International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-8 Issue-6S3, September 2019

Deep Learning Based Energy Efficient Scheme For Massive MIMO

TRV. Anandharajan, C. Murugalakshmi, B. Adhitya, K. Swetha

***Abstract:** This paper proposes a Deep Learning Energy Efficient Scheme (DLEE) for a massive multiple input multiple output system (MIMO). Massive MIMO is deployed using large number of antennas for multiple users. The proposed DLEE, learns the relationship between spatial beamforming pattern and the power consumption in a base station. In this work, we design a novel learning method where the spatial correlation across UE antennas are taken as input feature vector and find the output labels which give us the energy efficiency in a BS. Due to multipath propagation, other methods only try to address the energy efficiency problem through the bit rate and the power required for the throughput to be efficient. This paper discusses the unsupervised algorithm DLEE which is similar to an autoencoder by combining the power consumed due to radiation pattern through beamforming and the DL framework to address the energy efficiency to an extent of 12% in a BS.*

***Keywords:** deep learning; Massive MIMO; Beamforming.*

I. INTRODUCTION

This The power consumption of the communication technology industry and the carbon emission due to the power depletion in the base station are the major concerns in both societal and economic conditions. The green cellular network is a new research area of intense activity in both academy and industry. The power consumed in cellular networks is a major concern for the subscriber since the retail, data center and cellular infrastructure consume only 40% of the power on the whole where as the base stations consume 60% of the power. In a base station, the power supply, signal processing and air cooling consume 35% of the total base station power and the remaining 65% of the power is consumed by the power amplifier. Hence this motivates us towards this work of maximizing the energy efficiency in a base station [1].

Manuscript published on 30 September 2019.
* Correspondence Author (s)
Dr. TRV. Anandharajan, Professor, Dept. of ECE, Malla Reddy Engineering College for Women(Autonomous), Secunderabad-TS
trvanandharajan@gmail.com
C. Murugalakshmi, Assistant Professor, Einstein College of Engineering, Tirunelveli, Tamilnadu, India
laxmiraju792@gmail.com
B. Adhitya, Asst. Professor, Dept. of ECE, Malla Reddy Engineering College for Women(Autonomous), Secunderabad-TSaadi.mrecw@gmail.com
K. Swetha, Asst. Professor, Dept. of ECE, Malla Reddy Engineering College for Women(Autonomous), Secunderabad-TS
swetaelex2008@gmail.com

The energy efficiency of any communication system is measured in bits/joule i.e, it can be calculated by computing the ratio of total amount of data transmitted per second to the energy consumed to transmit that data which is given by:

$$EE = T/Pt = bits/sec\ /\ joule/sec = bits/joule$$

The energy efficiency is the use of energy in an optimum manner to achieve the required energy output. 5G systems have a very high carbon footprint [2], approximately around 235 Mt CO_2, due to the deployment of more than 45 base stations per km while compared to 4G systems which is 170 Mt of CO_2 for 8 BS/km respectively.

II. SYSTEM MODEL

A. Problem Statement

First, In cellular networks, we are much concerned about the energy consumption at the BS. These must operate at low consumption levels generating large gains to increase reliability. Such criteria we are defining in this context is the Energy Efficiency, which is defined as the throughput to the energy spent in achieving that throughput. To achieve an iMIMO in how the beamforming is being performed, we define the following problem

Problem 1: An DLEE is achieved by solving the following optimization problem

$$\max_{M \in Z,\ K \in Z,\ \rho \geq 0} EE^{(DLEE)} = \frac{K\left(1 - \frac{\tau_{sum} K}{U}\right) R}{\frac{B\sigma^2 \rho S_x}{\eta} K + \mathrm{P}_{cp}^{(ZF)}}$$

Subject to constraints, with respect to uplink and downlink configurations. Massive MIMO technology which is recently proposed, offers the spectral and energy efficiency gains. In massive MIMO technology, the base station which is equipped with the large number of M antennas serves the K single antenna user equipments (UE), where M>>K. The main aim of this paper is to achieve high energy efficiency by increasing the number of base station antennas without consuming more power. The gain can be achieved using simple linear processing techniques at the base station, such as Maximum Ratio Transmission/Combining (MRT/MRC), Zero-Forcing(ZF) and Minimum Mean Square Error (MMSE). In a single-cell multiuser MIMO system, we have to consider uplink and downlink which is operating at a bandwidth of B(Hz). The M antennas of the BS must communicate with K single antenna UEs, and that particular UE is selected by round robin fashion, from the large set of UEs within that coverage area.

