

# MDT 데이터 및 물리계층 측정값을 활용한 다층 퍼셉트론 기반 처리량 예측기법

박정균\*, 김동욱\*, 공규열\*\*, 최수용<sup>o</sup>

## A Multi Layer Perceptron Based Throughput Prediction Technique using MDT Data and Physical Layer Measurements

Junggyun Park\*, Dongwook Kim\*, Gyuyeol Kong\*\*, Sooyong Choi<sup>o</sup>

### 요약

본 논문에서는 MDT 데이터를 포함한 다양한 물리계층 측정값과 처리량과의 상관관계를 분석하고, 물리계층 측정값을 이용한 다층 퍼셉트론 기반 처리량 예측기법을 제안한다. 먼저 LTE 네트워크에서 단말에 제공하는 전송 속도에 영향을 미칠 수 있는 다양한 측정값을 선택하고, 실제 측정된 MDT 데이터 및 물리계층 측정값을 이용하여 처리량과의 상관관계를 분석한다. 또한, 다층 퍼셉트론 알고리즘을 활용하여 처리량 예측을 위한 모델을 구축하고, 처리량과 상관관계가 존재하는 물리계층 측정값을 이용하여 처리량 예측 모델을 학습한다. 실험을 통해 다층 퍼셉트론 기반 처리량 예측 수행 결과, 실제 측정값에 가까운 처리량을 예측하는 것을 확인할 수 있으며, 다른 예측기법에 비해 예측 정확도가 우수함을 확인할 수 있다.

**Key Words** : Minimization of Drive Test, physical layer measurements, LTE throughput prediction, multi layer perceptron

### ABSTRACT

In this paper, we analyze the relationship between various physical layer measurements including MDT data and throughput, and propose a multi layer perceptron(MLP) based throughput prediction method using the physical layer measurements. First, we analyze and select various physical layer measurements that affect the transmission rates to the UE in the LTE network using the actual measurements. In addition, we construct a MLP model for throughput prediction and train the prediction model using the physical layer measurements with throughput. From the simulation results, it is shown that the throughput prediction model based on the MLP shows the similar actual measured throughput, and has the improved results compared to other prediction techniques.

※ 본 연구는 과학기술정보통신부 및 정보통신기술진흥센터의 정보통신·방송 연구개발사업의 일환으로 수행되었습니다. [2017-0-00377, ICT융합기술개발지원사업(ICT R&D 바우처 지원) - 인공 지능 기반 무선망 품질 모니터링 시스템]

• First Author : (ORCID:0000-0001-5192-9432) School of Electrical and Electronic Engineering, Yonsei University, pjgstar@yonsei.ac.kr, 학생회원

◦ Corresponding Author : (ORCID:0000-0001-7413-1765) School of Electrical and Electronic Engineering, Yonsei University, csyong@yonsei.ac.kr, 중신회원

\* (ORCID:0000-0002-6490-0281) School of Electrical and Electronic Engineering, Yonsei University, kmdngwok@yonsei.ac.kr, 학생회원

\*\* (ORCID:0000-0003-4509-0153) School of Electrical and Electronic Engineering, Yonsei University, gykong@yonsei.ac.kr, 정회원  
논문번호 : KICS2018-04-076, Received April 3, 2018; Revised April 27, 2018; Accepted April 27, 2018

## I. 서 론

3GPP 표준이 진화함에 따라, 통신 속도를 높여주는 방안 이외에도 높은 데이터 전송률 요구량 만족, 커버리지 확장 및 음영 지역에 대한 지원을 위해 무선 망을 최적화시키려는 방안이 논의 중이다<sup>[1]</sup>. 일반적으로 무선망 초기 구축 시 또는 망 최적화 시, 기지국 또는 기지국 제어국은 자신의 셀 커버리지에 대한 LTE 품질 (처리량, 트래픽량, 무선링크실패 등)을 측정함으로써 현재 네트워크의 상태를 파악하는 작업을 수행한다<sup>[2]</sup>. 특히 하향링크 처리량은 네트워크 전반에 걸친 QoS (quality of service) 성능, 셀 내 음영지역 파악을 위해 매우 중요한 요소이다<sup>[2]</sup>.

이를 위해 측정자가 자동차에 측정 장비를 싣고 위치에 따른 하향링크 처리량 및 무선 환경 정보를 수집해야 하며, 이를 드라이브 테스트 (Drive Test)라고 한다<sup>[3]</sup>. 기존의 드라이브 테스트는 측정자가 반복적인 측정 업무를 장시간 수행하여야 하는 번거로움이 있으며, 여러 기지국이 발생시키는 신호를 다수의 지점에서 직접 측정함에 따른 시간 및 주파수 자원의 낭비를 야기한다. 이에 따라 무선망 최적화 비용 및 운영 비용을 증가시키는 문제가 있으므로, 드라이브 테스트를 최소화하고, 무선 환경의 분석 과정 및 수동설정을 개선하기 위한 연구가 MDT (Minimization of Drive Test) 라는 이름으로 진행되고 있다<sup>[3]</sup>. MDT의 핵심은 기지국이 단말에서 측정된 무선 환경 정보 및 기타 부가 정보인 MDT 데이터를 수집할 수 있게 허용하는 것이다. 이러한 MDT 데이터를 이용하여 LTE 네트워크를 최적화하려는 연구가 활발히 진행되고 있으며<sup>[3,4]</sup>, MDT 데이터를 포함한 물리계층 측정값을 이용한 하향링크 처리량 예측기법이 연구되고 있다<sup>[5]</sup>.

