논문 12-37B-11-01

한국통신학회논문지 '12-11 Vol.37B No.11
http://dx.doi.org/10.7840/kics.2012.37B.11.965

효율적인 네트워크 자원 관리를 위한 호스트의 접속 유형 판별에 관한 연구

허 민◆, 김 명 섭°

Research on the Identification of Network Access Type of End-Hosts for Effective Network Management

Min Hur◆, Myung-Sup Kim°

요 약

최근 스마트 디바이스 사용이 대중화 되고 업무 환경 또한 PC중심에서 스마트 디바이스로 확대되어 가면서 무선 트래픽 양이 급격하게 증가 되고 있다. Enterprise 네트워크 에서 무선 IP 대역과 무선 트래픽 대역폭 관리는 중요한 사항이 되고 있다. 네트워크 설계 시 단말 호스트들의 네트워크 접속 유형 판별은 효율적인 네트워크 설계와 관리에 큰 이점이 된다. 또한 판별된 단말 호스트의 지속적인 관리를 통해 효과적인 네트워크 운용이 가능하고, 접속 유형의 변화를 통해 NAT 사용 호스트를 판별 할 수 있다. 본 논문에서는 Enterprise 네트워크의 인터넷 접속점에서 수집된 트래픽의 RTT(Round-Trip-Time) 값을 이용하여 단말 호스트의 접속 유형을 판별하는 방법론을 제안한다. 또한, 실제 학내 망을 대상으로 제안하는 방법의 타당성을 증명한다.

Key Words : identification of connection type, connection type of terminal host, round-trip-time, network management, 접속 유형 판별, 단말 호스트의 접속 유형, 네트워크 관리

ABSTRACT

As the use of smart devices has become popular, the number of smart devices connected to network has increased and the amount of traffic from them has grown rapidly. The management of mobile traffic and IP address for smart devices in an enterprise network is crucial problem for efficient operation of network. The information about connection type of a terminal host to the network will be very useful for stable and efficient management of an enterprise network. Also, this information might be used to identify NAT device. In this paper, we propose a methodology to identify the connection type of a terminal host using RTT (Round-Trip-Time) value extracted from captured packets. We prove the feasibility of our proposed method in a target campus network.

Ⅰ. 서 론

최근 다양한 스마트 디바이스의 등장과 사용이 대중화 되면서 다양한 어플리케이션이 등장했다. 그로 인해 네트워크에서 무선 트래픽이 차지하는 비율은 갈수록 증가 하고 있다[1]. 급증하는 무선 트래픽의 관

※ 본 논문은 2012년 정부(교육과학기술부)의 재원으로 한국연구재단(2012R1A1A2007483)의 지원을 받아 수행된 연구임
◆ 주저자 : 고려대학교 컴퓨터정보학과 네트워크 관리 연구실, grrrhm@korea.ac.kr, 준회원
° 교신저자 : 고려대학교 컴퓨터정보학과 네트워크 관리 연구실, tmskim@korea.ac.kr, 종신회원
논문번호 : KICS2012-03-125, 접수일자 : 2012년 3월 19일, 최종논문접수일자 : 2012년 11월 7일

리는 효과적인 네트워크 자원 관리에 필수적인 사항이 되고 있다. 이러한 네트워크 관리는 Enterprise Network 에서 더욱 더 중요한 사항이다[2].

급증하고 있는 무선 기술 및 단말의 증가로 인해 네트워크 자원 관리에 대한 연구는 최근 활발하게 진행되고 있다. 중첩되어 존재하는 다양한 이기종 네트워크들의 자원을 하나의 자원과 같이 통합적으로 관리하는 통합자원관리 방안[3], 예약 기반의 네트워크 자원 할당으로 인한 네트워크 자원 관리 방안[4], 무선망의 다양한 기술을 적용한 네트워크 자원 관리 방안[5] 등이 있다. 본 논문에서는 효율적인 네트워크 자원 관리를 위하여 네트워크에 접속 하는 다양한 단말 호스트의 접속 유형을 판별하는 방법을 제시한다. 네트워크 접속유형 판별 결과는 초기 네트워크 대역폭 설계 및 대역폭 증설에 필요한 기초 정보로 활용될 수 있다.

Enterprise Network에서 무선 디바이스에 대한 효과적인 설계와 증설을 위해서는 단말 호스트의 접속 유형이 판별 되어야 한다[6]. 단말 호스트의 접속 유형이 많은 곳에 다른 접속 유형보다 많은 IP대역과 대역폭을 설정 할 수 있기 때문이다. 또한 이미 설계된 Enterprise Network에서 IP 대역과 대역폭의 재설정은 사용자에게 보다 안정적이고 효율적인 네트워크 사용을 제공할 수 있다. 따라서 본 논문에서 제안하는 호스트의 접속 유형 판별은 Enterprise Network 에서 네트워크 설계와 관리를 위한 유용한 정보를 제공할 것이다.

본 논문에서는 TCP 프로토콜에서 네트워크 연결설정 과정에서 발생하는 3-way handshaking에 SYN/ACK, ACK Packet을 이용하여 RTT값을 얻어내고, 이를 이용하여 단말 호스트의 접속 유형을 판별한다[7]. 선행 연구에서는 RTT의 최소값을 이용하여 호스트의 접속 유형을 판별했다[8]. 그러나 최소값을 이용한 판별 방법은 무선 호스트에 대한 분석률 및 정확도가 낮은 문제를 갖고 있었다. 따라서 본 논문에서는 이러한 문제를 극복하기 위해 RTT의 분포를 이용하여 판별하고, 네트워크 환경에 맞는 적절한 Threshold을 정하여 접속 유형 판별의 정확도 및 분석률을 향상 시키는 연구를 진행하였다.

본 논문은 다음과 같이 구성된다. 2장은 네트워크 자원 관리에 대한 관련 연구를 기술하며, 3장은 시스템의 흐름과 본 논문에서 제안하는 방법을 제시한다. 4장은 실험 환경과 분석 결과를 기술하며, 5장에서는 결론 및 향후 연구과제 대하여 기술한다.