Considering block flat-fading channels, where B_c (Hz) is the coherence bandwidth and T_c(sec) is the coherence time. Hence the channels are static within time-frequency coherence blocks of $U=B_cT_c$ symbols. If BS and UEs are synchronized and operate with time-division duplex(TTD) as shown in Figure 1.

A number of user equipments in a cell area is addressed by a corresponding base station b. A number of antennas are deployed in a single base station. The UE may or may not use the same base station antenna or nearby base station antenna. The gross rate is similar to data rate but overhead factors are not included for this measure, depending on a single antenna UE hence depends on two base station one in the cell at which it is localized and the other is the nearby neighbouring base station. The resulting gross rate can

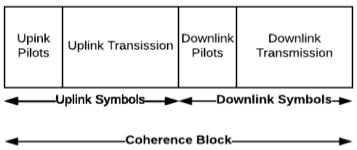

Figure 1: Illustration of TDD protocol, where $\zeta^{(ul)}$ and $\zeta^{(dl)}$ are the fractions of UL and DL transmission respectively.

In order to calculate total amount of data transmitted per one user we have to calculate total uplink and downlink user grossrate i.e

$$T=\Sigma\left(E\left(R_K^{(Ul)}\right) + E\left(R_K^{(Dl)}\right)\right)$$

And total power consumption is given by

$$P_t=P_{tx}^{(Ul)} + P_{tx}^{(Dl)} + P_{CP}$$

where, P_{cp}=circuit power

In this paper modeling can be done in order to achieve maximum EE by designing the link for P_{cp} as a function P_{cp}(M,K,R). While we go for the above modeling process there is an issue arising, known as optimization. That is:

$$EEmax=T/Pt$$

T is the amount of information in bits and P_t is the power consumed in joules. Thus this optimization problem can be solved by modelling technique using Zero Forcing Processing as shown below, in the section no. IV.

The circuit power is calculated using circuit power computation

$$P_SP_DL(index)=4*M*K*B/(tau_c*L_BS)$$

This shows the circuit power for the precoding vectors. Denominator of the function is the product of coherence block and the number of cells in a base station. This function can also be used for similar computations in terms of Uplink channels for other base stations for MR, ZF, MMSE etc. Some of our previous research we focussed on Machine Learning and now in this paper we have addressed Deep Learning for MIMO systems [14], [15], [16], [17].

B. Channel model and Linear processing:

The spectrum of massive MIMO consists of M antennas at the base station which are equally placed in such a way that the channel components between the BS antenna and single antenna UE's are uncorrelated. Rayleigh small scale fading distribution calculates the energy efficiency of massive MIMO antenna at BS and user location [7].

The channel vector h_k has $\{h_{k,n}\}$ entries that describes the instantaneous propagation channel between the **n**th antenna at the BS and the **k**th UE's. Where in Time Division Duplex the uplink data detection and downlink data precoding is done by Linear Processing. During the uplink data transmission the BS is able to acquire perfect Channel State Information (CSI) [3], [6] from the uplink pilots. Based on the channel vector h_k the uplink linear receives combining matrix G, which is given by

$$G=\left[g_1,g_2,\dots,g_K\right]\in\mathbb{C}^{M\times K}$$

Here, we are considering MRC/MRT, ZF, MMSE methods for uplink detection.

