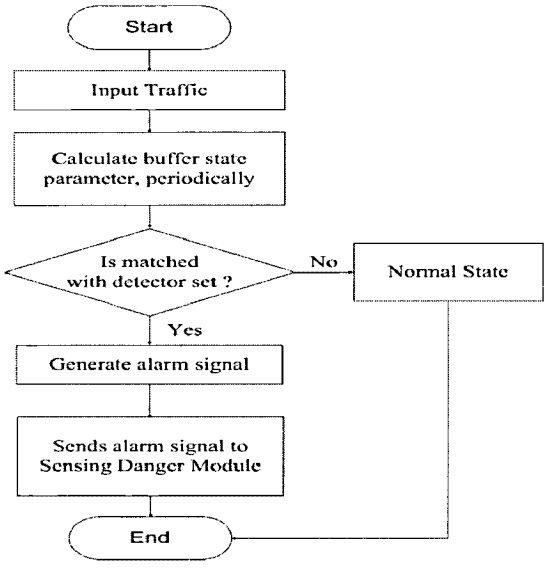

그림 10. 항원 인식 모듈의 경고 신호 발생 메커니즘
Fig. 10. Alarm signal generation mechanism

다면 정상 상태로서 부가적인 동작이 발생하지 않는다.

3.3.2 위험 감지 모듈

위험 감지 모듈은 그림 11과 같이 세 가지 기능 모듈로 구성하였으며 항원 인식 모듈에서 전달되는 경고 신호들을 주기적으로 수집하고 네트워크의 위험도를 계산함으로써 공격 상태 여부를 최종적으로 결정한다. 이는 위험 이론을 적용한 것으로 일시적인 네트워크의 상태를 기반으로 공격 상태 여부를 결정 내리기보다는 동적으로 변화하는 네트워크의 상태를 고려하기 위한 것이다.

분석기 모듈에서는 경고 신호 발생의 심각성을 측정하기 위한 위험도 측정 인자를 계산하며 경고 신호의 빈도와 경고 신호의 집중도를 사용한다. 제안한 탐지 기법에서는 동적으로 변화하는 네트워크의 환경을 고려하기 위하여 퍼지 이론을 적용하므로 두 인자 값을 퍼지 수로 변환하고 두 퍼지 수를 통합하는 기능을 퍼지 모듈에서 수행한다. 최종적으로 퍼지 모듈은 통합된 퍼지 수를 9개의 상태 집합과 유사도 검사를 수행하여 가장 유사도가 높은 상태 값으로 최종 위험도

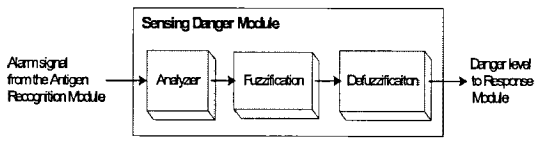

그림 11. 위험 감지 모듈의 구성도
Fig. 11. Block diagram of Sensing Danger Module

값을 결정하여 공격 대응 모듈에 전송하는 기능을 수행한다.

3.3.2.1 위험도 측정 인자 정의

제안한 위험 감지 모듈에서는 정상 상태를 공격 상태로 오판하거나 공격 상태를 정상 상태로 오판하는 오탐지를 감소시키기 위하여 일정 감시 주기 동안의 경고 신호 발생 회수와 집중도를 감시하여 최종 위험도를 결정한다. 위험이론의 위험 영역을 시간 영역으로 설정하여 위험 파라미터 값을 슬라이딩 윈도우 방식으로 일정 주기 동안 계산함으로써 위험 인자의 영향 범위를 제한하였다.

위험도 측정을 위한 파라미터 th_1, th_2을 식(4), 식(5)과 같이 정의한다.

$$th_1 = \sqrt{\frac{N_a}{M_a}} \tag{4}$$

$$th_2 = 1 - (I_{\max} / M_i) \tag{5}$$

N_a: 주기 동안 경고 신호 발생 수
M_a: 주기 동안 발생할 수 있는 최대 경고 신호 수
T_i : 주기 동안 i번째 경고 신호의 발생 시간
$I_{\max}$: 경고 신호 발생 시간 간격 중 최댓값
M_i : 주기 동안 발생할 수 있는 최대 경고 신호 발생 시간 간격

식(4)와 같이 첫 번째 매개변수 th_1은 감시 주기 동안 발생되는 경고 신호 발생 횟수에 비례한다. 두 번째 매개 변수 th_2는 경고 신호 발생의 집중도를 나타내는 것으로 발생된 경고 신호의 집중도가 높을수록 th_2의 값은 커지게 되어 위험도가 증가하게 된다.