기존의 하향링크 처리량 예측기법은 LTE 네트워크 스캐너에서 측정된 각종 데이터를 이용하여 전체 OFDM 심볼 중 정보 전송에 사용되는 NRS (non-reference signal)의 비율을 계산하여 처리량을 예측한다<sup>[5]</sup>. 하지만 NRS를 계산하기 위한 EPRE (energy per resource element)의 비율과 같은 셀 특징 변수는 LTE 네트워크 스캐너가 아닌 일반 단말은 수집할 수 없는 측정값이므로, 단말로부터 획득 가능한 MDT 데이터를 포함한 물리계층 측정값을 활용한 처리량 예측 기술이 필요하다.

한편 최근 컴퓨터 하드웨어의 성능이 올라가면서 복잡한 행렬 연산에 걸리는 시간이 크게 줄어들게 되어 인공 신경망(neural network) 기반의 심층 학습 (deep learning) 알고리즘들이 주목받기 시작했다<sup>[6]</sup>.

심층 학습 알고리즘은 기계학습의 한 분야로서, 여러 비선형 변환기법의 조합을 통해 다량의 데이터나 복잡한 자료들 속에서 핵심적인 내용 또는 기능을 요약하는 작업을 수행하는 알고리즘이다<sup>[7]</sup>. LTE 네트워크 기지국에서 단말의 MDT 데이터가 1일 100억 건가량 수집되는 실제 통신 상황을 고려하였을 때, 무선 환경 정보와 처리량 사이의 관계를 심층 학습 알고리즘을 이용하여 추정할 수 있다면 무선망 최적화 비용을 줄이며, 처리량 예측 정확도를 향상할 수 있을 것으로 예상된다.

본 논문에서는 국내에서 측정된 MDT 데이터를 포함한 물리계층 측정값을 이용하여, 각각의 측정값과 처리량과의 상관관계를 분석하고, 심층 학습의 한 종류인 다층 퍼셉트론 (multi-layer perceptron, MLP)을 활용하여 물리계층 측정값으로부터 처리량을 예측할 수 있는 다층 퍼셉트론 기반 처리량 예측기법을 제안한다. 이를 위해 II장에서는 국내에서 측정된 MDT 데이터 및 물리계층 측정값을 분석하여 처리량과의 상관관계를 알아보고, III장에서는 물리계층 측정값을 활용한 다층 퍼셉트론 기반 처리량 예측기법에 대해 설명하며, 예측 성능을 분석한다. 마지막으로 IV장에서는 결론으로 마무리된다.

## II. MDT 데이터 및 물리계층 측정값

본 장에서는 LTE 시스템의 무선 채널 상태를 나타내는 많은 지표 가운데 처리량 예측을 위하여 선정된 것들에 대해 설명하고, 국내에서 실제 측정된 결과를 이용하여 각 측정값과 처리량과의 관계를 분석한다.

### 2.1 처리량 예측을 위한 MDT 데이터 및 물리계층 측정값

MDT 데이터는 드라이브 테스트를 최소화하기 위한 목적으로, 단말에서 측정된 무선 채널 측정 정보 및 기타 부가 정보를 기지국으로 보고하는 3GPP TS 37.320에서 정의한 데이터의 규격이며, reference signal received power (RSRP), reference signal received quality (RSRQ), power headroom (PH), data volume, scheduled IP throughput 등이 있다<sup>[8]</sup>. 또한, 단말로 부터 signal to interference and noise ratio (SINR), rank indicator (RI), channel quality indicator (CQI), precoding matrix index (PMI), modulation and coding scheme (MCS), 하향링크 ACK/NACK (acknowledgement/not-acknowledgement) 와 같은 물리계층 측정값들도 기지국으로 보고

된다.

일반적으로 처리량은 LTE 네트워크에서 단말에 제공하는 전송 속도와 밀접한 관련이 있으며, 전송 속도는 전송 신호의 세기와 품질, 하향링크 채널의 상태에 따라 변화하게 된다<sup>4)</sup>. 본 논문에서는 전송 신호의 세기와 품질, 하향링크 채널의 상태를 나타내는 많은 지표들 가운데 LTE 기지국에서 안정적인 확보가 가능한 RSRP, RSRQ, SINR, RI, CQI, MCS를 처리량 예측을 위한 측정값으로 선정하였다.

1) RSRP : 주파수 전 대역에 걸쳐 전송되는 CRS (cell-specific reference signal)를 운반하는 모든 RE (resource element)의 수신전력을 평균한 값이다. 그림 1에서와같이 기지국은 정해진 패턴으로 참조 신호 (reference signal, RS)를 전송하는데 이들 가운데 수신되는 참조 신호의 세기를 평균하여 구해진다. 그림 1에서 PRB는 physical resource block의 약자로 LTE 네트워크의 무선 전송 자원단위이며, DTx (discontinuous transmission)는 단말의 전력 소모를 줄이기 위하여 수신 동작을 중지하고 sleep 모드로 진입하도록 제어하는 기능을 담당하는 신호이다.

