Received: 16 December 2022 | Revised: 3 May 2023 | Accepted: 15 May 2023

DOI: 10.4218/etrij.2022-0462

FEATURED ARTICLE

ETRI Journal WILEY

Enhanced deep soft interference cancellation for multiuser symbol detection

Jihyung Kim[1] | **Junghyun Kim**[2] | **Moon-Sik Lee**[1]

[1]Satellite Communication Research Division, Electronics and Telecommunications Research Institute, Daejeon, Republic of Korea

[2]Department of Artificial Intelligence, Sejong University, Seoul, Republic of Korea

Correspondence
Junghyun Kim, Department of Artificial Intelligence, Sejong University, Seoul, Republic of Korea.
Email: j.kim@sejong.ac.kr

Funding information
MSIT, Grant/Award Number: 2021-0-00794

Abstract

The detection of all the symbols transmitted simultaneously in multiuser systems using limited wireless resources is challenging. Traditional model-based methods show high performance with perfect channel state information (CSI); however, severe performance degradation will occur if perfect CSI cannot be acquired. In contrast, data-driven methods perform slightly worse than model-based methods in terms of symbol error ratio performance in perfect CSI states; however, they are also able to overcome extreme performance degradation in imperfect CSI states. This study proposes a novel deep learning-based method by improving a state-of-the-art data-driven technique called deep soft interference cancellation (DSIC). The enhanced DSIC (EDSIC) method detects multiuser symbols in a fully sequential manner and uses an efficient neural network structure to ensure high performance. Additionally, error-propagation mitigation techniques are used to ensure robustness against channel uncertainty. The EDSIC guarantees a performance that is very close to the optimal performance of the existing model-based methods in perfect CSI environments and the best performance in imperfect CSI environments.

KEYWORDS
channel state information, deep learning, interference cancellation, joint-symbol detection, multiuser system

1 | INTRODUCTION

Multiuser multiple-input multiple-output (MU-MIMO) technology is a key technology in modern wireless communication systems. MU-MIMO has received considerable attention because it can provide high throughput and spectral efficiency. However, several challenges are found in realizing the attractive benefits of MU-MIMO in practice. One of these is the simultaneous detection of multiuser signals in an environment with interference and incomplete channel-state information (CSI). Existing multiuser symbol detection algorithms, such as maximum a posteriori (MAP), which is based on optimal rules for joint recovery of symbols for all users, increase their computational complexity to an impossible level as the number of users increases. In addition, the technique that recovers each symbol individually while considering and processing the remaining symbols as noise can

reduce the complexity but is limited by the decrease in throughput. In imperfect CSI environments, serious performance degradation may occur.

The successive interference cancellation [1] technique removes interference from a received signal by demodulating and/or decoding the desired information and by using this information in conjunction with a channel estimate. By utilizing this technique for multiuser symbol detection, performance degradation can be alleviated and complexity can be reduced. Specifically, the technique sequentially removes hard decision values of user symbols from the received signal detected previously; however, error propagation may occur in this process. To alleviate this problem, a method [2] that repeatedly uses soft-decision values can be employed. Although this method significantly mitigates error propagation with controllable complexity and reduces the performance gap compared with the optimal MAP performance, it still suffers from severe performance degradation in imperfect CSI environments, which can occur at high speeds.

Recently, significant progress has been achieved in deep learning (DL) [3], which has been applied in many areas. In particular, the remarkable success of DL in challenging games, such as Go [4], Starcraft [5], and computer vision [6], has transformed the existing model-driven mindset into a data-driven mindset. Data-driven methods have two advantages compared with model-based approaches. First, data-driven techniques can work in scenarios where modeling is unknown because the system is difficult to interpret. Second, if the basic modeling method is known but the system is very complex and has not been sufficiently analyzed, data-driven algorithms can be used to extract features from the observed data.

Many researchers in the field of wireless communication have attempted to apply DL to communication technologies [7–10]. These attempts have included the use of neural networks to design channel estimators [11], channel denoisers [12], channel decoders [13, 14], and end-to-end transceivers [15–18] to manage multiuser interference [19–22] and design multiuser symbol detectors [23–25]. In this study, we focused on the use of DL for joint-symbol detection in multiuser MIMO systems. DL-based detectors are anticipated to be a highly promising alternative to MAP detectors—which exhibit an exponential increase in complexity as the number of users increases—because these methods have linearly increasing complexity with the number of users. Conventional MAP detectors suffer from significant performance degradation in environments with imperfect CSI. This is because the posterior probability used in the MAP detector does not adequately account for the impact of the uncertain channel information. However, DL-based detectors can prevent performance degradation by improving the generalization performance if sufficient training data are provided, even in situations wherein only uncertain channel information is available. In conclusion, the ability to model wireless communication environments ensures the generation of infinite training data, which is advantageous for ensuring a robust performance in imperfect scenarios [26, 27].

