A Study on Model Driven DL based Massive MIMO Detection

Saleem Ahmed, Sooyoung Kim*

IT Convergence Research Center, Jeonbuk National University, Jeonju, Korea

saleem.ahmed@jbnu.ac.kr, *sookim@jbnu.ac.kr

모델 기반 심층 학습 Massive MIMO 검출에 관한 연구

살림 아흐메드, 김수영*
전북대학교 IT 융합연구센터

Abstract

Signal detection plays an important role in massive multiple-input multiple-output (MIMO) systems. In this research work, a novel model driven deep learning (DL) method is proposed for massive MIMO detection. We propose to enhance the performance of approximate message passing (AMP) algorithm by adding trainable parameters. A popular strategy to avoid divergence of AMP algorithm is damping, which is a simple and very low complex solution. The suitable choice of values for damping parameters are important for overall performance of AMP algorithm. We propose to learn the damping parameters which can enhance the performance of the AMP algorithm.

Ⅰ. Introduction

Modern wireless communication systems employing massive multiple-input multiple-output (MIMO) systems can provide higher data rates and can improve the spectral efficiency [1]. The signal detection plays an important role in massive MIMO systems. A practical detector for massive MIMO systems should not only achieve acceptable error rate performance but also low computational complexity. However, achieving a reasonable balance between the complexity and detection performance is challenging.

The linear detectors such as zero forcing (ZF) and minimum mean square error (MMSE) detector exhibit poor bit error rate (BER) performance, while optimal maximum likelihood (ML) detection is computationally intractable. Deep unfolding also known as model driven deep learning (DL) unfolds the iterations of iterative algorithm into sequentially connected layers of neural networks to efficiently solve a range of tasks including MIMO detection problem [2]. Several model driven DL based MIMO detection methods such as orthogonal approximate message passing network (OAMP-Net) and DetNet can produce reasonable performance [2][3]. However, both methods exhibit higher computational complexity. The computational complexity of OAMP-Net is higher due to matrix inverse operation in each iteration and the DetNet method learns large number of parameters which makes it computationally complex.

The LArge MIMO AMP (LAMA) algorithm is suitable for massive MIMO systems due to lower complexity and better error rate performance [4]. Furthermore, iterative nature of LAMA makes it suitable choice for model driven DL architecture. We propose an efficient massive MIMO detection method by employing model driven DL method. We proposed to enhance the performance of LAMA algorithm by learning the damping parameters. The damping factor in each iteration of LAMA algorithm can enhance the performance by finding the suitable values of damping factors through DL.

Ⅱ. Proposed Model Driven DL based MIMO Detection

Consider a MIMO system where the transmitter is equipped with M antennas and the receiver is equipped with N antennas. The received signal can be represented in complex domain as:

$$\mathbf{y}_c = \mathbf{H}_c \mathbf{s}_c + \mathbf{n}_c \,, \tag{1}$$

where $\mathbf{y}_c$ is the received signal and $\mathbf{s}_c$ contains the transmitted symbols drawn from complex constellation C. The $\mathbf{H}_c$ represents the channel matrix and $\mathbf{n}_c$ is the Gaussian noise with variance N_0.

The complex signal model can be converted into a real equivalent model as:

$$\begin{bmatrix} \Re(\mathbf{y}_c) \\ \Im(\mathbf{y}_c) \end{bmatrix} = \begin{bmatrix} \Re(\mathbf{H}_c) & -\Im(\mathbf{H}_c) \\ \Im(\mathbf{H}_c) & \Re(\mathbf{H}_c) \end{bmatrix} \begin{bmatrix} \Re(\mathbf{s}_c) \\ \Im(\mathbf{s}_c) \end{bmatrix} + \begin{bmatrix} \Re(\mathbf{n}_c) \\ \Im(\mathbf{n}_c) \end{bmatrix}, \quad (2)$$

$$\mathbf{y} = \mathbf{Hs} + \mathbf{n}, \quad (3)$$

where $\mathbf{y} \in \mathbb{R}^{2N\times 1}$, $\mathbf{H} \in \mathbb{R}^{2N\times 2M}$, $\mathbf{s} \in \mathbb{R}^{2M\times 1}$ and $\mathbf{n} \in \mathbb{R}^{2N\times 1}$. In the real equivalent model, the elements of $\mathbf{s}$ are chosen from real equivalent constellation set.