Which gives,

$$G=\begin{cases} H & for\ MRC, \\ H(H^HH)^{-1} & for\ ZF, \\ \left(HP^{(ul)}H^H+\sigma^2I_M\right)^{-1}H & for\ MMSE, \end{cases}$$

Where,

H = All user channels $\left[h_1,h_2,\dots,h_K\right]$

σ^2 = Noise variance (joules/symbol)

$P^{(ul)}$ = Transmitted uplink power

$$P^{(ul)}=\left[P_1^{(ul)},P_2^{(ul)},\dots,P_K^{(ul)}\right]$$

Similarly,

For downlink transmission, linear precoding matrix V is given by

$$V=\left[V_1,V_2,\dots,V_K\right]\in\mathbb{C}^{M\times K}$$

For downlink transmission also we consider MRC, ZF, MMSE as precoding schemes, which gives

$$V=\begin{cases} H & for\ MRC, \\ H(H^HH)^{-1} & for\ ZF, \\ \left(HP^{(ul)}H^H+\sigma^2I_M\right)^{-1}H & for\ MMSE, \end{cases}$$

By combining the linear processing with proper power allocation we can achieve uplink rate

($\zeta^{(ul)}\overline{R}$), downlink rate ($\zeta^{(dl)}\overline{R}$), and uniform gross rate ($\overline{R}$) in bits/second for any active User Equipment which is explained below.

C. Uplink and Downlink

Generally, in telecommunications, a link is a communication channel that connects two or more devices in order to transmit the information. The communication which is going from kth antenna UE to BS is called uplink. The communication which is going from BS to kth antenna UE is called downlink [4], [5].

Retrieval Number: F13380986S319/2019©BEIESP
DOI:10.35940/ijeat.F1338.0986S319
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Uplink frequency is always higher than the downlink frequency such as 6/4 GHZ, 14/11 GHZ, 30/20 GHZ.

In Time Division Duplex, uplink is being received by the BS, at the same time downlink is being received by the cellular antenna. This type of communication is called Two-Way communication.

For a realistic power consumption model we take the power consumption of all components involved as shown in the next section.

D.Realistic circuit power consumption model

In order to calculate circuit power consumption we have to compute the average sum of all the power consumptions by various analog components and processing elements. We proposed a model to calculate all power consumptions by various components for multi user Massive MIMO [7] as shown below

Where,

Pfix= fixed power

Ptc=power consumed by no of transmission chains

Pce=power consumed for channel estimation

Pc/d=power consumed for coding and decoding

Pbh= power consumed for transfering U.L/D.L data between B.S & N/W

Plp=power consumed for linear processing purpose

E. Channel Estimation

Based on UE requirement, transmission process can be done in between BS and UE whose computational theoretical efficiency can be Lbs and Lue in joule or flops per watt. As there exists B/U coherence blocks per second in each link transmission, in order to require channel for each transmission process from BS and UE we modeled a pilot based CSI estimation.The pilot based CSI estimation can be done once per block for U.L and D.L as total B/U coherence blocks exist per second.

In Uplink, UE transmits pilot signal so that estimation for each UE channel can be done by multiplying with respective pilot sequence of length T(U.L)*K i.e, the transmitted pilot signal can be given as

$$M * T(U.L)\ K$$

i.e, the power consumption for channel estimation at uplink is given by,

$$Pce(UL) = B/U2T(UL)MK * K/Lbs\ \ watt$$

In Downlink, each active UE receives a pilot sequence of length T(D.L)*K so that CSI estimation can be done with effective precoded channel gain and some interference in addition to some noise, the transmitted pilot signal can be given as

$$M * T(D.L)\ K$$

i.e, the power consumption for channel estimation at down link is given by,

$$Pce(DL) = B/U4T(DL)K * K/Lue\ watt$$

Hence total power consumption for channel estimation becomes

$$Pce = Pce(U.L) + Pce(D.L)$$

In order to go for channel estimation purpose some amount of power is required.Thus the power consumption for channel estimation can be calculated by using the formula,

$$P_C E(index) = 3 * K * B/(tau_c * L_B S) * (M * tau_p + M^2/)$$

which shows that the amount of power required in which the numerator of RHS gives the product of total bandwidth required to the total number of UEs per cell. Denominator is the product of coherence block and pilot signals for M number of base stations to transmit.

F. Coding and Decoding:

In downlink as each user transmits signals to BS so that the coding and modulation can be done to K sequence of information. For each UE there is no same algorithm applied for decoding so that the power consumption Pc/d is always based on the no of symbols transmitted i.e, the number of bits which is given by,

$$Pc/d =^{T} = \Sigma(E(R_K^{(Ul)}) + E(R_K^{(Dl)})) * (Pcod + Pdoc)$$

Where,

Pcod= Power required for coding in watts bits/sec

Pdoc=power required for decoding in watts bits/sec

The overall circuit power for coding and precoding the vectors by using

where P_COD + P_DEC gives the total amount of power consumption for coding and decoding processes by consideration of required bandwidth which gives spectrum efficiency w.r.to number of base station antennas.