The recently proposed deep soft interference cancellation (DSIC) technique [25] was reported to demonstrate a performance close to that of the existing iterative soft interference cancellation (ISIC) [2] and exhibited outstanding performance even in imperfect CSI environments. By improving the DL-based technique, we propose a new model with near-optimal performance in both perfect and imperfect CSI environments. The main contributions of this study are as follows:

- The proposed scheme uses a new neural network structure. To predict each user symbol in an interference environment, one of the state-of-the-art schemes, DSIC, uses a basic deep neural network (DNN) model. However, to obtain high efficiency and performance improvement, herein, we use a bidirectional long short-term memory (biLSTM) model, which is more suitable for time series data.
- The proposed scheme predicts each user symbol in each iteration sequentially using the updated soft estimates in the current iteration, instead of predicting in parallel using the soft estimates in the previous iteration. This method not only speeds up convergence but also ensures the best performance within a limited processing time period.
- Because the proposed scheme operates in a sequential manner, to address the problem of error propagation that might occur in sequential methods, we applied mitigation techniques.

The remainder of this paper is structured as follows. Section 2 presents our system model and reviews related works. Section 3 proposes an enhanced DSIC (EDSIC) that overcomes the limitations of DSIC. Section 4 outlines the experimental results of the proposed EDSIC, and Section 5 provides the concluding remarks.

2 | SYSTEM MODEL AND RELATED WORKS

2.1 | System model

We consider a multiuser uplink MIMO system with K users demonstrating a single transmit antenna and a base station (BS) equipped with an array of N receiving

antennas. For a linear channel with additive white Gaussian noise (AWGN), the signals received at the BS are represented as

$$\boldsymbol{y} = \boldsymbol{h}\boldsymbol{s} + \boldsymbol{w}, \tag{1}$$

where $\boldsymbol{h} \in \mathcal{R}^{N \times K}$ is a channel matrix, $\boldsymbol{s} \triangleq [s_1, s_2, ..., s_K]^{\mathrm{T}}$ is a symbol vector of mutually independent elements drawn from a discrete constellation set $\mathcal{S} \triangleq \{\alpha_1, \alpha_2, ..., \alpha_M\}$, and $\boldsymbol{w} \in \mathcal{R}^{N \times 1}$ is a noise vector of a zero-mean multivariate Gaussian with covariance $\sigma_w^2 I_N$. In this study, we focused on real-valued channel scenarios, and the system model could be extended in a straightforward manner to complex-valued channel scenarios. In addition, the time index is omitted for simplicity. Figure 1 shows the system model.

For a discrete memoryless channel, the received signal $\boldsymbol{y}$ is represented by the stochastic mapping of the transmitted symbols $\boldsymbol{s}$. Thus, we can express their relationship using the conditional probability measure $p_{\boldsymbol{y}|\boldsymbol{s}}(\cdot|\cdot)$. Subsequently, we focus on the problem of individual reconstruction of the transmitted symbols $\boldsymbol{s}$ by each user from the received signals $\boldsymbol{y}$. In ideal environments, the optimal symbol detector is the MAP detector, which minimizes error probability. To describe this in more detail, let $p_{\boldsymbol{s}|\boldsymbol{y}}(\cdot|\cdot)$ be the posterior probability of $\boldsymbol{s}$ given $\boldsymbol{y}$. The MAP detector selects a set of symbols that maximizes the posterior probability; that is,

$$\hat{\boldsymbol{s}}_{MAP} \triangleq \arg\max_{\boldsymbol{s} \in \mathcal{S}^K} p(\boldsymbol{s}|\boldsymbol{y}). \tag{2}$$

When the transmitted symbols demonstrate equal probabilities, the MAP detector is equivalent to the maximum likelihood detector.

Although the MAP detector ensures optimal performance in ideal environments, it is still associated with several limitations: it requires an exhaustive search M^K, which is computationally infeasible, especially when K is large, and it also requires accurate knowledge of channels that are sometimes difficult to obtain in real-world communication environments.