2) RSRQ : RSRP가 신호의 세기를 나타내기 위한 지표라면 RSRQ는 신호의 품질을 나타내기 위한 지표로, 측정 대역폭 또는 서브 밴드에 따른 신호 대 잡음 간섭비를 의미한다. RSRQ는 RSRP와 RSSI (received signal strength indicator)를 통해 구해질 수 있는데  $N$ 을 RSSI를 측정하기 위해 사용된 RB (resource block)의 개수라고 했을 때 다음의 식 (1)로 주어진다.

$$RSRQ = N \frac{RSRP}{RSSI} \quad (1)$$

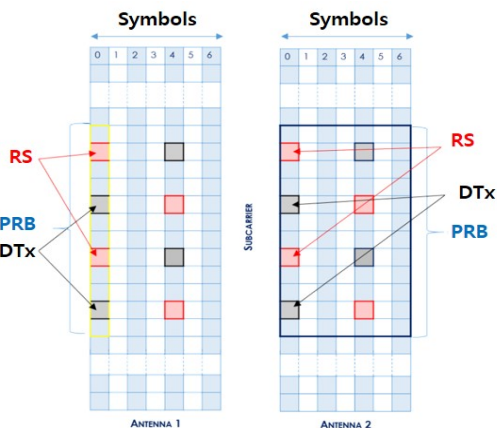


그림 1. OFDM 심볼 구조  
Fig. 1. OFDM symbol structure

여기서 RSSI는 신호의 세기를 나타내는 품질 지표로서 전체 주파수 대역에서 측정된 신호, 간섭, 열잡음을 모두 포함한 수신전력을 의미한다.

3) SINR : RSRQ와 마찬가지로 수신된 신호의 품질을 나타내는 지표로, 단말에서 측정된 신호 대비 다른 셀로부터의 간섭과 잡음의 비로 정의된다. RSRP나 RSRQ와 같이 기지국으로 보고되지는 않으나 CQI의 산출을 위해 활용된다.

4) RI : RI는 데이터 전송에 사용되는 채널의 rank에 대한 권고 값, 즉 해당 단말로의 하향링크 전송에 사용될 바라는 layer의 개수에 대한 정보를 전달하며, 단말이 네트워크에 제공해주는 CSI (channel state information)의 구성 요소이다.

5) CQI : CQI 역시 CSI의 구성 요소 중 하나로 단말로의 하향링크 채널의 품질에 대한 정보, 즉 단말이 권고하는 RI와 PMI (precoder matrix indicator)가 사용될 때 10% 이하의 블록 오류율 (block error rate, BLER)로 수신될 수 있는 PDSCH (physical downlink shared channel) 전송에 대한 가능한 가장 높은 MCS의 추정치이다. CQI는 정숫값으로 표현되는데 높은 값일수록 채널의 품질이 더 좋음을 나타낸다.

6) MCS : MCS는 MAC (medium access control)에서 기지국과 단말 간의 무선 채널 환경을 고려하여 결정되는 변조 및 코딩 방식에 대한 정보를 제공한다. 무선 채널 상태가 좋을수록 높은 MCS 인덱스를 할당하여 단위 PRB당 더 많은 데이터 비트 수를 전송할 수 있게 하고, 무선 채널 상태가 좋지 않을 때는 낮은 MCS 인덱스를 할당하여 무선 구간에 발생하는 데이터 수신 오류를 낮출 수 있게 한다.

## 2.2 MDT 데이터 및 물리계층 측정값 분석

물리계층 측정값과 처리량과의 상관관계 분석 및 처리량 예측을 위해 20MHz의 대역폭을 가지며 동시에 최대 2개의 데이터 스트림을 주고받을 수 있는 환경에서 실제 측정된 MDT 데이터 및 물리계층 측정값을 사용한다. 하나의 측정값 샘플은 동일한 시간에 측정된 대역폭, RI, RSRP, RSRQ, SINR, CQI, MCS, 처리량을 포함하고 있으며, 총 41,626개의 샘플을 사용하였다. 그림 2는 측정값의 분포를 나타낸 것이며 표 1은 각 측정값의 요약, 표 2는 일반적인 LTE 무선 접속 채널의 품질에 따른 환경의 분류 기준을 나타낸 것이다<sup>4)</sup>.