3.3.2.2 퍼지 이론을 이용한 위험도 결정

계산된 위험도 측정 인자를 통합하여 최종 이상 상태를 결정한다. 이때 기존의 임계값 기반 정책들을 사용하면 네트워크의 동적인 상황에서 오탐지 발생 확률이 상대적으로 높다. 따라서 본 논문에서는 이상 상태 결정에 퍼지 이론을 적용하였다. 이를 위하여 위험도 측정 인자를 퍼지 수로 변환하고 두 퍼지 수를 통합하여 최종 위험도 퍼지 수를 얻는 과정을 수행한다. 먼저 각 인자의 위험도를 그림 12와 같이 저, 중, 고수준 등 3 수준으로 분류하여 0과 1사이의 위험도 값을 가지는 위험도 측정 인자로 퍼지 수 변환과정이 수행

된다[8]. 퍼지 수 변환은 각 위험도 값 x와 그에 의해 얻어지는 소속(membership) 값에 따라 이루어진다. 그림 12와 같이 위험도 값 x는 저, 중, 고 영역 중 2개의 영역에 속해 있는 경우가 발생한다. 이때 어느 영역에 속하게 될 것인지는 확률 p에 의해 영역이 설정되면 그에 해당되는 소속 값을 이용하여 위험도 값 x에 해당하는 사다리꼴 표현의 퍼지 수로 변환된다.

그림 13에서는 사다리꼴로 표현한 두 퍼지 수 $A=(a_1,a_2,a_3,a_4;w_A)$와 $B=(b_1,b_2,b_3,b_4;w_B)$의 예를 보였다. 여기서 a_1부터 a_4까지는 사다리꼴을 정의하는 꼭짓점이고 w_A는 소속 값을 나타낸다.

그림 13에서 퍼지 수 A는 계산된 위험도 값이 중수준의 소속이 0.6으로 계산된 경우이고 퍼지 수 B는 고수준의 소속이 0.25로 계산된 경우이다. 따라서 $w_A=0.6$이고 $w_B=0.25$인 경우는 *A=(0.2, 0.39, 0.61, 0.8;0.6), B=(0.5, 0.58, 1.0, 1.0;0.25)*로 표현될 수 있다. 이와 같은 표현 방법을 이용하여 위험 측정인자 th_1과 th_2을 퍼지 수 $T_1=(a_1,a_2,a_3,a_4;w_{T_1})$과 $T_2=(b_1,b_2,b_3,b_4;w_{T_2})$로 변환하고 두 퍼지 수를 하나의 위험도 값 C로 통합하여 최종 위험도 값을 얻는다. 통합된 퍼지 수를 $C=(c_1,c_2,c_3,c_4;w_C)$로 정의하면 식(6)과 식(7)에 의해 각 꼭짓점과 소속 값을 구할 수 있다[8].

$$c_i=(a_i * w_{T_1}+b_i * w_{T_2})/(w_{T_1}+w_{T_2}) \quad \text{for } i=1,2,3,4 \tag{6}$$

$$w_C=\min(w_{T_1},w_{T_2}) \tag{7}$$

퍼지 수 C를 공격 대응 모듈이 인식할 수 있는 절대적 숫자 표현으로 변환하여 자신의 위험 심각성을 전달할 수 있어야 한다. 이를 위하여 일반화된 9개의 단어 집합으로 구성된 퍼지 수 집합과 비교하여 절대적인 위험도를 결정한다. 표 1에 정의한 집합을 이용하였으며 최종적으로 통합된 위험도 값인 퍼지 수 C와 9개의 퍼지 수들을 각각 비교하여 유사성이 가장 높은 퍼지 수를 선택하고 해당하는 단어를 절대적인 위험도 값으로 결정한다.