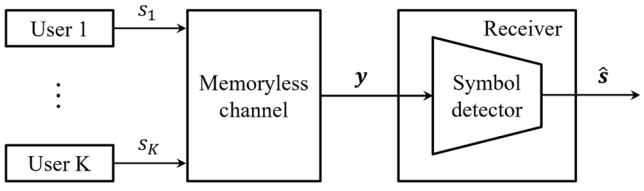

FIGURE 1 An uplink multiuser multiple-input multiple-output system model.

2.2 | ISIC

Interference cancellation is an effective strategy for implementing multiuser symbol detection with affordable computational complexity. This technique performs joint symbol detection iteratively by reconstructing each symbol based on the channel output and estimates of the remaining interfering symbols. Specifically, by removing the effects of the estimated interference, the algorithm uses the knowledge of the channel to facilitate the recovery of each symbol in the received signal.

To alleviate the phenomenon of error propagation inevitably exhibited by traditional interference cancellation techniques, a detector was proposed in [2] using soft values for interference cancellation. This scheme, referred to as ISIC, operates iteratively by combining multilevel interference cancellation with soft decisions. In each iteration, for every kth user, where $k \in \mathcal{K} \triangleq \{1,2,...,K\}$, an estimate of the conditional probability of s_k given $\boldsymbol{y}$ is calculated using the estimates of transmitted symbols $\{s_l\}_{l \in \mathcal{K} \setminus k}$ from other users obtained in the previous iteration and treated as interference for the kth user. In addition, the reliability of the conditional probability estimates is improved by repeating this procedure such that the symbol detector can accurately reconstruct each transmitted symbol from the output of the last iteration.

We assume that the ISIC detector comprises Q iterations. In each iteration, the detector generates K probability vectors $\hat{\boldsymbol{p}}_k^{(q)} \quad \mathcal{R}^M$, $k \in \mathcal{K}$, $q \in \{1,2,...,Q\}$. The probabilities of $\hat{\boldsymbol{p}}_k^{(q)}$ are estimates of the conditional probability corresponding to $\boldsymbol{s}_k$ for each possible symbol in $\mathcal{S}$ when the received signal $\boldsymbol{y}$ is given, and the interfering symbols $\{s_l\}_{l \in \mathcal{K} \setminus k}$ are distributed via $\{\hat{\boldsymbol{p}}_l^{(q-1)}\}_{l \in \mathcal{K} \setminus k}$. Subsequently, each iteration comprises two stages performed in parallel for each user. The first stage is interference cancellation and the second stage is soft decoding. For the qth iteration and kth user, the interference cancellation stage respectively calculates the expected values and variances of $\{s_l\}_{l \in \mathcal{K} \setminus k}$ by using the following two equations,

$$e_l^{(q-1)} = \sum_{\alpha_m \in \mathcal{S}} \alpha_m \cdot \left(\hat{\boldsymbol{p}}_l^{(q-1)}\right)_{\alpha_m}, \tag{3}$$

and

$$v_l^{(q-1)} = \sum_{\alpha_m \in \mathcal{S}} \left(\alpha_m - e_l^{(q-1)}\right)^2 \cdot \left(\hat{\boldsymbol{p}}_l^{(q-1)}\right)_{\alpha_m}, \tag{4}$$

Denoting h_l as the lth column of $\boldsymbol{h}$, the channel output following interference cancellation can be represented as

$$z_k^{(q)} \triangleq \boldsymbol{y} - \sum_{l \in \mathcal{K} \setminus k} h_l e_l^{(q-1)}. \tag{5}$$

In the soft-decoding stage, the estimated conditional probability is calculated as

$$\left(\hat{\boldsymbol{p}}_k^{(q)}\right)_{\alpha_m} \triangleq p(s_k = \alpha_m | \boldsymbol{y}) = \frac{p(z_k^{(q)} | \alpha_m)}{\sum_{\alpha_{m'}} p(z_k^{(q)} | \alpha_{m'})} = \frac{\exp\left(-\frac{1}{2}\left(z_k^{(q)} - h_k\alpha_m\right)^T \Sigma_{W_k^{(q)}}^{-1}\left(z_k^{(q)} - h_k\alpha_m\right)\right)}{\sum_{\alpha_{m'} \in \mathcal{S}} \exp\left(-\frac{1}{2}\left(z_k^{(q)} - h_k\alpha_{m'}\right)^T \Sigma_{W_k^{(q)}}^{-1}\left(z_k^{(q)} - h_k\alpha_{m'}\right)\right)}, \tag{6}$$

where

$$\Sigma_{W_k^{(q)}}^{-1} = \sigma_n^2 I_N + \sum_{l \in \mathcal{K} \setminus k} v_l^{(q-1)} h_l h_l^T. \tag{7}$$

