Journal of the Korea Institute of Information and Communication Engineering

한국정보통신학회논문지 Vol. 25, No. 1: 50~55, Jan. 2021

Deep Q 학습 기반의 다중경로 시스템 경로 선택 알고리즘

정병창^{1*}·박혜숙²

Path selection algorithm for multi-path system based on deep Q learning

Byung Chang Chung^{1*} · Heasook Park²

^{1*}Senior Researcher, Electronics and Telecommunications Research Institute, Daejeon, 34129 Korea ²Principal Researcher, Electronics and Telecommunications Research Institute, Daejeon, 34129 Korea

요 약

다중경로 시스템은 유선망, LTE망, 위성망 등 다양한 망을 동시에 활용하여 데이터를 전송하는 시스템으로, 통신 망의 전송속도, 신뢰도, 보안성 등을 높이기 위해 제안되었다. 본 논문에서는 이 시스템에서 각 망의 지연시간을 보상으로 하는 강화학습 기반 경로 선택 방안을 제안하고자 한다. 기존의 강화학습 모델과는 다르게, deep Q 학습을 이용하여 망의 변화하는 환경에 즉각적으로 대응하도록 알고리즘을 설계하였다. 네트워크 환경에서는 보상 정보를 일정 지연시간이 지나야 얻을 수 있으므로 이를 보정하는 방안 또한 함께 제안하였다. 성능을 평가하기 위해, 분산 데이터 베이스와 텐서플로우 모듈 등을 포함한 테스트베드 학습 서버를 개발하였다. 시뮬레이션 결과, 제안 알고리즘이 RTT 감소 측면에서 최저 지연시간을 선택하는 방안보다 20% 가량 좋은 성능을 가지는 것을 확인하였다.

ABSTRACT

Multi-path system is a system in which utilizes various networks simultaneously. It is expected that multi-path system can enhance communication speed, reliability, security of network. In this paper, we focus on path selection in multi-path system. To select optimal path, we propose deep reinforcement learning algorithm which is rewarded by the round-trip-time (RTT) of each networks. Unlike multi-armed bandit model, deep Q learning is applied to consider rapidly changing situations. Due to the delay of RTT data, we also suggest compensation algorithm of the delayed reward. Moreover, we implement testbed learning server to evaluate the performance of proposed algorithm. The learning server contains distributed database and tensorflow module to efficiently operate deep learning algorithm. By means of simulation, we showed that the proposed algorithm has better performance than lowest RTT about 20%.

키워드: 강화학습, 지연시간 보정, 다중경로 시스템, 다중경로 TCP, 경로 선택

Keywords: Reinforcement learning, Delay compensation, Multi-path system, Multi-path TCP, Path selection

Received 30 October 2020, Revised 10 November 2020, Accepted 23 November 2020

* Corresponding Author Byung Chang Chung (E-mail:bcchung@etri.re.kr, Tel:+82-42-860-3932) Senior Researcher, Electronics and Telecommunications Research Institute, Daejeon, 34129 Korea

Open Access http://doi.org/10.6109/jkiice.2021.25.1.50

print ISSN: 2234-4772 online ISSN: 2288-4165

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(http://creativecommons.org/li-censes/by-nc/3.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © The Korea Institute of Information and Communication Engineering.

I. 서 론

최근 통신망의 신뢰도 및 속도, 보안성을 향상하려는 방법으로 다중경로 전송이 제시되고 있다. 다중경로 전송 시스템이란 말 그대로 통신 장치가 여러 개의 인터페이스를 보유하고 데이터를 전송하는 방식으로, multipath TCP[1], LTE-WiFi link aggregation 등의 방식이 대표적이다. 이렇게 다중경로를 이용해 데이터를 전송하게 되면, 데이터를 더욱 빠르게 보낼 수 있어 속도에서 우위를 가지거나, 여러 경로를 사용하기 때문에 한경로에 장애가 생겨도 바로 대체할 수 있으며, 패킷을 나누어 보내게 되어 특정 경로가 감청당해도 데이터 일부를 보호할 수 있다.



Fig. 1 Example of multi-path system

그림 1은 다중경로 시스템의 예시를 보여준다. 4개의 다른 매체를 가지는 망이 있으며, 각각 다른 품질을 가 진다. 품질의 예시로는 지연시간, 대역폭, 손실률 등을 생각할 수 있다. 품질에 따라 적절한 망을 선택하여 운 용할 수 있다.