G.Backhaul

As signal transmission can be done in bw UE and BS and from Bs to UE hence there exists the requirement of transferring the uplink/downlink data in b/w BS and Network

So the transferring process can be done by backhaul. The power consumption of the backhaul is given by sum of two parts i.e first is fixed power Pfix and the second is average data transmitted or bandwidth. So the power consumption for backhaul Pbh is given by the sum of uplink and downlink

$$Pbh = P_t = P_{tx}^{(Ul)} + P_{tx}^{(Dl)} + P_{CP} *\ \ Pbt\ watts$$

Where, Pbt is the backhaul traffic power

The power consumption for backhaul process can be calculated by

$$P_B H_M R = P_B T * B * sumSE_M R$$

where LHS is power required for backhaul transformation process w.r.to number of base station antennas. And RHS shows P_BT is the power required to calculate the transmission bandwidth (no of packets transmitted) w.r.to calculation of total available B.W of product term B which gives the spectral efficiency.

III. DEEP LEARNING

The spatial local correlation by enforcing a local connectivity pattern can be applied among the neurons of adjacent layers to exploit the recent and popular conventional neural networks (CNNs) for the encoder and decoder design. The overview of the proposed DL architecture, named DLEE. The real and imaginary parts of H being the input, the first layer of the encoder being the convolutional layer. This layer uses kernels with dimensions of 3 X 3 to generate two feature maps for the spatial correlation of the antenna patterns using MMSE, ZF and MR and DLEE. To generate a codeword s, which is a real-valued vector of size M, we reshape the feature maps into a vector and use a fully connected layer after the convolution layer. The first two layers mimic the projection of Compressive Sensing(CS) 's' and serve as encoders. However, in contrast to random projections in CS, DLEE attempts to translate the extracted feature maps into a codeword. The first layer of the decoder is a fully connected layer that considers s as input and outputs two matrices of size NcNt , after we obtain the codeword s, we use several layers (as a decoder) to map it back into the channel matrix H. This NcNt serves as an initial estimate of the real and imaginary parts of H. The initial estimate is then fed into several "RefineNet units" that continuously refine the reconstruction. The rectified linear unit (ReLU), ReLU(x) = max(x, 0), is used as the activation function, and we introduce batch normalization to each layer. In RefineNet unit, the first layer is the input layer. All the remaining 3 layers use 3×3 kernels. Each RefineNet unit consists of four layers. The second and third layers generate 8 and 16 feature maps, respectively, and the final layer generates the final reconstruction of H. Using appropriate zero padding, the feature maps produced by the three convolutional layers are set to the same size as the input channel matrix size $N_c \times N_t$.

Two features of a RefineNet (RN) unit is that the output size is equal to the channel matrix size [8], [9]. To reduce dimensionality, like we down-sample nearly all conventional implementations of CNNs involve pooling layer. In contrast to traditional experiments, our target is refinement rather than dimensionality reduction. Second, in the RefineNet unit, we introduce resolving shortcut connections that directly pass data flow to upper layers which is inspired by the deep Residual Network [10], [11], which avoids the vanishing gradient problem caused by multiple stacked non-linear transformations.

Experiments reveal that two RefineNet units produce good performance. Once the channel matrix has been refined by a series of RefineNet units, the channel matrix is input into the final convolutional layer, and the sigmoid function is used to scale values to the [0, 1] range.

To train DLEE, we use end-to-end learning for all the kernel and bias values of the encoder and decoder. This training procedure differs from the two-step approach used in [10]. Notably, the input and output of DLEE are normalized channel matrices, whose elements are scaled in the [0, 1] range. The set of parameters is updated by the ADAM algorithm. The loss function is the mean squared error (MSE). The architecture of DLEE is a decoder output. The difference between the obtained channel and original H is measured by a normalized MSE as

$$NMSE = E\{\|H - \hat{H}\|_2^2/\|H\|_2^2\}$$

We then calculate the cosine similarity and then using NMSE and Rho we find the output of the normalized channel matrices back to their original levels.