Each symbol is detected after the final iteration Q by using a symbol in $\mathcal{S}$ that maximizes the estimated conditional probability:

$$\hat{s}_k = \arg\max_{\alpha_m \in \mathcal{S}} \left(\hat{\boldsymbol{p}}_k^{(Q)}\right)_{\alpha_m}. \tag{8}$$

The entire process of ISIC described above is summarized in Algorithm 1.

Algorithm 1 Iterative soft interference cancellation (ISIC)

1: Input: Received signals $\boldsymbol{y}$.
2: Initialization: Set $q = 1$, and initialize the probabilities $\{\hat{\boldsymbol{p}}_k^{(0)}\}_{k=1}^K$.
3: **for** $q = 1, 2, \ldots, Q$ **do**
4: For each $k \in \mathcal{K}$ compute $e_k^{(q-1)}$ and $v_k^{(q-1)}$, obtained by (3) and (4), respectively.
5: For each $k \in \mathcal{K}$ compute $\boldsymbol{z}_k^{(q)}$ and $\hat{\boldsymbol{p}}_k^{(q)}$, obtained by (5) and (6), respectively.
6: **end for**
7: Output: A set of hard-detected symbols $\hat{\boldsymbol{s}}$ obtained by (8).

The ISIC detector is a model-based multiuser symbol detector that significantly reduces complexity via the above iterative operations. However, because the symbols are estimated using channel information, they still suffer from severe performance degradation in imperfect CSI environments, such as cases wherein MAP detectors are used.

2.3 | Deep hard interference cancellation (DHIC)

Successive interference cancellation is another effective method used to detect multiuser symbols. To avoid confusion, we refer to this method as hard interference cancellation (HIC). Using this technique, the receiver first detects a symbol with the maximum channel gain, subtracts the detected symbol from the received signal, and then sequentially detects the remaining symbols. Following the basic concept of HIC, [23, 24] proposed DL approaches for downlink signal detection in MU-MIMO-NOMA. In both [23] and [24], only the downlink environment was considered, particularly in [23], where only two users were assumed. In [24], it was extended to K users. Therefore, we modified the model by considering the uplink environments for K users. We refer to this technique as DHIC. Figure 2 describes the structure of DHIC.

The DHIC sequentially detects user symbols using as many DNN blocks as possible. Each DNN block consisted of fully connected neural networks. A single DNN block was used for each user. The output of each DNN block represents the detected symbol for the corresponding user, and the input is the received signal after all previously detected symbols are subtracted. This operation is performed sequentially. Therefore, the input for the current DNN block can be obtained by subtracting the output of the previous DNN block from the input. In other words, starting with the user with the highest channel quality, the process sequentially detects the transmitted symbols of a single user at each stage using separate DNN blocks. Algorithm 2 summarizes the DHIC procedure.

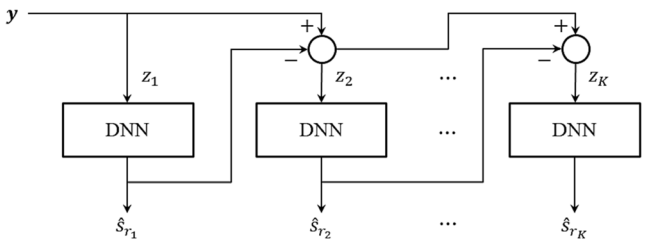

FIGURE 2 Structure of deep hard interference cancellation (DHIC)

Algorithm 2 Deep hard interference cancellation (DHIC)

1: Input: Received signal $\boldsymbol{y}$.
2: Initialization: Set $z_1 = \boldsymbol{y}$ and sort the indices of $\boldsymbol{s}$ in descending order of channel gains to $\{r_1, r_2, \dots, r_K\}$.
3: Estimate $\hat{s}_{r_1}$ from z_1 using the 1st DNN.
4: **for** $k = 2, 3, \dots, K$ **do**
5: Set $z_k = z_{k-1} - \hat{s}_{r_{k-1}}$.
6: Estimate $\hat{s}_{r_k}$ from z_k using the kth DNN.
7: **end for**
8: Output: A set of hard-detected symbols $\hat{\boldsymbol{s}}$.