그림 2의 물리계층 측정값 분포와 표 2에 제시된 기준값을 비교해보면, RSRP 측면에서 측정값의 55%

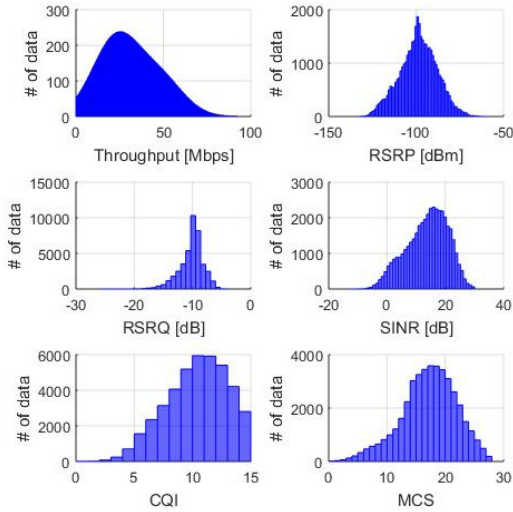


그림 2. 물리계층 측정값의 분포  
Fig. 2. Distribution of physical layer measurements

표 1. 물리계층 측정값 요약  
Table. 1. Summary of the physical layer measurement

Bandwidth		20MHz		
# of data		41,626		
	RSRP (dBm)	RSRQ (dB)	SINR (dB)	
Max	-59.3	-3.8	30.0	
Min.	-133.2	-24.9	-12.0	
Avg.	-98.6	-10.2	13.8	
	CQI	MCS	Throughput (Mbps)	
Max	15	28	92.1	
Min.	0	0	0	
Avg.	10.3	17.2	32.0	

표 2. 무선 접속 채널 품질에 따른 측정 환경의 분류  
Table. 2. The classification of measurement environments per the RF channel quality

		RSRP (dBm)	RSRQ (dB)	SINR (dB)
RF condition	Excellent	≥ -80	≥ -10	≥ 20
	Good	-80 to -90	-10 to -15	13 to 20
	Mid cell	-90 to -100	-15 to -20	0 to 13
	Cell Edge	≤ -100	< -20	≤ 0

가 ‘우수 (Good)’ 또는 ‘보통 (Mid cell)’ 환경에서 측정된 값이며 나머지 45%의 측정값은 ‘나쁜 (cell edge)’ 환경에서의 측정된 결과임을 알 수 있다.

RSRP 기준으로는 ‘매우 우수 (Excellent)’ 환경과

‘우수’한 환경에서의 측정값이 각각 55%와 45%의 비율을 갖는 것을 알 수 있으며, SINR 기준에서는 ‘매우 우수’와 ‘우수’, ‘보통’, ‘나쁜’ 환경에서 측정된 측정값의 비율이 각각 20%, 37%, 38%, 5%임을 알 수 있다. 측정값의 평균 기준으로는 RSRP 측면에서는 ‘보통’, RSRQ 및 SINR 측면에서는 ‘매우 우수’한 환경에서 측정된 결과임을 알 수 있다.

하나의 측정값 샘플은 동일 시간에 측정된 여러 변숫값의 집합이므로 서로 연계되어 있다. 따라서 측정값 샘플 전체를 변수별로 나열하여 각각의 변숫값들이 변화하는 모양을 비교하면 처리량과의 관계를 유추할 수 있다. 이를 위해 측정 변수 중 하나를 기준으로 하여, 기준 변수가 증가하는 순서대로 측정값 샘플을 정렬하고 변수별로 나열한 결과가 그림 3에 나타나 있다.

이때 각 변수의 범위 및 단위가 다르므로, 정규화를 통해 하나의 범위로 변환한다.  $X = [x_1, x_2, \dots, x_n]$  라는 변수에 대해  $x_i$ 의 정규화된 값  $\tilde{x}_i$ 는 식 (2)와 같다.

$$\tilde{x}_i = \frac{x_i - \min\{X\}}{\max\{X\} - \min\{X\}} \quad (2)$$

여기서  $\min\{X\}$ 와  $\max\{X\}$ 는 각각  $X$ 의 최솟값과 최댓값을 의미한다.

그림 3은 각 그래프 상단에 표기된 기준 변수의 오름차순으로 모든 변수가 나열된 결과이다. 물리계층 측정값과 처리량과의 관계는 세 가지 양상으로 관찰되는데, 첫 번째로 RSRP는 처리량과 상관관계를 갖지만, 다른 물리계층 측정값과 비교하면 상관도가 높지 않은 것으로 추정된다. 또한, RSRQ는 처리량과 비례 관계와 반비례 관계가 교차하는 형태를 보인다. RSRP, RSRQ를 제외한 SINR, CQI, MCS는 처리량과 비례 관계를 갖는 것으로 보이며, 각각의 측정값의 증가량과 처리량의 증가량이 유사한 것을 확인할 수 있다.

이와 같은 결과를 토대로 알 수 있는 것은 각각의 물리계층 측정값은 처리량과 상관관계가 존재한다는 것이다. 하지만 물리계층 측정값마다 처리량과의 상관관계가 다르며, 이는 물리계층 측정값으로 처리량을 표현할 때, 각각의 측정값에 대한 기중치가 다르다는 것을 의미한다. 본 논문에서는 다음 장에 소개될 다층 퍼셉트론을 활용하여, 각 측정값의 기중치를 구하여 처리량을 추정한다.