IV. NUMERICAL RESULTS

To generate the training and testing samples, we create a channel matrix through the COST 2100 channel model [13] the outdoor rural scenario at the 300MHz band. All parameters follow their default setting in [13]. The BS is positioned at the center of a square area with lengths of 20 and 20m for indoor and outdoor scenarios, respectively, whereas the UEs are randomly positioned in the square area per sample.

We use the ULA with $N_t = 200$ antennas at the BS and $N_c = 1024$ subcarriers. We compare the throughput and Energy consumed for the three techniques suh MMSE, ZF and MR. When transforming the channel matrix into the angular-delay domain, we retain the first 32 rows of the channel matrix. That is, H is 32 × 32 in size. The training, validation, and testing sets contain 100,000, 30,000, and 20,000 samples, respectively. All testing samples are excluded from the training and validation samples. The epochs, learning rate, and batch size are set as 1000, 0.001, and 100, respectively.

TABLE I. *MAXIMAL EE FOR ALL SCHEMES WITH RESPECT TO THE M & K RATIOS*

Scheme	(M,K)	Maximal EE (Mbit/Joule)	Area throughput (Gbit/s/km^2)	PC (W)
M-MMSE	(60, 20)	44.00	17.33	24.62W
RZF	(90, 30)	39.33	20.97	33.34W
MR	(70, 20)	20.14	8.3	25.75W
DLEE	(50, 20)	45.13	15.2	24.12W

We compare DLEE with three state-of-the-art methods, namely, MMSE, ZF, MR with our proposed DLEE and is as shown in Table 1 [12]. Among these algorithms, ZF provides the bottom line result of the energy efficiency problem by considering only the simplest sparsity prior.

We also provide the corresponding results for DLEE, which only learns to recover CSI from CS measurements. The architecture of DLEE is identical to that of the decoder of ZF. From the experiments done we find that by applying deep learning strategies we were able to recover better patterns from the channel matrices.

Retrieval Number: F13380986S319/2019©BEIESP
DOI:10.35940/ijeat.F1338.0986S319
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

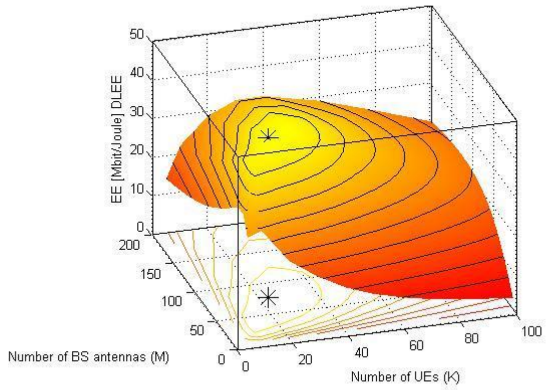

Figure 2 : EE per cell as a function of M and K with DLEE.

The EE per cell when (M,K) value is (50,20) we get the efficiency as 45.13 Mbit/Joule as shown in Figure 2. The amount of energy efficiency possible using our proposed DLEE is as shown and it shows that there is a minimum of 12% increase in energy efficiency when compared to average of all other available traditional strategies.

V. CONCLUSION

We used DL in DLEE, a novel power efficiency mechanism. DLEE performed well at varied M and K densities and reduced time complexity. We believe that its reconstruction quality can be further improved by applying advance DL technology, and we hope this paper encourages future research in this direction.