유사도란 일반화된 퍼지 수 쌍에 대한 유사성 또는 부합 정도를 나타내기 위해 사용되는데 그 값이 높아질수록 두 수의 유사도가 높다. 일반화된 퍼지 수들의 무게 중심과 유사성을 계산하기 위한 "simple center of gravity method (SCGM)"를 이용하였다.[9] 퍼지 수 A에 대한 무게중심(Center of Gravity: COG) (x_A,y_A)는 식(8)과 식(9)에 의해 계산된다.

$$x_A=\frac{y_A(a_3+a_2)+(a_4+a_1)(w_A-y_A)}{2w_A} \tag{8}$$

$$y_A=\frac{w_A\left(\frac{a_3-a_2}{a_4-a_1}+2\right)}{6} \tag{9}$$

퍼지 수 B에 대한 COG (x_B,y_B)도 동일한 방법으

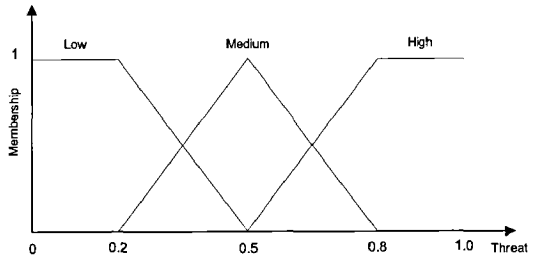

그림 12. 세 가지 상태 표현을 위한 퍼지 집합
Fig. 12. Fuzzy set for representing three state

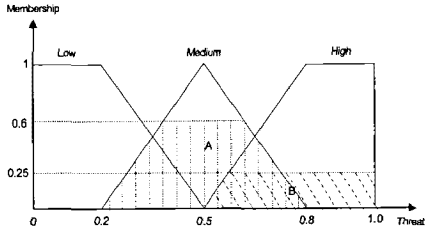

그림 13. 두 퍼지 수(A, B)의 사다리꼴 표현
Fig. 13. Two trapezoidal fuzzy number, A and B

표 1. 9가지 퍼지 단어 집합
Table 1. The nine linguistic fuzzy set

Linguistic Terms	Generalized Fuzzy Numbers
Absolutely-low(0)	(0.0, 0.0, 0.0, 0.0 ; 1.0)
Very-low(1)	(0.0, 0.0, 0.02, 0.07 ; 1.0)
Low(2)	(0.04, 0.1, 0.18, 0.23 ; 1.0)
Moderately-low(3)	(0.17, 0.22, 0.36, 0.42 ; 1.0)
Medium(4)	(0.32, 0.41, 0.58, 0.65 ; 1.0)
Moderately-high(5)	(0.58, 0.63, 0.80, 0.86 ; 1.0)
High(6)	(0.72, 0.78, 0.92, 0.97 ; 1.0)
Very-High(7)	(0.93, 0.98, 1.0, 1.0 ; 1.0)
Absolutely-high(8)	(1.0, 1.0, 1.0, 1.0 ; 1.0)

로 계산되며 이때 두 퍼지 수 사이의 유사도 $S(A,B)$는 식(10)에 의해 계산되어진다. 이 때 $S(A,B)$의 값이 클수록 두 퍼지 수 사이의 유사도가 높아진다.

$$S(A,B)=\left[1-\frac{\sum_{i=1}^{4}|a_i-b_i|}{4}\right]\times(1-|x_A-x_B|)\times\frac{\min(y_A,y_B)}{\max(y_A,y_B)} \quad (10)$$

예를 들어 두 퍼지 수를 통합한 퍼지 수 C가 (0.36, 0.51, 0.84, 0.86; 0.26)인 경우 식 (8)과 식(9)에 의해서 C의 무게 중심은 (0.64, 0.11)로 계산되어진다. 이 때 최종 위험도 값을 결정하기 위해서는 9개의 퍼지 단어 집합에서 정의한 각각의 퍼지 수와 통합 퍼지 수 C의 유사도를 모두 구해야 하므로 퍼지 수 각각에 대한 무게중심을 구하는 과정을 거치게 되며 구해진 유사도 값은 표 2와 같다. 본 논문에서는 유사도 결과 값이 6 이상인 경우 위험 상태로 간주하여 공격 대응 모듈에 신호를 전송하고 공격 트래픽 제어가 이루어지도록 한다.