Ⅱ. 관련연구

네트워크 자원 관리는 사용자에게 보다 효율적인 네트워크 환경을 제공하고, 관리자에게는 네트워크 증설 및 관리에서 발생하는 경제적인 문제를 해결 가능하게 한다. 효과적인 네트워크 자원 관리는 사용자에게 인터넷 속도 향상과 패킷 손실률을 줄임으로써 안정적인 네트워크 사용을 가능하게 해준다. 또한 네트워크 자원 부족으로 발생하는 사용자의 연결 끊김 현상을 방지할 수 있다. 지속적인 네트워크 자원 관리는 관리자에게 무분별한 네트워크 자원 증설에 사용되는 비용을 절감할 수 있고, 네트워크 사용자에 대한 서비스 불만을 감소시킬 수 있다. 따라서 최근 스마트 디바이스에 등장으로 무선 트래픽 분석에 대한 연구와 무선 자원에 대한 연구가 활발하게 진행되고 있다.

논문[3]은 중첩되어 존재하는 다양한 이기종 네트워크들의 자원을 하나의 자원과 같이 통합적으로 관리하는 통합자원관리(Common Radio Resource Management: CRRM) 방안을 제안하였다. 하지만 해당 방법론은 무선 접속기술의 무선 네트워크 환경에 중점을 두고 연구를 하였다. 따라서 유무선이 혼재하는 전체 네트워크 자원 관리에는 적용하기 어려운 문제가 있다.

논문[4]에서는 통합자원관리시스템을 개발하여 네트워크 자원의 예약 및 할당 가능한 네트워크를 개발하는 방법론을 제안하였다. 사용자와 라우터 간에 인터페이스 메시지를 통해 네트워크 자원을 예약하고 할당하는 테스트베드 네트워크를 개발하였다. 하지만 해당 방법론은 현재 개발 단계에 있으며 사용자가 많은 대용량 고속 네트워크에는 적용하기 어려운 문제점이 있고, 많은 사용자로부터 발생하는 네트워크 접속 Delay 문제점이 있다.

이 외에도 논문[5]에서 네트워크 자원 관리에 대해서 연구를 진행하고 있다. 하지만 진행되고 있는 연구들은 무선 트래픽에 대한 네트워크 자원 관리에 초점이 맞춰져 있다. 따라서 본 논문에서는 무선 트래픽의 증가로 인한 전체 네트워크에 자원 관리를 위하여 유무선 사용자의 수를 파악하여 보다 효율적이고 안정적인 네트워크 제공하기 위한 방법론을 제안한다.

본 논문에서는 TCP 연결 과정에서 발생하는 3-way handshake 패킷에 Round-Trip-Time을 이용하여 전체 네트워크에서 사용하는 단말의 접속 유형을 판별하는 방법을 제시한다. 판별된 정보를 이용하여 네트워크 자원 관리에 활용한다.

Ⅲ. 유무선 판별 방법

본 장에서는 RTT(Round-Trip-Time)를 이용하여 단말 호스트의 네트워크 접속 유형을 판별하는 방법을 기술한다. RTT 설정 방법과 시스템에 판별을 통한 각각의 단말 호스트에 상태 변화에 대해 기술한다. 또한, 본 논문에서 제안하는 시스템 알고리즘 및 다양한 네트워크 환경에 적용하기 위한Threshold설정에 대해 기술한다.

3.1. Round-Trip-Time

TCP 프로토콜을 사용하는 네트워크는 3-way handshaking 통해 연결 설정이 이뤄진다. 본 논문에서 제안하는 시스템은 연결 설정 과정에서 발생하는 SYN, SYN/ACK, ACK packet의 수집 시간을 이용한다. 정확한 분석을 위해 유무선 구간인 Client-Router 사이에서 발생하는 SYN/ACK, ACK packet의 수집 시간을 이용하여 RTT 값을 구한다. 본 연구에서는 유무선 접속유형의 대상은 Enterprise Network에 포함된 종단 호스트이고, Packet을 수집하는 지점은 Router 이다.

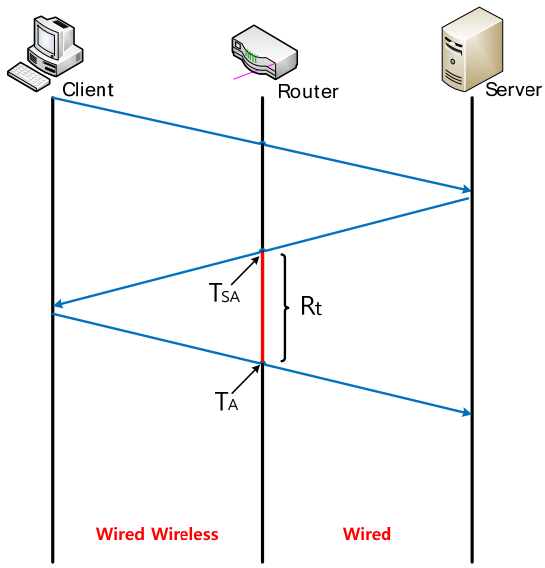

그림 1. TCP 3-way handshake & RTT

그림 1은 Client-Server 간에 연결 설정 과정에서 사용되는 3-way handshaking을 통한 RTT(R_t)값을 얻는 방법을 표현하고 있다. 본 논문에서는 수식(1)을 이용하여 값을 구한다.

$$R_t = T_A - T_{SA} \quad (1)$$

R_t는 응답시간이며, T_A는 Router에서 ACK Packet이 수집된 시간이며, T_{SA}는 Router에서 SYN/ACK Packet이 수집된 시간이다. R_t는 호스트에서 발생되는 모든 Flow를 대상으로 한다.

Enterprise Network에서 무선 접속은 AP을 통해 연결한다. 이 과정에서 무선 연결은 주위에 건물이나 장애물 등으로 인해 연결이 지연된다.[9] 따라서 유선의 RTT값은 무선의 RTT값보다 상대적으로 낮은 값을 가질 것이다. 또한, 같은 호스트에서 발생하는 Flow의 RTT값은 변화는 적을 것이고 비슷한 값을 갖고 있는 분포를 나타낼 것이다. 따라서, 본 논문에서는 접속 유형에 따른 RTT분포의 차이를 이용하여 단말 호스트의 접속 유형을 판별한다.

3.2. State Transition

시스템의 분석 대상은 1분 단위로 수집한 트래픽에서 Flow을 발생시킨 모든 호스트를 대상으로 한다. State Transition은 접속유형판별 과정에서 하나의 호스트 상태가 다양한 입력에 따라 변화하는 동작을 나타낸 것으로 그림 2와 같이 표현된다. 호스트의 상태는 시스템 분석을 통해 판별된 유선, 무선 상태 그리고 판별되지 않은 NULL 상태의 3가지 상태로 표현된다. 호스트의 상태는 1분 데이터가 수집이 완료된 후에 시스템 조건과 결과에 따라 변화한다.