REFERENCES

1. Gandotra, P., Jha, R. K., & Jain, S. (2017). Green communication in next generation cellular networks: a survey. IEEE Access, 5, 11727-11758.
2. Beck, P. Stoica, and J. Li, "Exact and approximate solutions of source localization problems," IEEE Trans. Signal Process., vol. 56, no. 5, pp. 1770–1778, May 2008.
3. K. Do˘ganc¸ay, "Passive emitter localization using weighted instrumental variables," Signal Process., vol. 84, no. 3, pp. 487–497, 2003.
4. M. Gavish and A. J. Weiss, "Performance analysis of bearing-only target location algorithms," IEEE Trans. Aerosp. Electron. Syst., vol. 28, no. 3, pp. 817–828, Jul. 1992.
5. J.Weiss, "Direct position determination of narrowband radio frequency transmitters," IEEE Signal Process. Lett., vol. 11, no. 5, pp. 513–516, May 2004.
6. Kwihoon Kim, Joohyung Lee,"Deep Learning based Pilot Allocation Scheme(DL-PAS) for 5G Massive MIMO System," IEEE Communication Process., DOI 10.1109LCOMM.2018.2803054.
7. Emil Bjornson,"Optimal Design of Energy-Efficient Multi-User MIMO Systems: Is Massive MIMO the Answer?," IEEE Transactions on Wireless Communications., vol. 14, no. 6, June 2015.
8. Lohit et al. (2017). Convolutional Neural Networks for Non-Iterative Reconstruction of Compressively Sensed Images. [Online]. Available: http://arxiv.org/abs/1708.04669
9. Mousavi, G. Dasarathy, and R. G. Baraniuk. (2017). DeepCodec: Adaptive Sensing and Recovery via Deep Convolutional Neural Networks. [Online]. Available: http://arxiv.org/abs/1707.03386
10. Telagam, N., Kandasamy, N., Nanjundan, M., & Arulanandth, T. S. (2017). Smart Sensor Network based Industrial Parameters Monitoring in IOT Environment using Virtual Instrumentation Server. International Journal of Online Engineering (iJOE), 13(11), 111-119.
11. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. CVPR, Las Vegas, NV, USA, 2016, pp. 770–778.
12. Emil Björnson, Jakob Hoydis and Luca Sanguinetti (2017), "Massive MIMO Networks: Spectral, Energy, and Hardware Efficiency", Foundations and TrendsR in Signal Processing: Vol. 11, No. 3-4, pp 154–655. DOI: 10.1561/2000000093.
13. L. Liu et al., "The COST 2100 MIMO channel model," IEEE Wireless Commun., vol. 19, no. 6, pp. 92–99, Dec. 2012.
14. B.Hari Krishna , J.Surendiran, C. Ashok Kumar," Smart and connected health care system using IoT", International Journal of Pure and Applied Mathematics Volume 119 No. 12 2018, 14405-14412 ISSN: 1314-3395.
15. Kandasamy, N., Telagam, N., VR, S. R., & Arulananth, T. S. (2017). Simulation of Analog Modulation and Demodulation Techniques in Virtual Instrumentation and Remote Lab. International Journal of Online Engineering (iJOE), 13(10), 140-147.
16. J.Surendiran1 , R.Sridhar2 ,Rajiv arya 3," GPRS BASED MONITORING SYSTEM USING IMAGE SENSOR", International Journal of Pure and Applied Mathematics Volume 119 No. 12 2018, 14387-14392 ISSN: 1314-3395.
17. Anitha, P., T. R. V. Anandharaja, and C. Murugalakshmi. "Regularization and its effects on generalization for character recognition." In Electrical, Instrumentation and Communication Engineering (ICEICE), 2017 IEEE International Conference on, pp. 1-3. IEEE, 2017.
18. N. Nithiyanandam, K. Venkatesh, M. Rajesh, Transfer The Levels Of The Monitored Carbon, Nitrogen Gases From The Industries, International Journal of Recent Technology and Engineering, Volume-7 Issue-6S3 April, 2019.
19. Sivanesh Kumar, A., Brittoraj, S., Rajesh, M., Implementation of RFID with internet of things, Journal of Recent Technology and Engineering, Volume-7 Issue-6S3 April, 2019.
20. Rajesh, M., Sairam, R., Big data and health care system using mlearningJournal of Recent Technology and Engineering, Volume-7 Issue-6S3 April, 2019.
21. Rajesh, M., and J. M. Gnanasekar. "Path Observation Based Physical Routing Protocol for Wireless Ad Hoc Networks." Wireless Personal Communications 97.1 (2017): 1267-1289.

Retrieval Number: F13380986S319/2019©BEIESP
DOI:10.35940/ijeat.F1338.0986S319
Journal Website: www.ijeat.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)