위와 같은 구조에서는 망을 적절하게 선택하는 것이 중요하다. 예를 들어 특정 경로가 좋다고 모든 단말이 특정 경로만을 이용하게 되면, 그 망은 결과적으로 포화 하여 지연시간이 길어질 것이다. 그렇다고 모든 경로에 대해 동일한 비율로 망을 이용하게 되면, 다중경로 망이 가지는 효율성을 떨어뜨린다. 따라서 올바른 경로 선택 알고리즘을 가지는 것이 중요하다.

본 논문에서는 이러한 다중경로 시스템에서의 경로 선택 문제를 강화학습을 이용하여 풀고자 한다. Ⅱ장에 서는 강화학습에 대하여 간단히 대해서 소개하고, Ⅲ장 에서는 논문의 제안 시스템에 강화학습 기법을 어떻게 적용하는지에 관해 설명하고자 한다. Ⅳ장에서는 제안 방안의 성능 검증을 위해 제안 방안을 수행할 수 있는 학습 서버를 구축하고 시뮬레이션을 적용한 결과를 설명 한다. 마지막으로 V장에서 논문을 결론짓도록 하겠다.

Ⅱ. 강화학습과 Deep Q learning

본 논문에서는 경로 선택 문제를 deep Q 학습을 이용하여 풀고자 한다. 먼저 Q 학습[2]을 소개하면 다음과같다. 어떤 상황(s: state)에서 어떤 행동(a: action)를 취할 때 나오는 보상(r: reward)을 Q(s,a)라는 함수에 저장하여, 특정 상황(s*)에서 Q(s*,a*)를 최대화하는 특정 행동(a*)을 취하는 것이다. 이렇게 알고리즘을 구성하면시행착오를 겪으며 Q 함수가 정확해지고, 학습이 진행될수록 최적의 선택을 하게 된다.



Fig. 2 Typical Q learning algorithm

하지만 이러한 Q 학습은 상황과 행동의 숫자가 많지 않은 환경에서 높은 효과를 가진다. 상황이나 행동의 차원이 늘어나게 되면 Q 함수를 모두 업데이트하기 어려워서, 시간이 오래 걸리거나 최적의 값으로 수렴하지 않게 된다.

이런 단점을 개선하기 위해 Q 함수 테이블을 deep learning을 이용해 예측하는 것이 deep Q 학습이다. Deep Q 학습을 이용하면 상대적으로 상황이나 행동의 숫자가 증가하여도 높은 예측도를 가지고 행동을 결정할 수 있다.

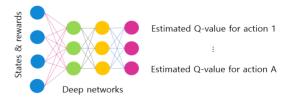


Fig. 3 Deep Q learning algorithm

그림 2와 그림 3을 통해 일반적인 Q 학습 알고리즘과 Deep Q 학습 알고리즘의 차이를 파악할 수 있다. 기본 적으로 States, Rewards, Actions의 구조는 동일하다. 하지만 일반적인 Q 학습 알고리즘에서는 Q table을 관리하며 이 테이블에 모든 상황/행동 조합에 대한 Q 값을 저장한다. 반면 deep Q 학습 알고리즘 모든 조합에 대해 Q 값을 저장하지 않고, 추론을 통해 Q 값을 예측하기 위해 [3]과 같은 일반적인 인공신경망을 이용하여 상황/행동에 따른 보상을 학습시킨다. 따라서 deep Q 학습에서는 겪어보지 않은 상태더라도 기존의 유사한 상태를 겪어봤다면 deep network의 Q값 추론을 통해 적절한 행동을 선택할 수 있다.

Ⅲ. 제안하는 경로 선택 알고리즘

위에서 언급한 deep Q 학습을 다중경로 시스템에 적용하기 위해서는 적절한 상황, 행동, 보상을 결정해야한다. 본 논문에서는 상황, 행동, 보상을 아래와 같이 정의하였다.