표 2. 9가지 퍼지 단어 집합과의 유사도 계산 결과
Table 2. Similarity of nine trapezoidal fuzzy numbers

Linguistic Terms	COG	S(A,B)
Absolutely-low(0)	(0.00, 0.09)	0.101
Very-low(1)	(0.03, 0.10)	0.132
Low(2)	(0.14, 0.10)	0.235
Moderately-low(3)	(0.29, 0.11)	0.421
Medium(4)	(0.49, 0.11)	0.713
Moderately-high(5)	(0.72, 0.11)	0.813
High(6)	(0.85, 0.11)	0.625
Very-High(7)	(0.97, 0.10)	0.398
Absolutely-high(8)	(1.00, 0.13)	0.348

3.4 공격 대응 모듈

공격 대응 모듈은 현재 공격을 발생시키고 있는 공격의 근원지 노드를 찾아 네트워크로의 트래픽 유입을 차단하도록 하는데 그 목적이 있으며 그림 3에서 제시한 바와 같이 역추적 모듈과 경고 모듈로 구성된다. 역추적 모듈은 최종 위험도가 6이상으로 결정되었을 때 공격의 근원지를 역추적하는 기능을 수행한다. 경고 모듈의 공격 대응 메커니즘은 그림 14에 보였다. 경고 모듈은 두 가지 경우에 위험도 값을 수신하게 되므로 먼저 어느 노드로부터 수신한 위험도 값인지를 확인하여야 한다. 자신의 위험 감지 모듈로부터 수신한 경우는 위험도 값이 6 이상인지를 확인하여 공격 상태라고 판단되면 역추적 모듈에게 공격 라우터의 IP 주소를 요청하여 공격 라우터에게 위험도 값을 전송함으로써 공격 트래픽을 제어하도록 한다. 또한 수신

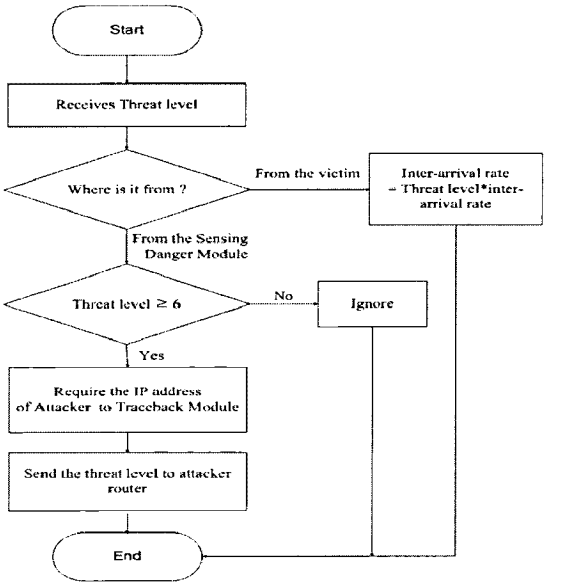

그림 14. Caution 모듈의 공격 트래픽 제어 메커니즘
Fig. 14. Attack response mechanism of Caution Module

한 위험도 값이 다른 라우터로부터 도착하였다면 자신 노드가 다른 노드에 과도한 트래픽을 전송하고 있는 것이므로 자신의 패킷 도착률을 조절함으로써 공격 트래픽을 제한하는 역할을 수행한다.

Ⅳ. 실험 및 성능 평가

제안한 기법의 성능 평가를 위해 OPNET을 이용하여 NSFNET 네트워크 모델을 설계하였다. 네트워크 모델은 14개의 노드와 하나의 제어 노드로 구성하였다. 제어노드는 노드를 초기화하고 최소 홉 기반으로 경로를 결정한다. 성능 평가를 위한 플러딩 공격 발생 경로는 그림 15와 같다.