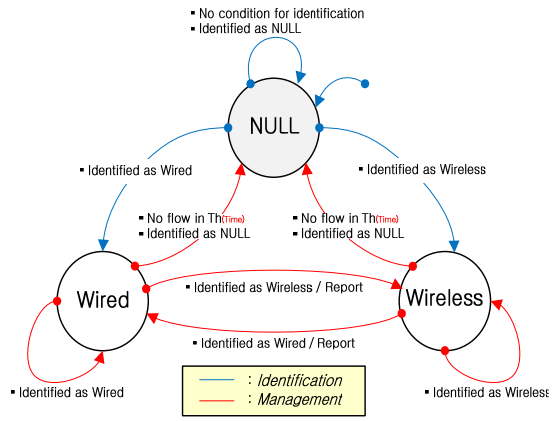

그림 2. State Transition

제안하는 시스템은 판별, 관리 모듈로 구성된다. 각 모듈에 따라 호스트 상태가 변화하는 조건은 각각 다르다. 판별 모듈은 호스트가 최초 상태에서 시스템에 의해 판별될 경우에 동작하며, 관리 모듈은 시스템에 의해 판별된 호스트 상태에서 동작한다.

분석 호스트의 최초 상태는 NULL이다. 분석 시스

템은 최초 판별을 위하여 최소 $Th_{(Time)}$시간 동안 Flow 수집이 필요한 조건이 있다. 따라서 이 조건에 만족하지 못한 호스트는 계속해서 NULL 상태를 유지한다. 또한, 시스템은 판별을 위해서 $Th_{(Flow)}$의 개수가 필요한 조건이 있다. 호스트에서 수집 시간 동안 모아진 Flow의 개수가 조건에 만족하지 못할 경우 계속해서 NULL 상태를 유지한다. 앞에서 언급한 2가지의 조건을 만족시킨 호스트는 본 논문에서 제안하는 알고리즘에 의해 판별 한다. 시스템 판별 결과는 유선 상태, 무선 상태, NULL 상태로 판별된다. 판별 결과가 NULL 상태인 경우는 수집된 Flow의 비율이 유선인지 무선인지 모르는 판별 불가능한 경우이다. 이 경우에는 지속적으로 Flow을 수집하여 특정 접속 유형으로 판별될 때까지 NULL 상태를 유지한다. 유선 상태 및 무선 상태로 판별될 경우에는 해당 상태로 변화하는 동작을 한다.

시스템은 지속적으로 호스트에 접속 유형을 관찰하며 관리하게 된다. 따라서 호스트가 판별된 상태에서는 시스템 관리 모듈을 통해 호스트에 상태가 변화 한다. 시스템 관리 모듈에서는 호스트 상태에 대해 지속적인 관찰을 위해 분 단위로 계속 판별 한다. 시스템에 의해 판별된 상태 이후에 1분 동안 발생한 Flow와 기존에 수집된 Flow와 같이 판별 한다. 하지만 호스트에 접속 유형 변화를 관찰하기 위해서 가장 최근 발생한 Flow부터 $Th_{(Flow)}$개수만큼 판별 한다. 예를 들어, 호스트에 상태가 판별된 이후에 1분 동안 20개 Flow가 발생하게 되면 호스트에서 발생시킨 최초에 20개 Flow을 제외하고 판별 한다. 시스템 판별 결과는 판별 모듈과 같이 유선 상태, 무선 상태, NULL 상태로 판별 된다. 해당 접속 유형 상태에서 판별 결과가 같으면 계속해서 상태를 유지하며, 같지 않으면 다른 접속 유형 상태로 변화한다. 이 경우는 하나의 호스트에서 2가지 접속 유형이 관찰됐기 때문에 NAT을 사용하는 호스트로 의심할 수 있다. 따라서 해당 호스트를 관리자에게 Report 한다. 또한, 호스트에 상태가 판별된 이후에 $Th_{(Time)}$시간 동안 한 개의 Flow도 발생하지 않을 경우 NULL 상태로 변화한다. 호스트에서 더 이상에 Flow가 발생하지 않는 것은 네트워크 사용이 중단된 것으로 예상되기 때문이다. 따라서 해당 호스트는 다시 최초 판별을 위한 조건을 만족시킬 때까지 NULL 상태로 유지된다.

3.3. 시스템 알고리즘

본 논문에서 제안하는 분석 시스템은 1분 단위 트래픽을 수집하여 TCP 연결과정에서 발생하는 3-way handshake 패킷을 이용하여 판별한다. 유선과 무선의 패킷 전송 시간의 차이를 이용하기 위해 RTT 값을 추출한다. 1개의 호스트에서 같은 접속 유형으로 트래픽을 발생시킬 경우에는 RTT 값의 분포가 균일하게 발생할 것이다. 따라서, 유선과 무선의 RTT 값의 분포 차이를 이용하여 유선 호스트와 무선 호스트를 판별한다. 그림 3은 학내 망에 있는 한 개의 유선과 무선 호스트의 RTT 값의 분포를 나타낸다. 분석 시스템은 그림 3의 결과를 바탕으로 시스템 알고리즘을 설계하였다.

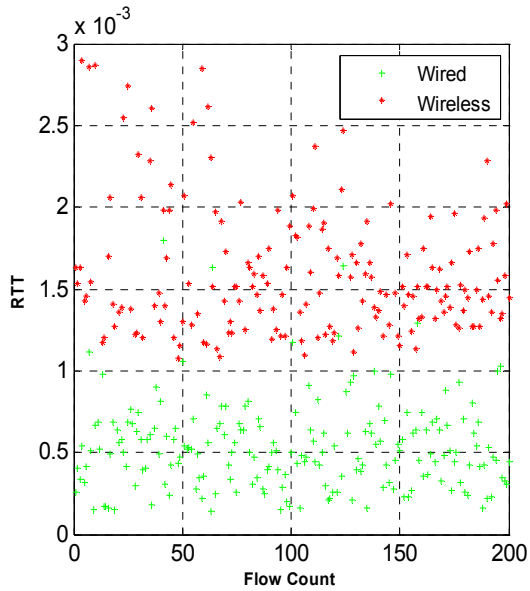

그림 3. RTT Distribution of wired, wireless

유선의 RTT 분포는 무선의 RTT 분포보다 비교적 낮게 형성되고 있으며, 무선에 비해 분포의 범위가 적은 것을 확인할 수 있다. 이는 유선 호스트의 네트워크 사용이 비교적 안정적으로 사용되는 것으로 판단할 수 있다. 반면, 무선의 RTT 분포는 유선에 비해 분포의 범위가 산발적인 것을 확인할 수 있다. 이는 무선 호스트의 네트워크 사용이 다양한 장애요소에 방해되는 것으로 판단할 수 있다.