정의에 앞서 주요 notation을 소개하겠다. 각각의 망을 $N=\{1,2,...,N\}$ 으로, 그 망의 지연시간 수치를 d_n^t 라 정의하였다. 여기서 t는 시간의 의미이다. 따라서 d_n^{t-1} 은 바로 이전에 얻은 지연시간의 수치가 될 것이다. 논문에서는 해당 notation을 이용하여 Q 학습의 상황(state)을 다음과 같이 정의하였다.

$$s^{t} = (d_{1}^{t}, d_{1}^{t-1}, \cdots d_{1}^{t-W}, \cdots, d_{N}^{t-W})$$

$$\tag{1}$$

즉, 모든 망의 최근 W개 지연시간 정보를 포함한 벡터를 상황으로 정의하였다. 예를 들어, 망이 N=4, W=10인 경우 상황은 각 4개 망의 최근 10개 지연시간 값들로이루어진 총 길이 40의 벡터로 정의할 수 있다.

위와 같이 상황을 정의하였을 때, 일반적인 Q 학습을 사용하기 어려운 이유는 지연시간 수치 d_n '가 몇 개의 범주로 분류되지 않는 무한한 범위를 가지는 값이기 때문 이다. 게다가 망별로 W개의 입력값을 가지기 때문에 전체 상황의 개수가 얼마나 될지 추정할 수 없다. 하지만 deep Q 학습을 통해서라면 비슷한 지연시간 수치 패턴에 대해서는 유사한 Q 값을 예상하도록 학습시킴으로써 기존의 일반적인 Q 학습의 테이블 관리 문제를 해결할 수 있다.

덧붙여 multi-armed bandit 모델 기반의 일반적인 Q 학습으로 경로 선택 문제를 풀지 않는가에 대한 의문이 생길 수 있다. Multi-armed bandit 모델은 망의 지연시간이 정적이거나 일정한 확률 분포를 가진 상황에서는 효율적이다. 하지만 망의 지연시간이 동적이고, 각각 망의배경 트래픽이 시시각각 변하는 환경에서는 위와 같이지연시간의 시계열 변화를 모두 고려하여 상황을 정의하는 게 효율적일 것이다. 대표적으로 최근의 무선통신망을 생각할 수 있다. 커버리지 또는 무선 특성이 다른 망들을 상황에 맞게 이용하는 이종망 네트워크라면 위와 같이 최근 시계열 지연시간 데이터를 모두 보는 것이효과적이다.

이어서 제안 시스템의 행동(action)을 정의하고자 한다. 행동은 직관적으로 어떤 망을 선택하느냐가 될 것이다. 따라서 행동의 집합은 다음과 같이 정의한다.

$$a^{t} \in \{ n = 1, n = 2, \dots, n = N \}$$
 (2)

그리고 제안 시스템의 보상(reward)은 특정 행동에 따른 지연시간으로 계산한다. 즉, n이라는 행동을 취했을 때 얻을 수 있는 보상은 $r^l = d_n^{l+k}$, k > 0이 된다. 여기서 k는 뒤에 언급할 지연시간 보정을 위해 정의한다.

위와 같이 상황, 행동, 보상을 정의하면 deep Q 학습 방안을 적용할 수 있다. 그림 4는 실제 deep Q 학습을 적 용한 제안 방안의 블록 다이어그램을 보여준다.

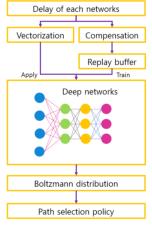


Fig. 4 Block diagram for proposed learning algorithm

Replay buffer 블록은 과적합 학습을 방지하기 위해 존재한다 [4]. 네트워크에서 발생하는 모든 값을 인공신 경망에 넣게 되면 과적합 문제가 발생할 수 있다. 과적 합이란 비슷한 데이터들을 너무 많이 학습시켰을 때 나 타나는 증상으로, 인공신경망이 학습된 데이터에만 심 하게 종속되는 경우를 말한다. 따라서 발생하는 모든 데이터를 replay buffer에 넣어두고, 모든 데이터를 학습시키는 것이 아니라 buffer에서 일정 확률로 데이터를 꺼내 인공신경망에 학습시킨다.

강화학습을 네트워크 시스템에 적용할 때의 문제점은 네트워크 시스템에서 얻을 수 있는 보상이 지연된다는 점이다. 다시 말해서, 기존의 강화학습 시스템이 상황과 행동에 따라서 얻는 보상이 즉각적이라면, 이 시스템에서는 상황과 행동에 따라서 얻는 보상이 일정 지연시간 후에 일어난다는 뜻이다. 따라서 이를 학습 알고리즘의 입력값으로 반영할 때 이 지연시간만큼을 보정해서 학습시켜야 한다. 즉, 지연시간 d_n' 를 보상으로 얻었다면, 이를 인공신경망에 입력할 때는 d_n' 시간만큼 이전의 상황과 행동과 매칭해서 넣어야 한다. 앞서 설명한보상의 정의 $r'=d_n'^{+k}$, k>0가 상황/행동-보상 간 지연시간 보정을 뜻한다. 해당 동작을 수행하는 부분이 그림 4의 Compensation 블록이다.