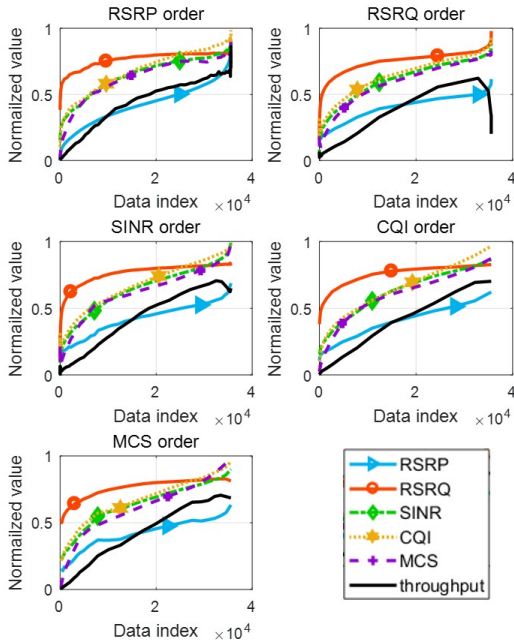


그림 3. 기준 변수의 오름차순으로 나열된 측정값  
Fig. 3. The measurements in ascending order of reference variable

### III. 다층 퍼셉트론 기반 처리량 예측기법

본 장에서는 처리량 예측을 위한 다층 퍼셉트론 알고리즘을 소개하고, 다층 퍼셉트론 기반의 처리량 예측 결과에 대해 분석한다.

#### 3.1 다층 퍼셉트론 (Multi Layer Perceptron)

인공 신경망은 생물학적 뉴런의 구조 및 기능을 단순화하여 수학적으로 표현한 것으로, 시냅스의 결합으로 네트워크를 형성한 인공 뉴런(노드)이 학습을 통해 시냅스의 결합 세기를 변화시켜 문제 해결 능력을 가지는 모델 전반을 가리킨다. 즉, 인공 신경망은 신경망의 입력으로부터 가중치를 자동으로 학습하는 능력을 가지는 모델로, 퍼셉트론은 이러한 인공 신경망을 구현하는 알고리즘의 한 종류이다<sup>6)</sup>. 그림 4는 퍼셉트론 알고리즘 개념을 도식화한 것이다.

퍼셉트론은 다수의 훈련 데이터를 이용하여 지도 학습을 수행하는 알고리즘으로 그림 4에서처럼 입력값  $x_1 \sim x_n$ 에 대한 실제 결괏값이  $y$ 일 때, 입력값  $x_1 \sim x_n$ 에 가중치  $w_1 \sim w_n$ 를 곱한 값을 모두 더하여 하나의 값으로 만들고, 시그모이드 (sigmoid), tanh (hyperbolic tangent), ReLU (rectified linear unit) 등의 활성화 함수 (activation function)를 이용하여 변환

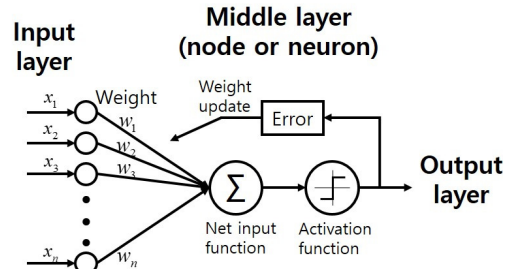


그림 4. 퍼셉트론 알고리즘  
Fig. 4. Perceptron algorithm

한 값이 실제 결괏값  $y$ 의 변환 결과와 같아질 때까지 가중치  $w_1 \sim w_n$ 를 수정한다. 그림 4처럼 중간층이 하나의 노드로 구성되어 중간층과 출력층의 구분이 없는 구조를 단층 퍼셉트론이라 하며, 그림 5처럼 은닉층 (hidden layer)라고도 불리는 중간층을 구성하는 노드가 여러 개이고, 이러한 중간층이 다수로 구성되어 있는 구조를 다층 퍼셉트론이라 한다.

다층 퍼셉트론은 학습 단계와 예측 단계로 구분된다. 학습 단계는 입력 값에 따라 각 뉴런의 입력과 활성화 함수를 사용하여 출력을 산출하는 전방향처리 (Feedforward) 과정과 목표값과 출력값과의 차이를 역으로 진행하면서 오차가 최소가 되도록 가중치를 변화, 조절하는 역방향처리 (Feedback) 과정에 의해 수행된다. 이러한 가중치를 조절하는 대표적인 학습 방법으로 오류 역전파 알고리즘이 있다<sup>6)</sup>. 예측단계는 입력이 주어지면 학습 단계에서 최종적으로 획득한 가중치를 이용하여 출력을 계산한다.