무선 RTT 분포는 대부분 1.5×10^{-3}초에 나타나고 있으며, 유선 RTT 분포는 0.5×10^{-3}초에 나타나고 있다. 따라서, 본 논문에서는 유선과 무선의 분류하기 위한 $Th_{(RTT)}$값을 1.0×10^{-3}초로 정의하였다. 또한, 학내 망에 존재하는 호스트에서 발생하는 유선과 무선의 RTT 값은 대부분 비슷한 분포를 나타내고 있다. 그러나, 같은 호스트에서 발생시킨 RTT 중에 1.0×10^{-1}초 이상 값을 갖는 RTT을 발견하였다. 보

다 정확한 분석이 필요하겠지만 재전송 패킷으로 의심된다. 분석 시스템에서는 판별의 정확성을 높이기 위해 1.0×10^{-1}초 이상의 RTT 값은 제외하였다.

단말 호스트의 접속 유형 판별 시스템은 크게 2개 모듈로 구성된다. 판별 모듈은 n분 동안 호스트에서 발생하는 m개의 Flow를 유선과 무선을 판별하기 위한 기준 값과 비교하여 판별하고, 관리 모듈에서는 최초 판별 이후 지속적으로 호스트에서 발생시킨 Flow를 수집하여 판별하고, 호스트의 접속 유형 변화를 관찰하여 접속 유형의 변화가 생긴 호스트를 관리자에게 Report 한다.

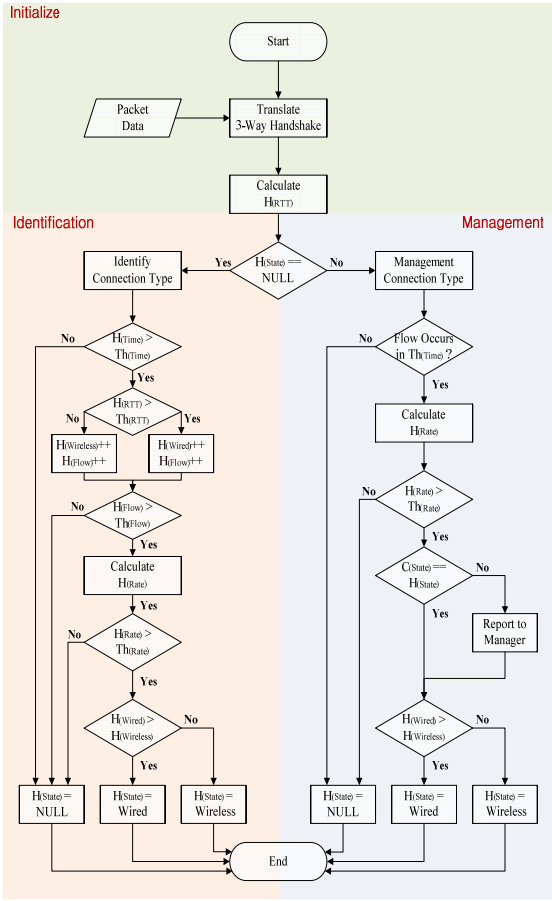

그림 4. Algorithm

그림 4는 분석 시스템의 전체 알고리즘 순서도를 표현한다. 분석 시스템은 학내 망에서 발생시킨 모든 패킷 을 1분 단위로 수집하여 판별하고, 분석 대상은 1분 동안 트래픽을 발생시킨 모든 종단 호스트이다. 따라서, 분석 시스템의 알고리즘은 1분 동안 모든 트래픽을 수집한 후에 개별 호스트를 대상으로 수행한다. 본 논문에서 $H_{(\ )}$는 하나의 호스트를 의미하며, $Th_{(\ )}$는 접속 유형 판별을 위한 Threshold 값을 의미한다. $C_{(\ )}$는 시스템 현재 상태를 의미한다.

시스템 초기 단계에서는 수집된 Packet Data에서 TCP 연결 과정에서 발생하는 3-way handshake 패킷만 추출하여 handshake 파일로 변환한다. 변환된 파일은 수식(1)을 이용하여 각각의 호스트 별로 $H_{(RTT)}$ 값을 구한다.

● 판별(Identification) 모듈

판별 모듈은 $H_{(State)}$가 NULL 상태일 경우에 판별한다. 호스트의 최초 판별을 위해서 $Th_{(Time)}$동안 Flow을 수집한다. 따라서 $H_{(Time)}$ > $Th_{(Time)}$조건을 만족시킬 때까지 호스트는 NULL 상태를 유지한다. $H_{(Time)}$은 최초에는 Flow가 발생한 시점을 기록하며, 다음 시스템 알고리즘 시작 시점에 1분씩 증가한다.

$Th_{(Time)}$동안 Flow 수집이 완료되면 수집된 Flow의 $H_{(RTT)}$값과 $Th_{(RTT)}$을 비교한다. $H_{(RTT)}$값이 $Th_{(RTT)}$값보다 크면 $H_{(Wireless)}$(무선 유형)을 증가 시키고 낮으면 $H_{(Wired)}$(유선 유형)을 증가 시킨다. 또한, 전체 Flow의 개수를 확인하기 위해 동일하게 $H_{(Flow)}$를 증가시킨다. $H_{(Flow)}$는 $Th_{(Time)}$시간 동안 수집된 전체 Flow의 개수를 나타낸다. 시스템은 $H_{(Flow)}$ > $Th_{(Flow)}$조건을 비교하고, 조건을 만족하지 못한 호스트는 NULL 상태를 유지한다. $H_{(Flow)}$가 낮을 경우에는 적은 Flow을 이용하여 판별 할 경우 판별 결과의 신뢰성을 낮추기 때문에 $H_{(State)}$를 NULL로 유지한다. 조건에 만족한 호스트는 접속 유형 비율을 나타내는 $H_{(Rate)}$를 계산한다. $H_{(Rate)}$ > $Th_{(Rate)}$ 조건을 비교하고, 조건에 만족하지 않은 호스트는 판별 비율이 낮기 때문에 더 많은 Flow 수집을 위하여 $H_{(State)}$를 NULL로 유지한다. 조건에 만족한 호스트는 접속 유형이 유무선 판별을 위해 $H_{(Wired)}$ > $H_{(Wireless)}$조건을 비교하여 호스트의 접속 유형을 판별한다. 조건에 만족한 호스트는 $H_{(State)}$ = Wired로 변경하고, 만족하지 못한 호스트는 $H_{(State)}$ = Wireless로 변경하고 판별 모듈을 종료한다.