마지막으로 인공신경망을 통해 Q 값을 예측하면, 무조건 최적의 Q 값을 고르는 것이 아니라 최적의 Q 값을 높은 확률로 고를 수 있도록 설계한다. 무조건 최적의 Q 값을 고르지 않는 이유는 학습 알고리즘 특성상 다양한 시도를 통해 다양한 경우의 Q 값을 확보하고, 특정 경로의 상황이 변하는 것에 대처하기 위함이다. 강화학습에서는 이를 exploitation-exploration 상호관계라고 말한다. 강화학습이 효율적으로 이루어지기 위해서는 좋은행동을 다수 실행하는 것(exploitation)과 변화 감지를위해 다른 행동을 수행해보는 것(exploration)의 조화가필요하다. 본 연구에서는 Boltzmann distribution에 따라행동을 결정하였다 [5]. 여기서 $p(a_i)$ 는 행동 a_i 를 고를확률이고, c에 따라 최적의 Q 값을 고르는 확률 분포가 달라진다.

$$p(a_i) = \frac{e^{Q(s,a_i)/c}}{\sum_{k=1}^{N} e^{Q(s,a_k)/c}}$$
(3)

전체 알고리즘은 그림 4의 블록 다이어그램에 따라수행된다. 각각 네트워크로부터 지연시간 정보를 받으면 이를 기존의 상황, 행동 정보와 함께 replay buffer에 넣는다. Replay buffer에서 임의의 확률로 상황, 행동, 보상 정보를 인공신경망에 넣고, 이것으로 예상 Q 값을 학습한다. 패킷 전송 시 현재 상황을 인공신경망에 넣으면

예상 Q 값을 얻게 되고, 이를 수식 (3)에 대입하여 행동을 결정한다.

기존의 라우팅 분야에 학습 알고리즘을 적용한 연구들[6-7]과 비교하였을 때, 제안 방안은 더 높은 확장성을 가질 수 있다. 기존의 알고리즘은 시간 차이를 보정하는 부분을 반영하여도 고전적인 Q 학습만을 도입하거나, 수렴성과 효율을 위해 최단 경로만 선택지로 활용하는 등의 방식을 채택하였다. 반면, 제안 방안에서는 알고리즘을 deep network로 구성하였기 때문에 state와 action 수가 증가하더라도 알고리즘이 수렴할 확률이 높아진다.

Ⅳ. 제안 방안 개발 및 성능 분석

제안하는 알고리즘을 평가하기 위해 tensorflow, HBase, spring framework 등을 활용하여 제안하는 강화학습 알고리즘을 구동할 수 있는 학습 서버를 구축하였다. 학습 서버의 구성도는 아래 그림 5와 같다. 위 그림 4에서 언급한 학습 모듈은 그림 5의 tensorflow 모듈에 반영된다.

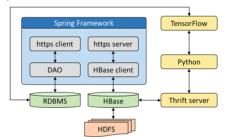


Fig. 5 Design of learning server

제안 방안에서 실제 지연시간 측정은 기지국 또는 게이트웨이 노드에서 수행되므로 범용적인 인터페이스인 REST API를 통해 지연시간을 수신하고, 데이터 용량이 커지는 상황을 고려하여 분산 데이터베이스의 일종인 HBase에 저장하도록 하였다 [8]. 제안 알고리즘에서는 데이터량이 이미지나 영상만큼 크지는 않으나, 향후 네트워크 플로우 정보 등의 다른 정보를 학습 알고리즘의 입력값으로 사용할 수 있도록 분산 데이터베이스를 이용해 저장소를 구성하였다. REST API와 HBase client 가 Java spring framework에서 동작하는 반면 tensorflow는 python 기반으로 동작한다. 따라서 HBase 에서 python으로 데이터베이스를 가져오기 위해 thrift