공격은 총 120초 동안 20초부터 40초, 60초부터

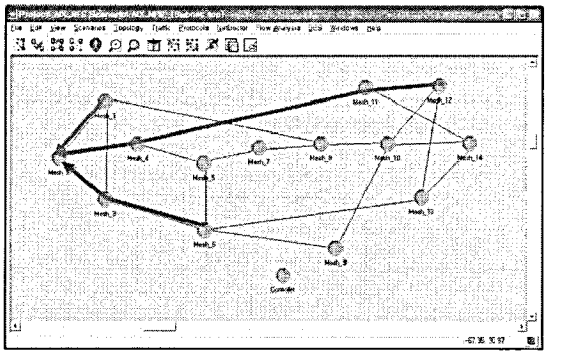

그림 15. 플러딩 공격 발생 경로
Fig. 15. Paths of Flooding Attack

100초 사이에 발생하며 노드 Mesh_12, Mesh_1, Mesh_6에서 동시에 다량의 패킷이 노드 Mesh_2로 전송된다. 제안한 공격 탐지 모듈의 성능을 평가하기 위하여 공격 탐지율, 탐지 정확성 그리고 오탐지율을 성능 평가 요소로 결정하고 TP, FP, FN, TN을 정의하였다. TP(True Positive)는 공격이 발생한 시간에 공격으로 탐지된 시간이고, FP(False Positive)는 정상 상태를 공격으로 잘못 감지한 시간이다. FN(False Negative)은 공격이 발생한 시간에 공격으로 감지하지 못한 시간이고, TN(True Negative)은 정상 상태를 정상으로 감지한 시간이다. 따라서 공격 탐지율, 탐지 정확도, 오탐지율은 다음과 같이 계산된다.

$$\text{공격 탐지율} = \frac{TP}{TP+FN} \quad (11)$$

$$\text{탐지 정확률} = \frac{TP}{TP+FP} \quad (12)$$

$$\text{오탐지율} = \frac{FP}{TN+FP} \quad (13)$$

표 3은 기존의 부정 선택 기반 탐지 기법, 임계값 기반의 탐지 기법과 제안한 탐지 기법에 의해 탐지된 TP, FP, FN, TN의 결과를 보였다. 임계값 기반의 탐지 기법에서는 네트워크의 정상 상태 시에 수집한 버퍼 점유율의 평균값을 임계값으로 결정하였고 실험 결과 4.74로 계산되었다. 따라서 임계값 기반의 탐지 기법에서는 1초 동안 수집된 10개의 버퍼 점유율 값의 평균을 계산하여 임계값보다 큰 경우 공격 상태로 판단하였다. 부정 선택 기반의 탐지 기법은 항원 인식 모듈에서 경고 신호가 발생하면 검출기 집합과 일치하는 경우이므로 공격 상태로 판단하였다.

그림 16은 각 기법의 공격 탐지율, 탐지 정확률, 오탐지율의 결과를 보였다. 제안한 공격 탐지 기법의 경우 공격 탐지율이 평균 97.5%로 기존의 탐지 기법들과 유사한 성능을 보였으나 정확률은 평균 95%, 오탐지율은 평균 5.5%로 성능이 기존의 탐지 기법들보다 우수함을 확인할 수 있었다. 그림 17은 위험도 값을 기반으로 공격 트래픽 제어를 수행했을 때 공격 대상 노드 Mesh_2의 평균 버퍼 패킷 손실률의 변화를 보였다. 공격 트래픽이 발생하는 동안에는 노드 Mesh_2 버퍼의 오버플로로 인하여 패킷 서비스를 거부한 평균 패킷 손실률은 정상 상태보다 22~32배 정도의 증가를 보였다. 그러나 제어를 수행한 후에는 정상 상태와 유사한 상태로 전환됨을 확인할 수 있다.

표 3. 평가 요소별 성능 비교
Table 3. Performance comparison of detection schemes

탐지기법 / 평가 요소	NS -based	Threshold -based	Proposed
FP	16s (21.1%)	14s (18.9%)	1s (1.6%)
TP	60s (78.9%)	60s (81.1%)	60s (98.4%)
FN	2s (4.5%)	2s (4.3%)	1s (1.9%)
TN	42s (95.5%)	44s (95.7%)	53s (98.1%)