● 관리(Management) 모듈

관리 모듈에서는 $H_{(State)}$가 NULL이 아닌 호스트에 대해서 수행한다. NULL가 아닌 호스트는 판별 모듈을 통해서 접속 유형이 판별된 상태이다. 관리 모듈에서는 판별된 호스트에서 더 이상의 Flow가 발생되지 않을 경우 NULL 상태로 변경하고, 지속적으로 수집된 Flow을 이용하여 접속 유형의 변화를 관찰하는 역할을 수행한다.

시스템은 Flow Occurs in Th(Time)조건을 통하여

호스트에서 $Th_{(Time)}$동안 Flow 발생이 있었는지 확인한다. $Th_{(Time)}$동안 1개의 Flow도 발생하지 않을 경우에는 네트워크 사용이 중단된 것으로 판단하고 호스트의 $H_{(State)}$를 NULL 로 변경한다. 조건에 만족한 호스트는 기존 판별 모듈을 통해 수집된 $H_{(Flow)}$와 판별 모듈 이후에 발생한 Flow를 이용하여 $H_{(Rate)}$를 계산한다. 그러나, 호스트에서 계속적으로 Flow가 발생하게 되면 수집된 Flow가 많아지게 된다. 이는 시스템의 성능 저하와 판별 결과의 정확성 및 접속 유형 변화를 관찰할 수 없는 문제점이 발생한다. 따라서 관리 모듈에서는 1분 동안 발생한 Flow 개수만큼 과거 수집된 Flow을 제외하고 판별한다. 계산된 $H_{(Rate)}$는 판별 모듈과 같이 $Th_{(Rate)}$을 비교한다. 조건에 만족하지 못한 호스트는 $H_{(State)}$를 NULL로 변경한다. 조건에 만족한 호스트는 관리 모듈에 판별 결과인 $C_{(State)}$와 $H_{(State)}$가 같은지 비교한다. 비교 결과가 같지 않을 경우 호스트의 접속 유형이 변화된 것이다. 이 경우는 하나의 호스트에서 2개의 접속 유형이 판별 된 것을 확인 할 수 있다. 해당 호스트에 대한 지속적인 연구가 필요하겠지만 NAT 사용 호스트로 의심할 수 있다. NAT 사용에 관한 보다 정확한 연구는 향후 연구를 통해 풀어야 할 과제이다. 따라서 해당 호스트 정보를 관리자에게 Report 한다. $C_{(State)}$와 $H_{(State)}$가 같은 경우 $H_{(Wired)}$와 H(Wireless)를 비교하여 기존 호스트의 상태를 유지하고 수행을 종료한다.

3.4. Threshold

Enterprise Network 환경은 각기 다르기 때문에 환경에 맞는 Threshold 값이 필요하게 된다. Threshold은 분석 시스템에 정확도와 분석률을 향상시키기 위해 네트워크 환경에 맞추어 설정한다. 다음에 나오는 4가지 요소들 분석 시스템에 미치는 영향과 역할에 대해 설명한다. 요소들은 분석 시스템을 적용하기 전에 네트워크 환경에 맞게 실험을 통해 설정해야 한다.

- $Th_{(RTT)}$

유선과 무선을 판별 하는 기준 값이다. 기준 값의 설정은 유선 값의 분포와 무선 값의 분포에 경계선이 되는 값으로 설정 한다. Enterprise Network 환경마다 통신을 방해하는 요소들이 다르기 때문에 분석 시스템에 정확도를 향상시키기 위해 설정해야 한다. 본 논문에서는 학내 망의 여러 번 실험 결과를 통해 얻은 1.0×10^{-3}초로 설정한다.

- $Th_{(Time)}$

판별에 필요한 데이터를 수집하는 시간이다. 시간 설정은 시스템의 정확한 판별을 위해 Flow를 수집하는 시간을 정의하고, 판별된 호스트에서 더 이상의 Flow가 발생하지 않을 경우 시스템 판별에 제외시키기 위해 필요한 설정 값이다. 시간 설정은 지속적인 데이터의 수집으로 발생하는 시스템 성능 저하와 메모리 낭비 문제를 해결할 수 있다. 본 논문에서는 시스템 성능 향상과 판별의 정확성을 높이기 위한 5min 값으로 설정한다.

- $Th_{(Flow)}$

판별에 필요한 Flow의 개수를 설정하는 값이다. Flow의 개수는 정확한 판별을 위해 필요한 최소한의 값이다. 이 값을 설정 하지 않을 경우에는 많은 데이터를 수집할 경우 메모리 낭비 문제를 갖게 되고 적은 데이터를 수집할 경우에는 정확한 판별을 할 수 없는 문제가 생기게 된다. 1분당 발생하는 Flow의 평균 숫자를 기준으로 값을 정한다. 본 논문에서는 1분당 40개 Flow의 평균 발생으로 계산하여 5min 동안 수집을 위한 200개로 설정한다.

- $Th_{(Rate)}$

하나의 단말 호스트에서 판별된 접속 유형 비율을 설정하는 값이다. 판별 비율은 분석 시스템에 전체 분석률에 영향을 주게 된다. 설정 값이 높아지면 판별된 호스트의 정확도가 올라가지만 판별 되지 않는 호스트가 많아 진다. 또한 설정 값이 낮아지면 판별 되는 호스트가 많지만 판별된 호스트의 정확도가 낮아지게 된다. 본 논문에서는 시스템 정확성 향상에 중점을 두고 설정한다. 여러 번의 실험을 통해 적지 않은 호스트를 판별하고 시스템 판별 결과의 정확성이 높은 75%로 설정한다.

Ⅳ. 실험 및 성능 평가

본 절에서는 단말 호스트의 네트워크 접속 유형 판별을 위해 구성한 시스템 환경 및 실험 결과에 대해 기술한다. 또한, 제안하는 방법론을 학내 망의 실제 트래픽의 적용해서 그 타당성을 증명하고, 실험 결과를 분석하여 네트워크 자원 관리에 활용 가능한 정보에 대해서 기술한다.