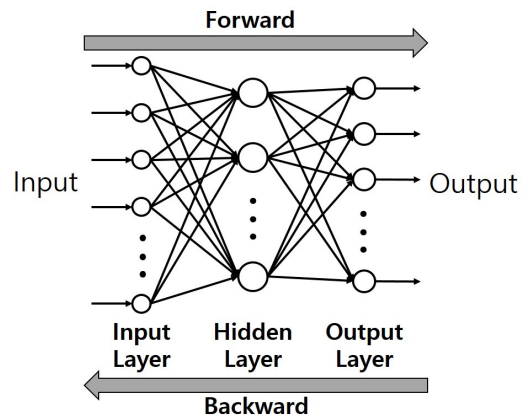


그림 5. 은닉층이 하나인 다층 퍼셉트론 구조  
Fig. 5. The structure of multi layer perceptron with 1 hidden layer



### 3.2 다층 퍼셉트론 기반의 처리량 예측기법

인공 신경망은 은닉층의 개수에 따라 단층 신경망과 심층 신경망으로 구분된다. 단층 신경망은 은닉층과 출력층의 구분이 없는 구조로서 구조가 간단하지만, 비선형적으로 분리되는 데이터에 대해서는 학습이 불가능하다는 단점이 있다. 이를 극복하기 위한 방안으로 입력층과 출력층 사이에 여러개의 은닉층을 두어 비선형적으로 분리되는 데이터에 대해서도 학습이 가능하도록 구성한 인공 신경망을 심층 신경망 (deep neural network) 이라 하며, 심층 신경망을 학습하기 위해 고안된 특별한 알고리즘들을 딥러닝 (deep learning)이라 한다.

앞 장에서 살펴본 MDT 데이터 및 물리계층 측정값과 처리량 사이에는 비선형적 관계가 존재할 것으로 추정되기 때문에, MDT 데이터 및 물리계층 측정값을 이용한 처리량 예측을 위해서는 비선형 특성을 근사 및 추정하는데 특화된 심층 신경망이 적합하다. 본 논문에서는 처리량 예측을 위해 심층 신경망의 가장 기초적인 구조 및 알고리즘인 다층 퍼셉트론을 적용하였다.

다층 퍼셉트론 기반 처리량 예측을 위한 학습 및 성능평가는  $k$ 겹 교차검증 (Stratified  $k$ -fold cross validation)을 통해 수행된다.  $k$ 겹 교차검증은 측정값 샘플을  $k$ 개의 세트로 나눈 후  $k-1$ 개의 세트는 학습을 시키고 나머지 1개의 집합으로 성능을 측정하는 방법이다. 본 논문에서는  $k$ 를 4로 설정하였으며, 학습을 위한 세트와 성능 측정을 위한 세트는 임의 추출을 통해 선택하였다.

학습을 위한 측정값 샘플 중에서 RSRP, RSRQ, SINR, RI, CQI, MCS는 다층 퍼셉트론의 입력값으로, 처리량은 가중치 조절을 위한 목표값으로 사용하였으며, 은닉층의 출력값 변화를 위한 활성화 함수로는 ReLU를 사용하였다. 이때 ReLU의 출력은 0과 1 사이의 값만을 반환하기 때문에, 다층 퍼셉트론의 입력값들 역시 0과 1 사이의 값을 갖도록 정규화된 값으로 변환하여 사용하였다.

다층 퍼셉트론의 은닉층 개수 및 뉴런 수에 따른 처리량 예측 성능을 평가하기 위해 은닉층의 수와 은닉층 별 뉴런의 수는 다양하게 고려하였으며, 각 층의 가중치 갱신은 경사 하강법 (gradient descent) 기반의 역전파 알고리즘을 적용하였다. 학습 횟수는 500,000 회, 학습률은 0.001로 설정하였으며, 학습 과정에서 출력 값과 목표값의 차이는 평균 제곱 오차 (mean square error, MSE)를 통해 계산하였다. 이상의 학습 모델 설정 결과가 표 3에 정리되어 있다.

표 3. 처리량 예측을 위한 다층 퍼셉트론 모델의 설정값  
Table. 3. Set values of the multi layer perceptron model for throughput prediction

Learning target	Throughput
Input for MLP	RSRP, RSRQ, SINR, CQI, MCS, RI
# of hidden layer	2 or 3
# of neuron per layer	100 or 150
Activation function	ReLU
Learning method	Backpropagation method with Gradient descent algorithm
Learning number	500,000
Learning rate	0.001
Cost function	Mean square error

학습 단계에서 다층 퍼셉트론은 다음과 같이 동작한다.

- 1) 학습을 위한 다층 퍼셉트론 구조 결정 및 각 층에서의 가중치와 바이어스 초기화
  - 2) 학습을 위한 데이터에 대해서 각 층에서 순입력 함수값(net input function)을 계산하고 최종적으로 활성화 함수에 의한 출력값을 계산
  - 3) 출력층의 활성화 함수에 의한 결괏값과 목표값의 오차를 반영하여 각 층의 가중치를 갱신
- 학습이 완료된 후  $k$ 겹 교차검증으로 나뉘진 성능 측정을 위한 측정값 샘플 세트를 활용하여 처리량을 예측한다.

### 3.3 처리량 예측 결과

본 논문에서는 II장에서 분석한 측정값 샘플을 사용하여 처리량 예측을 위한 다층 퍼셉트론 학습 및 다층 퍼셉트론 기반의 처리량 예측을 수행하였다. 다층 퍼셉트론 학습 횟수에 따른 학습 단계의 오차가 그림 6에 나타나 있다. 다층 퍼셉트론 은닉층 수에 따라 학습 오차가 감소하는 속도가 다른 것을 확인할 수 있으며, 은닉층의 수가 2개인 다층 퍼셉트론 모델이 3개인 모델에 비해 학습 오차가 빨리 감소한다. 하지만 은닉층 뉴런 수에 따른 학습 오차 차이는 관찰되지 않았으며, 학습 횟수가 증가할수록 모든 경우의 학습 오차가 같은 값으로 수렴하는 것을 확인할 수 있다.