The AMP algorithm estimating various parameters in each iteration are given as:

$$\begin{aligned}
&\textbf{for } l=1 \text{ to } L \textbf{ do} \\
&\quad \hat{\mathbf{s}}^{l+1} = \mathrm{F}(\mathbf{z}^l, \rho^l \mathbf{g}) \\
&\quad \tau^{l+1} = \mathrm{G}(\mathbf{z}^l, \rho^l \mathbf{g}) \\
&\quad \overline{\tau}^{l+1} = \theta_\tau (\tau^{l+1} \mathbf{g}^T) + (1-\theta_\tau)\overline{\tau}^{l} \\
&\quad \rho^{l+1} = \theta_\rho (\overline{\tau}^{l+1} + N_0)\mathbf{g} + (1-\theta_\rho)\rho^l \qquad (4) \\
&\quad \mathbf{v}^l = \frac{\overline{\tau}^{l+1}}{\tau^l + N_0}(\mathbf{z}^l - \hat{\mathbf{s}}^l) \\
&\quad \mathbf{z}^{l+1} = \tilde{\mathbf{y}}^{MF} + \tilde{\mathbf{G}}\hat{\mathbf{s}}^{l+1} + \mathbf{v}^l \\
&\textbf{end for}
\end{aligned}$$

In (4), $\mathrm{F}(\mathbf{z}^l, \rho^l \mathbf{g})$ and $\mathrm{G}(\mathbf{z}^l, \rho^l \mathbf{g})$ are mean and variance estimates of signal estimate $\mathbf{z}^l$. Furthermore, $\overline{\tau}^l$ and ρ^l are the damping parameters estimated for signal mean and variance, respectively. The Onsager term is represented as $\mathbf{v}^l$. At the end of each loop the new signal estimation is calculated $\mathbf{z}^{l+1}$. The terms $\tilde{\mathbf{y}}^{MF}$ and $\tilde{\mathbf{G}}$ are normalized matched filtering, and normalized gram matrix, respectively. Furthermore, $\mathbf{g} = \mathrm{diag}(\mathbf{G})$. The θ_τ and θ_ρ are the damping constants. The damped message to be passed in iteration l is computed as a weighted average of the message in iteration $l-1$ and the message computed at the lth iteration with a damping factor ranging in [0, 1].

The convergence speed of AMP algorithm plays important role in its performance and complexity. Since the fast converging algorithm require less iteration for desired performance. The damping factors can make the algorithm converge faster. There has been several studies on damping factors for various AMP algorithms. However, optimization of damping constants through DL is not been performed. Instead of fixed damping constants, we proposed to optimize them through DL algorithm by employing model driven DL approach. We propose to generate L different values for both damping parameters through DL. By learning the damping contants θ_τ and θ_ρ, the error rate performance of LAMA method can be improved by 1dB-2dB.

Ⅲ. Conclusion

In this paper, we proposed a novel deep learning based AMP algorithm. We proposed a method which can optimize the damping constants of the LAMA algorithm. The deep unfolding architecture is employed to learn the damping parameters by unfolding the LAMA algorithm iterations into DL layers. The proposed method can reduce the computational complexity of LAMA algorithm by achieving the desired error rate performance with less number of iterations.

ACKNOWLEDGMENT

This work was supported by National Research Foundation of Korea with Brain Pool (BP) program, Grant No. RS-2023-00217283.

REFERENCES

[1] Strinati E. C., Belot D., Falempin A., and Dore, J. B. "Toward 6G: From new hardware design to wireless semantic and goal-oriented communication paradigms," arXiv preprint arXiv:2107.01019, 2021.

[2] He H., Wen C. K, Jin S., and Li G. Y., "A model-driven deep learning network for MIMO detection," in Proc. IEEE Global Conf. Signal Inf. Process (GlobalSIP), pp. 584-588, Anaheim, CA, USA, Sep. 2018.

[3] N. Samuel T. Diskin, and A. Wiesel, "Deep MIMO detection," in Proc. 18th IEEE Int. Workshop on Signal Processing Advances in Wireless Commun.(SPAWC), pp. 1-5, Sapporo, Japan, 2017.

[4] C. Jeon, R. Ghods, A. Maleki, and C. Studer, "Optimality of large MIMO detection via approximate message passing," in Proc. IEEE Int. Symp. Inf. Theory (ISIT), pp. 1227-1231, Hong Kong, China, June, 2015.