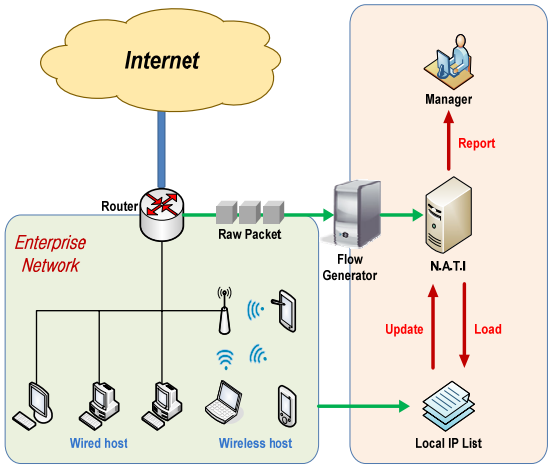

그림 5. The Experimental Environment

실험 환경 구성은 그림 5와 같이 다양한 유선과 무선이 공존하는 학내 네트워크를 대상으로 한다. 패킷은 외부 인터넷이 연결되어 있는 최상위 Router에서 미러링 받아 분석 시스템에 전달하는 형태로 구성하였다. 본 논문에서 제안하는 분석 시스템은 N.A.T.I(Network Access Type Identifier)로 정의한다. 분석 시스템은 수집된 패킷을 이용하여 학내 네트워크에 존재하는 호스트의 접속 유형을 판별한다. N.A.T.I 시스템은 학내 네트워크의 Local IP List를 통해 트래픽을 발생시킨 호스트의 정보를 추출한다. 각각의 호스트에서 발생하는 패킷 정보를 이용하여 RTT값을 추출하여 접속 유형을 판별한다. 판별된 결과는 호스트의 접속 유형 관찰을 위해 Local IP List에 Update 한다. 또한, 지속적인 호스트의 접속 유형 관찰하여, 변화가 생긴 해당 호스트의 정보를 관리자에게 Report 한다.

표1은 실험에 사용된 트래픽 트래이스와 실험 기간을 나타낸다. 트래이스는 학내 망에서 사용하는 3,000여대의 호스트에서 사용하는 다양한 종류의 응용 프로그램의 트래픽으로 구성된다. 실험 기간은 매달 2째주에 일주일 데이터를 5개월간 수집하여 실험하였다.

표 1. Traffic Trace

Date	# of flows	# of packets	Bytes
2011-10-10 ~ 2011-10-16	342M	53,974T	87,433TB
2011-11-07 ~ 2011-11-13	366M	57,927T	93,246TB
2011-12-05 ~2011-12-11	407M	64,041T	103,121TB
2012-01-09 ~ 2012-01-15	212M	33,656T	52,232TB
2012-02-06 ~2012-02-12	257M	39,221T	59,665TB

트래픽 트레이스는 학내 망의 특성에 맞게 학업 기간(10월, 11월, 12월) 데이터의 비해 방학 기간(1월, 2월) 데이터의 양이 낮은 것을 보여주고 있다.

본 논문에서 제안하는 방법론의 성능 평가를 위해 수집된 5개월 데이터를 다양한 관점으로 실험했다. 1일 동안에 학내 망 사용자의 접속 유형 판별 및 월별 사용자의 접속 유형을 실험을 통해 분석했다. 그림 6은 5개월 데이터를 실험하여 하루 동안 1분 단위로 판별된 호스트의 접속 유형 평균을 표현한다.

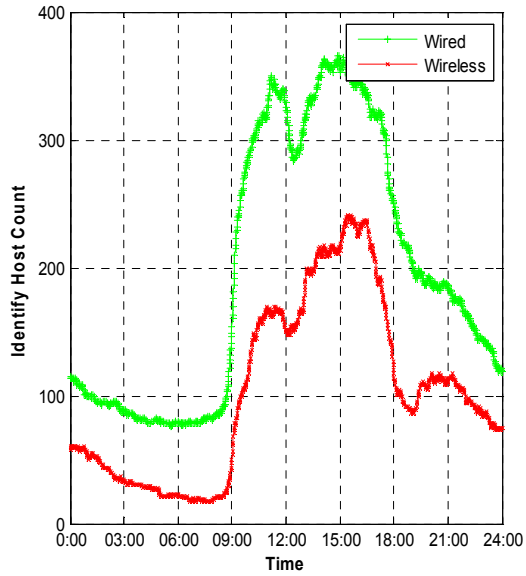

그림 6. Connection Type of Host by Time

본 논문에서는 학내 망을 대상으로 실험했기 때문에 수업이 시작하는 09시부터 판별되는 호스트의 수가 증가하고 18시 이후부터 감소하는 것을 확인 할 수 있다. 유선과 무선 사용자의 사용 패턴은 비슷하게 유지되고 있다. 또한, 피크 시간(15시)에 판별된 무선 호스트의 수는 약 230대 정도를 나타내고 있다. 따라서 사용자에게 효율적인 네트워크 사용 환경과 효과적인 네트워크 자원 관리를 위해서는 피크 시간에 1분당 최대 250대의 무선 호스트 사용자가 접속 가능해야 한다.

학내 망 특성에 맞는 효과적인 네트워크 자원 관리를 위해서 5개월 데이터를 기간 별로 분석했다. 학내 망의 특성 분석을 위해 학업 기간과 방학 기간에 호스

트 접속 유형을 분석했다. 그림 7과 8은 월별 하루 동안 1분 단위로 판별된 각각의 호스트 접속 유형 평균을 나타낸다.