를 이용하였다. 마지막으로, 각각 기지국이나 게이트웨이 같은 주요 제어 정보들이 분산 데이터베이스에 저장되는 것은 비효율적이다. 따라서 이 정보들은 MariaDB에 저장되도록 설계하였고, 이것을 python으로 가져오는 것은 happybase 라이브러리를 이용해 구현하였다. 아래 그림 6은 구성한 학습 서버 내 hbase shell 구동 예시를 보여준다. 그림 6과 같이 분산 데이터베이스를 관리하면, 학습 데이터가 지속적으로 수집되어도 용량 문제를 쉽게 해결할 수 있다. 더불어, 학습 모델이 추후 개선되더라도 기존에 저장된 데이터를 이용하여 학습 알고리즘을 새로이 적용할 수 있는 확장성을 얻을 수 있다.



Fig. 6 Example operation of distributed database

위와 같이 학습 서버를 구축한 뒤, 네트워크 상황을 구성하였다. 네트워크 구조는 4개의 품질이 다른 망을모델링하였다. 해당 모델링에서는 4개의 개별 무선망을고려하였으며, 기지국의 위치, pathloss 모델에 차이를두어 특성이 달라지게 구성하였다 [9]. 덧붙여 무선의특징 이외에도 코어망에서도 각 망의 배경 트래픽에 따라 각각의 망이 다른 service rate를 가지도록 모델링하였다. 각 망의 배경 트래픽은 무선망의 성능 변화보다긴 시간 스케일로 변화하도록 설정하였다. 그 외 시뮬레이션 파라미터로 W=10, c=2.5를 사용하였고, 인공신경망은 4개의 레이어로 구성하였다. 즉, 40개의 입력값으로 시작해 최종적으로 보상값 1개를 예측하는 신경망을구현하였다.

그림 7은 시간에 따른 전체 네트워크의 평균 지연시간을 보여준다. x축은 시간으로, 1번의 iteration이 5000번의 타임슬롯을 가진다고 가정한다. y축인 평균 지연시간은 5000번의 타임슬롯 동안 각각 망이 얻은 지연시간을 모두 합쳐 평균값으로 계산하였다. 비교 방안으로는 임의로 선택하는 방안(random selection)과, MPTCP

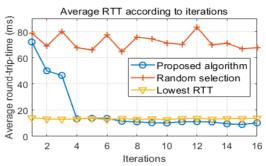


Fig. 7 Average RTT of comparable schemes

표준에서 제시한 매 순간 최소 지연시간을 선택하는 방안(lowest RTT)을 적용하였다 [10].

결과를 보면, 초반의 iteration에서는 각각 행동의 따른 망의 변화를 학습하지 못한 상태이므로 임의 선택한 방안과 유사한 결과를 가지게 된다. 하지만 iteration이 반복될수록, deep Q 학습 알고리즘은 어떤 망을 많이 선택할수록 지연시간이 단축되는지를 인지하게 되고, 점점 최적의 망 선택을 수행한다. 상대적으로 임의 선택방안의 성능 열화가 크기 때문에 그래프 상에서 제안 방안과 최소 지연시간 선택 방안의 차이가 두드러지게 나타나지는 못하지만, 제안 방안과 최소 지연시간 선택 방안의 RTT 차이를 비교하면 약 21% 정도 제안 방안이 우수한 것을 확인할 수 있다.

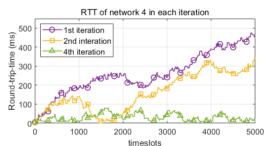


Fig. 8 RTT of particular network in each iteration

학습 알고리즘의 특징을 더욱 잘 확인하기 위해서 그림 8에는 가장 품질이 좋지 않은 망의 지연시간을 초반, 중반, 후반 iteration으로 나누어 타임슬롯에 따라 표현하였다. 초반 iteration의 경우, 학습 알고리즘이 해당 망의 성능을 잘 인지하지 못하여 상대적으로 많은 패킷을 할당하고, 이로 인해 망의 지연시간이 많이 늘어나는 결과를 보인다. 하지만 학습이 진행될수록, 학습 알고리즘은 망의 지연시간에 민감하게 반응한다. 시간이 지날수

록 지연시간이 늘어나서 포화하는 경우가 발생하지 않는 것을 볼 수 있다. 다시 말해, 제안 알고리즘은 지연시간을 보상으로 학습을 진행하였으므로 다중경로 네트워크의 지연시간을 최적화하는 방향으로 경로 선택을하는 것을 확인할 수 있다.