그림 7은 은닉층 개수 및 은닉층 뉴런 수에 따른 처리량 예측 결과를 나타내었다. 표 4는 처리량 예측의 정확도를 정리한 것이다. 제안 기법의 예측 성능평가를 위해 실제 측정된 처리량과 통계적 예측에 이용하는 기법의 하나인 다중 회기 분석<sup>[9]</sup> 기반 처리량 예측 결과도 함께 나타내었다. 처리량 예측 정확도를 평

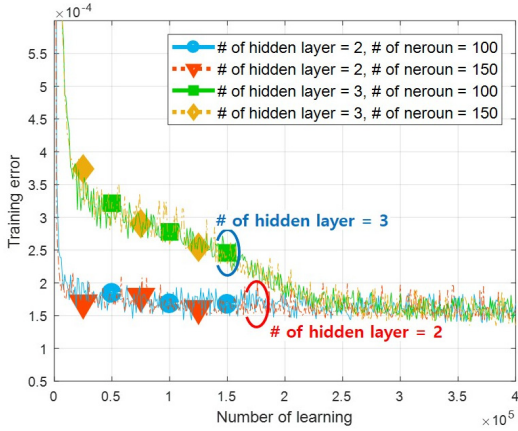


그림 6. 학습 횟수에 따른 학습 오차  
Fig. 6. The training error according to the number of learning

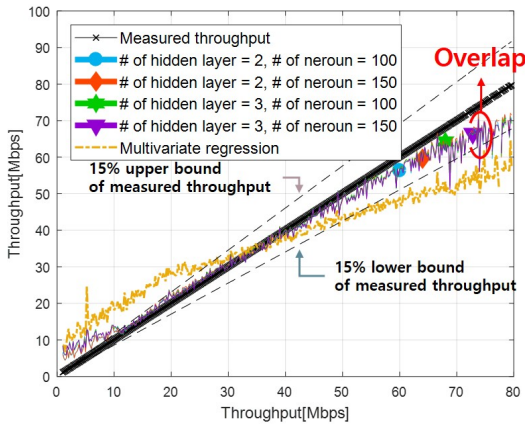


그림 7. 처리량 예측 결과  
Fig. 7. Throughput prediction results

가하기 위한 척도는 평균 제곱 오차와  $n\%$  outage 오차 확률  $P_{out,n}$  두 가지를 사용하였다. 평균 제곱 오차는 각 기법의 예측 결과와 실제 측정값 사이의 오차의 제곱에 대해 평균을 취한 값이며,  $n\%$  outage 오차 확률  $P_{out,n}$ 은 예측 결과가 실제 측정값 대비  $P_{out,n}$ 의 확률로 오차범위  $n\%$  안에 발생함을 의미한다. 본 논문에서는 15% outage 확률을 성능평가 척도로 설정하였다.

처리량 예측 결과를 보면 다층 퍼셉트론 기반의 처리량 예측 결과는 은닉층 수 및 뉴런 수에 큰 영향을 받지 않으며, 다중 회귀 분석 기법에 비해 실제 측정된 처리량에 가까운 예측 결과를 보이는 것을 알 수 있다. 또한, 그림 2의 측정된 처리량 분포와 그림 7의 처리량 예측 결과를 비교하면, 측정된 처리량이 많은 영역에서는 다층 퍼셉트론 기반 처리량 예측 모델이

표 4. 예측 방법에 따른 처리량 예측 정확도  
Table. 4. Accuracy of throughput prediction based on prediction method

Throughput prediction scheme		Measure	
		MSE	$P_{out,15}$
Multi layer perceptron	# of hidden layer = 2 # of neuron = 100	17.27	84.34%
	# of hidden layer = 2 # of neuron = 150	16.31	83.59%
	# of hidden layer = 3 # of neuron = 100	18.54	82.58%
	# of hidden layer = 3 # of neuron = 150	18.90	83.08%
Multivariate regression		110.50	30.81%

실제 처리량을 정확하게 예측하지만, 측정된 처리량의 수가 적은 영역에서는 처리량 예측 결과가 실제 측정된 처리량과 큰 차이가 발생하는 것을 확인할 수 있다. 처리량 예측 방법에 따른 평균 정확도 차이는 표 4를 통해 확인할 수 있는데, 다층 퍼셉트론 기반의 처리량 예측기법은 MSE 측면에서 다중 회귀 분석에 비해 낮은 예측 오차를 가지며  $P_{out,15}$  측면 또한 다중 회귀 분석에 비해 높은 예측 정확도를 갖는다. 특히 다층 퍼셉트론 은닉층의 수가 2개이고, 뉴런의 개수가 100개 일 때 84.34%의 가장 높은 예측 정확도를 보인다. 이를 통하여 다층 퍼셉트론을 이용한 처리량 예측 기술이 매우 효율적인 것으로 분석된다.