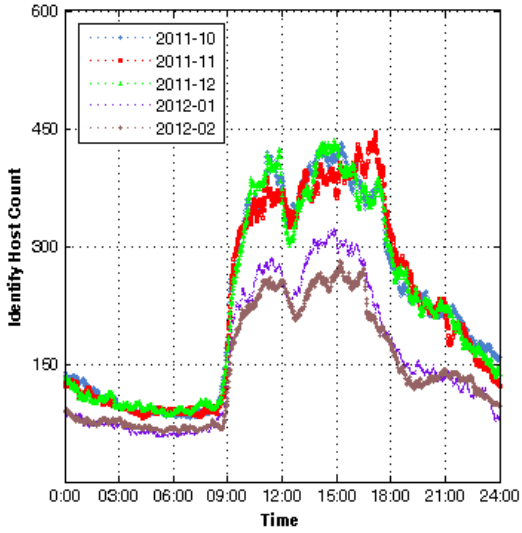

그림 7. Wired Host by month

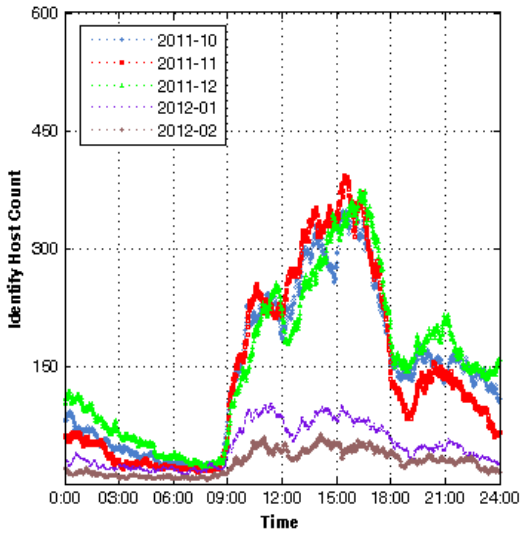

그림 8. Wireless Host by month

그림 7과 8에서 학업 기간에 유무선 호스트의 사용이 방학 기간보다 많은 것을 확인 할 수 있다. 2 기간에 판별된 유선 호스트의 차이는 최대 100대 미만이다. 그러나 무선 호스트의 차이는 피크 시간(15시)에 약 200대 차이가 나는 것을 확인 할 수 있다. 이는 유선의 경우 방학 기간에도 대학원 연구실 및 교직원 등의 사용으로 판별 결과의 차이가 작은 것으로 예상되고, 무선의 경우는 학업 기간에 학생들의 스마트 디바이스 사용이 많기 때문에 판별 결과의 차이가 큰 것으로 예상된다.

기간 별 호스트의 접속 유형 결과를 효율적인 네트워크 자원 관리에 정보 활용하기 위해 기간 별 호스트의 접속 유형 평균을 분석했다. 그림 9는 월별 호스트의 접속 유형 평균을 표현한다.

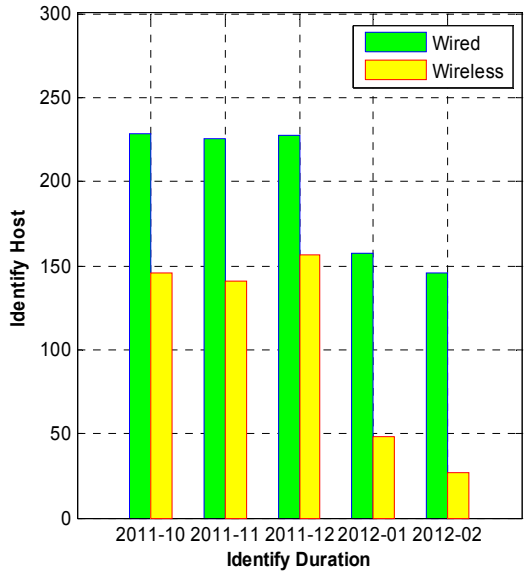

그림 9. Connection Type by Period

그림 9와 같이 방학 기간보다 학업 기간에 무선 사용자가 많은 것을 확인 할 수 있다. 평균적으로 100대 이상의 호스트 차이가 난다. 또한, 기간에 상관없이 유선과 무선의 비율은 전체적으로 6:4 정도에 비율을 유지하고 있다. 이 정보는 학내 망을 최초 설계 시에 네트워크 대역폭 설정에 유용하게 사용될 것이다. 또한, 방학 기간에는 무선에 사용이 적기 때문에 무선 대역을 줄이고 유선 대역을 확장시키는 것이 사용자에게 보다 나은 네트워크 환경을 제공할 것이다.

본 논문에서 제안하는 유무선 판별 방법의 정확성을 측정하기 위해서 사용된 정답지는 학내 망의 IP 대역을 이용하여 생성하였다.

무선 호스트의 정답지는 학내 망의 특정 IP 대역을 사용하고 있는 약 1,000대의 호스트에 대해 정답지를 생성하였고, 유선 호스트의 정답지는 컴퓨터 실습실, 학과 PC실, 전산실 등 공유기 및 무선에 사용이 불가능한 약 500대의 호스트에 대해 정답지를 생성하였다.

본 논문에서 제안하는 유무선 판별 방법의 정확한 성능 평가를 위해 정확도를 다음과 같이 정의하였다. 정확성은 학내 망의 IP 대역을 이용하여 생성한 정답을 알고 있는 호스트의 수 중에서 정확하게 접속 유형을 판별한 단말의 비율을 나타낸다.

$$Accuracy = (N_{CT} \div N_{GT}) \times 100 \qquad (2)$$

위 수식(2)는 정확성을 구하는 것으로 N_{GT}는 정답을 알고 있는 전체 호스트의 수이며, N_{CT}는 정확하게 시스템을 통해 접속 유형을 판별한 호스트의 수이다. 다음 표2는 본 논문에서 제안하는 단말의 접속 유형 판별 방법에 대한 정확도를 나타낸다.

표 2. Accuracy

	Previous Method		Proposed Method	
	Wired	Wireless	Wired	Wireless
Accuracy	100%	78.11%	97.07%	100%

표2 는 5개월 데이터를 기존 방법과 본 논문에서 제안하는 방법으로 유무선 판별 시스템의 정확도를 나타낸다. 기존 방법론에서는 유선에 대해서는 100% 완벽하게 판별했지만, 무선 호스트에 대해서는 평균 78%의 정확도로 판별했다. 기존 방법은 최소 값을 이용하여 판별하기 때문에 무선 호스트에서 $Th_{(Rate)}$보다 낮은 Flow가 발생하게 되면 다른 Flow가 높게 발생하더라도 유선으로 판별되기 때문이다. 따라서 기존 방법은 유선에 대해서는 완벽하게 판별 가능하지만 무선 호스트에 대한 판별 신뢰성은 낮다.

본 논문에서 제안하는 방법론은 무선 호스트에 대한 분석 시스템의 정확도는 100% 완벽하게 판별했다. 반면에, 유선 호스트에 대한 분석은 약 97%의 정확도로 판별했다. 이는 특정 유선 호스트에서 높은 RTT값을 갖는 Flow을 발생시키고 있기 때문이다. 높은 RTT값을 갖는 호스트는 트래픽 발생이 많은 호스트에서 나타난다. 다음 그림 10은 미분류 되는 유선 호스트에서 발생되는 Flow의 RTT을 표현한다.