Ⅴ. 결 론

본 논문에서는 다중경로 시스템의 경로 선택 문제를 강화학습을 이용해서 해결하였다. 강화학습에 적용할수 있도록 지연시간 기반의 상황, 행동, 보상을 모델링하였다. 모델링 구조 상 전체 상황의 개수가 많아져 기존의 일반적인 강화학습으로는 문제를 풀기 어려우므로, 인공신경망을 이용하는 deep Q 학습 방안을 제안하였다. 덧붙여 네트워크의 특성에 맞게 replay buffer, delay compensation 등의 과정을 적용하여 전체적인 알고리즘을 설계하였다. 제안하는 알고리즘의 성능 검증을 위해 학습 데이터를 수집하고, 그로부터 학습을 수행할 수 있는 학습 서버를 개발하였다. 이 학습 서버에서 시뮬레이션을 통해 알고리즘의 성능을 검증한 결과, 해당 알고리즘을 적용하였을 때 전체적인 평균 지연시간이 감소하고 품질이 좋지 않은 경로의 포화가 줄어든 것을 확인하였다.

ACKNOWLEDGEMENT

This work was supported by the Institute for Information and Communications Technology Promotion (IITP), Korea Government (MSIT), through the Development of Adaptive Network Technology With Multi-Media Multi-Path, under Grant 2017-0-00282.

REFERENCES

[1] M. S. Kim, J. Y. Lee, and B. C. Kim, "Design of MPTCP congestion control based on BW measurement for wireless networks," *Journal of the Korea Institute of Information and Communication Engineering*, vol. 21, no. 6, pp. 1127-1136, Jun. 2017

- [2] K. Jang, "Reinforcement learning for node-disjoint path problem in wireless ad-hoc networks," *Journal of the Korea Institute of Information and Communication Engineering*, vol. 23, no. 8, pp. 1011-1017, Aug. 2019.
- [3] S. Park, B. Lim, and H. Jung, "CNN-Based Toxic Plant Identification System," *Journal of the Korea Institute of Information and Communication Engineering*, vol. 24, no. 8, pp. 993-998, Aug. 2020.
- [4] V. Mnih, K. Kavukcuoglu, D. Silver, A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, pp. 529-533, Feb. 2015.
- [5] B. C. Chung and D. H. Cho, "Semidynamic cell-clustering algorithm based on reinforcement learning in cooperative transmission system," *IEEE Systems Journal*, vol. 12, no. 4, pp. 3853-3856, Dec. 2018.
- [6] S. P. M. Choi and D. Y. Yeung, "Predictive Q-Routing: A Memory-based Reinforcement Learning Approach to Adaptive Traffic Control," in *Proceedings of Advances in Neural Information Processing Systems*, pp. 945-951, 1996.
- [7] S. Hoceini, A. Mellouk, and Y.Amirat, "K-Shortest Paths Q-Routing: A New QoS Routing Algorithm in Telecommunication Networks," in *Proceedings of International Conference on Networking* 2005, pp. 164-172, 2005.
- [8] Apache HBase, [Internet]. Available: https://hbase.apache.org/.
- [9] M. Series, "Guidelines for evaluation of radio interface technologies for IMT-Advanced," *International Telecommunication Union (ITU)*, Geneva, Switzerland, Technical Report ITU-R M.2135-1, Dec. 2009.
- [10] C. Raiciu, M. Handley, and D. Wischik, "Coupled congestion control for multipath transport protocols," RFC 6356, IETF, Oct. 2011. [Internet]. Available: https://tools.ietf.org/html/rfc6356.

20 20 20 20 20 20

정병창(Byung Chang Chung)

2011년 한국과학기술원 전기 및 전자공학과(공학사) 2013년 한국과학기술원 전기 및 전자공학과(공학석사) 2017년 한국과학기술원 전기 및 전자공학과(공학박사) 2017년 ~ 한국전자동선연구원 지능화융합연구소

국방ICT융합연구실(선임연구원) ※관심분야 : 머신러닝, 네트워크, 게임이론



박혜숙(Heasook Park)

1994년 부산대학교 전자계산학과(이학석사) 2005년 충남대학교 컴퓨터과학과(이학박사) 1992년 ~ 한국전자통신연구원 지능화융합연구소 국방ICT융합연구실(책임연구원/실장)

※관심분야: 국방 네트워크, 다중경로 네트워크