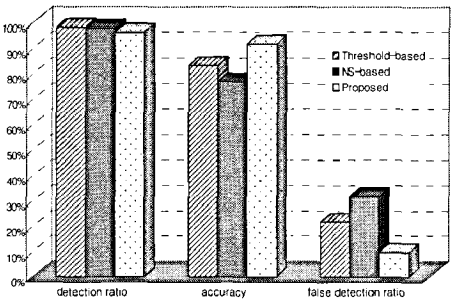

그림 16. 공격 탐지 기법들의 성능 비교
Fig. 16. Performance Comparison of Detection Schemes

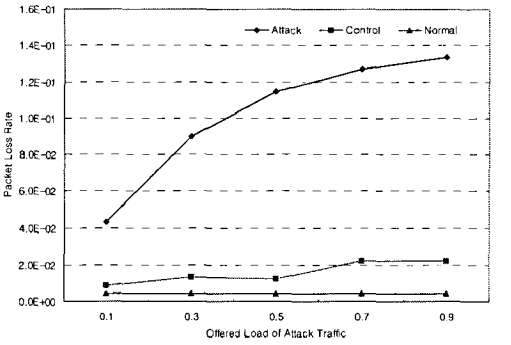

그림 17. 노드 Mesh_2 버퍼의 평균 패킷 손실률
Fig. 17. Buffer loss ratio of Mesh_2 node

V. 결 론

본 논문에서는 네트워크 공격을 효율적으로 탐지하면서 오탐지율을 줄이기 위한 인공면역체계 기반의 공격 탐지 및 대응 기법을 제안하였다. 먼저 버퍼 점유율 시퀀스를 상태 인자로 결정하고 부정 선택 알고리즘을 이용한 검출기 집합을 생성하였다. 생성한 검출기 집합은 실험에서 평균 97% 이상의 높은 공격 상

태 탐지율을 보였다. 실험 결과 부정 선택 기반의 검출기 집합만을 이용하여 공격을 탐지하는 기존의 자기/비자기 구별 기반의 공격 탐지 기법보다 제안한 탐지 기법이 오탐지율이 낮을 뿐만 아니라 정확률이 높게 나타났다. 오탐지율 측면에서는 제안한 기법은 공격 부하가 0.9일 때 3%의 오탐지율을 보여 기존의 부정 선택 기반의 탐지 기법이 22%, 임계값 기반의 탐지 기법이 21%를 보인 것에 비해 좋은 성능을 보였다. 정확률 측면에서도 부정 선택 기반의 탐지 기법이 82%, 임계값 기반의 탐지 기법이 83%를 보인데 비해 제안한 탐지 기법은 92%로 높은 정확도를 보였다. 제안한 공격 대응 기법은 다양한 공격 부하를 발생시킨 실험에서 패킷 손실률이 공격 트래픽 제어를 수행하였을 때 제어 이전보다 평균 6.7배 감소하여 효과적인 공격 제어가 가능함을 보였다.

참 고 문 헌

[1] Dasgupta D., "Advances in Artificial Immune System," IEEE Computational Intelligence Magazine, Vol.1, pp.40-49, Nov. 2006.

[2] M. S. Abadeh, J. Habibi, M. Daneshi, M. Jalali and M. Khezrzadeh, "Intrusion Detection using a Hybridization of Evolutionary Fuzzy Systems and Artificial immune Systems," Proc. of CEC 2007, pp.3547-3553, Sept. 2007.

[3] Li Zhi-tang, Li Yao and Wang Li, "A Novel Fuzzy Anomaly Detection Algorithm based on Artificial Immune System," Proc. of HPCASIA '05, pp.5-9, Nov. 2005.

[4] H. Groux, N. Fournier, and F. Cottrez, "Role of Dendritic cells in the generation of regulatory T cells", Seminars in Immunology, Vol.16, No.2, pp.99-106, 2004.

[5] Aickelin. U and Cayzer. S, "The Danger Theory and Its Application to Artificial Immune Systems," Proc. of ICAARIS2002, pp.141-148, 2002.

[6] Zhen Yu Zhou, JianJing Shen, and XinPeng Zhang, "A Danger Theory Inspired Multi-agent Fusion Model for Network Security Assessment," Proc. of the ICC2007, Vol.3, pp.599-603, Aug. 2007.

[7] S. Forrest, A. perelson, L. Allen, and R. Cherukuri, "Self-nonself discrimination in a computer," Proc. of the IEEE Symposium on Research in Security and Privacy, pp.202-212, 1994.

[8] Y. Yu and J. Graham, "Threat Evaluation for Intrusion Detection based upon Fuzzy Number Theory," Proc. of the Symposium on Information Assurance: Intrusion Detection and Prevention, pp.81-87, June 2006.

[9] S. J. Chen and S. M. Chen, "A new simple center-of-gravity method for handling the fuzzy ranking and the defuzzification problems," Proc. of the 8th National Conference Fuzzy Theory Application, pp.103-110, 2000.

유 경 민 (Kyung-Min Yoo) 정회원
한국통신학회 논문지 제 35권 제 1호 참조

양 원 혁 (Won-Hyuk Yang) 정회원
한국통신학회 논문지 제 33권 제 8호 참조
현재 전북대학교 컴퓨터공학과 박사과정

김 영 천 (Young-Chon Kim) 종신회원
한국통신학회 논문지 제 33권 제8호 참조
현재 전북대학교 IT정보공학부 교수