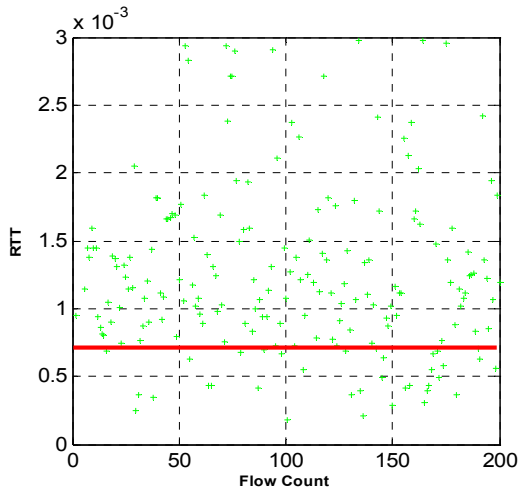

그림 10. RTT of Unidentified Wired Host

그림 10과 같이 미분류 유선 호스트의 RTT 분포는 특정 영역에 분포되지 않고 전체적으로 분포되어 있다. 따라서 유선 호스트의 RTT 분포는 붉은 색의 1.0×10^{-3} $Th_{(RTT)}$값으로 판별 불가능하다. 이러한 분포를 갖는 유선 호스트는 특정 호스트에서 가끔씩 발견된다. 본 논문에서 제안하는 방법론의 정확도를 향상시키기 위해서는 미분류 유선 호스트에 대한 보다 정확한 연구가 필요하다. 따라서 미분류 호스트에 대한 연구는 향후 연구를 통해 풀어야 할 과제이다.

V. 결론 및 향후 연구

최근 스마트 디바이스에 발전과 다양한 어플리케이션에 등장으로 네트워크 자원을 사용하는 무선 트래픽이 증가하였다. 따라서 네트워크 자원 관리에 대한 다양한 연구가 진행되고 있고, 여러 기술들을 적용하여 네트워크 자원을 효율적으로 사용자에게 제공하는 연구를 진행하고 있다. 본 논문에서는 Round-Trip-Time을 이용하여 Enterprise Network에 단말 호스트에 접속 유형 판별 방법론을 제안했다. 판별된 접속 유형 정보를 통해 네트워크 자원 관리에 다양한 정보를 제공했다. 또한 다양한 네트워크 환경에 본 논문에서 제안하는 방법을 적용하기 위해 4가지에 Threshold 값을 정의했다. 제안하는 방법론에 대해 다양한 실험을 통해 하루 동안 사용되는 접속 유형 판별 정보와 5개월 기간에 접속 유형 판별 정보를 제공했다. 끝으로, 지속적인 관리 모듈을 통해 단말 호스트의 접속 유형 변화를 관찰하고 NAT 사용 호스트를 판별 했다. 하지만 NAT 사용 호스트에 대한 정확한 분석 결과를 갖고 있지 않기 때문에 향후 연구를 통해

해결해야 할 문제이다. 또한, 분석 시스템의 정확도 향상을 위해서 미분류 호스트에 대한 연구도 해결되어야 한다.

향후 연구로는 다양한 네트워크에 적용하기 위한 Threshold 값 설정을 위한 자동으로 값을 추출하는 연구와 NAT로 판별되는 호스트에 대한 연구와 미분류 호스트에 대한 연구도 계속해서 진행 할 계획이다.

References

[1] Cisco, "Cisco visual networking Index: Global Mobile Data Traffic Forecast Update," *White Paper*, Feb. 1, 2011.

[2] David Kotz, Kobby Essien, "Analysis of a Campus-Wide Wireless Network", *WIRELESS NETWORK 11*, pp. 115-133, Jan. 2005.

[3] C. Y. Shin and J. S. Cho, "An ANP-based Resource Management Scheme in Heterogeneous Wireless Networks Considering Multiple Criteria", *KICS*, vol. 36, no. 8, pp. 910-920, Oct. 2011.

[4] H. K. Lim, J. H. Moon, J. U. Kong, J. S. Han and Y. W. Cha, "A Reservation based Network Resource Provisioning Testbed Using the Integrated Resource Management System", *KICS*, vol. 36, no. 12, pp. 1450-1458, Dec. 2011

[5] J.D. Mallapur, Syed Abidhusain, Soumya S. Vastrad, Ajaykumar C. Katageri, "Fuzzy Based Bandwidth Management for Wireless Multimedia Networks", *Communications in Computer and Information Science* vol. 70, no. 10, pp. 81-90, Mar. 2010

[6] Wei Wei, Bing Wang, Chun Zhang, Jim Kurose, Don Towsley, "Classification of Access Network Types: Ethernet, Wireless LAN, ADSL, Cable Modem or Dialup?", *INFOCOM 2005*, pp. 1060-1071, Dec. 2008.

[7] H. S. Lee, Y. S. Oh, S. W. Lee and M. S. Kim "Host information gathering using the traffic analysis", *In proceedings of Korea Information Processing Society(KIPS) 2009*, pp. 1202-1205, Apr. 2009.

[8] J. M. Park and M. S. Kim, "A Study on the Determinative Method of the Access Type of End-Host", *In proceedings of Korea Information and Communication Society(KICS)*, pp. 315, Nov. 2010.

[9] Kamal Jain, Jitendra Padhye, Venkaka N.Padmanabhan, Lili Qiu, "Impact of Interference on Multi-Hop Wireless Network Performance", *WIRELESS NETWORK 11*, pp.471-487, Jul. 2005.

허 민 (Min Hur)

2011년 고려대학교 컴퓨터정보학과 학사

2011년 3월~현재 고려대학교 컴퓨터정보학과 석사과정

<관심분야> 네트워크 관리 및 보안, 트래픽 모니터링 및 분석

김 명 섭 (Myung-Sup Kim)

1998년 포항공과대학교 전자계산학과 학사

2000년 포항공과대학교 컴퓨터공학과 석사

2004년 포항공과대학교 컴퓨터공학과 박사

2006년 Post-Doc.. Dept. of ECE, Univ. of Toronto, Canada.

2006년~현재 고려대학교 컴퓨터정보학과 조교수

<관심 분야> 네트워크 관리 및 보안, 트래픽 모니터링 및 분석, 멀티미디어 네트워크