#### IV. 결론

본 논문에서는 다층 퍼셉트론을 이용하여, 국내에서 측정된 물리계층 측정값을 통한 LTE 처리량 예측 기법을 제안하였으며, 예측된 결과를 측정치와 비교하여 다음과 같은 결론을 도출하였다.

- 1) 다층 퍼셉트론 기반 처리량 예측은 다중 회귀 분석 기반 처리량 예측에 비해 높은 예측 정확도를 보인다.
- 2) MSE 측면의 최적 다층 퍼셉트론 구조와  $P_{out,15}$  측면의 최적 다층 퍼셉트론 구조가 다른 것을 확인할 수 있다.
- 3) 처리량 예측을 위한 다층 퍼셉트론 학습에 사용된 측정값의 분포에 따라 처리량 예측 정확도가 달라지는 것을 확인할 수 있다.

이상의 연구를 바탕으로 향후 다층 퍼셉트론의 다양한 구조에 따른 예측 정확도 분석을 통해 처리량 예측에 적합한 최적의 구조를 탐색하는 연구가 필요하다.

며, 보다 다양한 측정결과를 이용하여 수집한 처리량 예측 결과에 대한 분석 또한 필요하다. 또한, 처리량 예측뿐만 아니라 단말에서 간헐적으로 발생하는 VoLTE drop과 같은 중요한 이벤트를 예측하는 연구로의 발전도 기대할 수 있다.

References

[1] M. Shafi, et al., “5G: A tutorial overview of standards, trials, challenges, deployment, and practice,” *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1201-1221, Jun. 2017.

[2] R. Wang, J. Zhang, S.H. Song, and K. B. Letaief, “Average throughput analysis of downlink cellular networks with multi-antenna base stations,” *IEEE 25th Annu. Int. Symp. Pers.*, 2014.

[3] J. Johansson, W. A. Hapsari, S. Kelley, and G. Bodog, “Minimization of drive tests in 3GPP release 11,” *IEEE Commun. Mag.*, Nov. 2012.

[4] B. J. Kim, “The analysis of the relationship between the throughput and the wireless channel conditions of a LTE mobile communication system,” *J. KIES*, vol. 10, no. 2, pp. 219-26, Feb. 2015.

[5] K. Chang, R. P. Wicaksono, and S. Kunishige, “Scanner based load estimation for LTE networks,” *Int. Conf. ICT Convergence*, pp. 413-418, Oct. 2015.

[6] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press. 2016,

[7] Y. Bengio, A. Courville, and P. Vincent, “Representation learning: A review and new perspectives,” *IEEE Trans. PAMI, Special Issue Learning Deep Architectures*, 2013.

[8] 3GPP TS 37.320, *Radio measurement collection for minimization of drive tests*, v.13.1.0. 2016.

[9] MG. Evans, “A monte carlo study of the effects of correlated method variance in moderated multiple regression analysis,” *Organ. Behav. Hum. Decis. Process.*, vol. 36, no. 3, pp. 305-323, Dec. 1985.

박 정 균 (Junggyun Park)



2009년 2월 : 연세대학교 전기 전자공학과 졸업  
 2009년 3월~현재 : 연세대학교 전기전자공학과 박사과정  
 <관심분야> Wireless communication system, Artificial Intelligence / Machine Learning Algorithms for Communication Systems

김 동 욱 (Dongwook Kim)



2017년 8월 : 연세대학교 전기 전자공학과 졸업  
 2017년 9월~현재 : 연세대학교 전기전자공학과 석사과정  
 <관심분야> Wireless communication system, Artificial Intelligence / Machine Learning Algorithms for Communication Systems

공 규 열 (Gyuyeol Kong)



2007년 8월 : 연세대학교 전기 전자공학과 졸업  
 2014년 2월 : 연세대학교 전기 전자공학과 박사  
 2014년 3월~2016년 3월 : 삼성전자 메모리사업부 책임연구원

2016년 4월~현재 : 연세대학교 산학협력단 Research Professor  
 <관심분야> Signal Processing and Error Correction Codes for Communication Systems, Artificial Intelligence / Machine Learning Algorithms for Communication Systems



최 수 용 (Sooyong Choi)



1995년 8월: 연세대학교 전자  
공학과 졸업

1997년 8월: 연세대학교 전자  
공학과 석사

2001년 8월: 연세대학교 전자  
공학과 박사

2002년 3월~2004년 8월:

University of California, San Diego,  
Postgraduate Researcher

2004년 9월~2005년 7월: Oklahoma State  
University, Research Assistant Professor and  
Researcher

2005년 9월~2011년 8월: 연세대학교 전기전자공학  
과 조교수

2011년 9월~2016년 8월: 연세대학교 전기전자공학  
과 부교수

2016년 9월~현재: 연세대학교 전기전자공학과 정교  
수

<관심분야> Massive MIMO, FBMC, Interference  
management/coordination/cancellation/mitigation,  
Detection and estimation, Iterative receivers, 무  
선 통신 시스템, Signal processing and machine  
learning for communication systems