2.4 | DSIC

A DL-based soft interference cancellation scheme, referred to as DSIC, was recently proposed in [25]. It replaces the model-based computations in ISIC by learning the DNNs. Neural networks predict multiuser symbols by learning only the relationship between the input and output data without prior knowledge of channel information. Figure 3 illustrates the DSIC architecture. Two types of DSIC were identified. The first type uses end-to-end training, whereas the second uses sequential training. In this study, we renamed the sequential training in [25] as iterative training to prevent confusion with our fully sequential method.

The end-to-end scheme simultaneously updates the parameters $\boldsymbol{\theta}$ of Q neural network blocks, each consisting of K DNN subblocks, and thus exhibits a cost function via

$$\mathcal{L}(\boldsymbol{\theta}) = \frac{1}{J}\sum_{j=1}^{J}\sum_{k=1}^{K} -\log \hat{\boldsymbol{p}}_k^{(Q)}\left(\boldsymbol{y}_j, \left\{\hat{\boldsymbol{p}}_{j,l}^{(0)}\right\}_{l\in\mathcal{K}\setminus k}, (\hat{\boldsymbol{s}}_j)_k; \boldsymbol{\theta}\right), \quad (9)$$

where $\hat{\boldsymbol{p}}_k^{(Q)}\left(\boldsymbol{y}_j, \left\{\hat{\boldsymbol{p}}_{j,l}^{(0)}\right\}_{l\in\mathcal{K}\setminus k}, (\hat{\boldsymbol{s}}_j)_k; \boldsymbol{\theta}\right)$ is the estimated conditional probability corresponding to the estimated symbol $(\hat{\boldsymbol{s}}_j)_k$ at the kth subblock when the jth training data sample $\boldsymbol{y}_j$ and initial probabilities $\left\{\hat{\boldsymbol{p}}_{j,l}^{(0)}\right\}_{l\in\mathcal{K}\setminus k}$ are given.

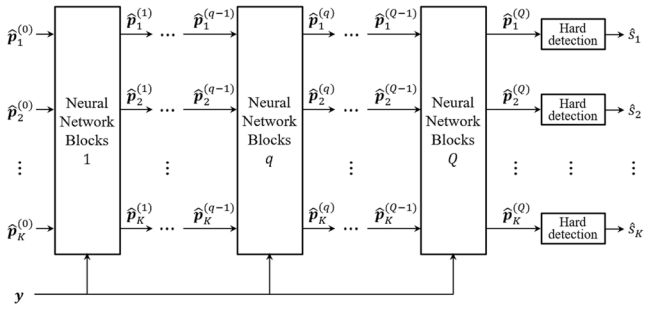

FIGURE 3 Deep soft interference cancellation (DSIC) structure.

If the initial probabilities are consistently selected from a uniform distribution, the term can be omitted from the input. Finally, J denotes the number of training data samples.

Compared with end-to-end training, which jointly learns all trainable parameters, iterative training uses the same set of data to tune neural network blocks individually with trainable parameters. This approach requires significantly fewer training data samples at the expense of performance degradation and potential delays resulting from iterative tasks. At the qth iteration, the cost function of the iterative training is represented according to

$$\mathcal{L}\left(\boldsymbol{\theta}_k^{(q)}\right) = \frac{1}{J}\sum_{j=1}^{J} -\log \hat{\boldsymbol{p}}_k^{(q)}\left(\boldsymbol{y}_j, \left\{\hat{\boldsymbol{p}}_{j,l}^{(q-1)}\right\}_{l\in\mathcal{K}\setminus k}, (\hat{\boldsymbol{s}}_j)_k; \boldsymbol{\theta}_k^{(q)}\right), \quad (10)$$

where $\hat{\boldsymbol{p}}_k^{(q)}\left(\boldsymbol{y}_j, \left\{\hat{\boldsymbol{p}}_{j,l}^{(q-1)}\right\}_{l\in\mathcal{K}\setminus k}, (\hat{\boldsymbol{s}}_j)_k; \boldsymbol{\theta}_k^{(q)}\right)$ is the currently estimated conditional probability corresponding to the estimated symbol $(\hat{\boldsymbol{s}}_j)_k$ at the kth subblock when jth training data sample $\boldsymbol{y}_j$ and the previously estimated probabilities $\left\{\hat{\boldsymbol{p}}_{j,l}^{(q-1)}\right\}_{l\in\mathcal{K}\setminus k}$ are given, and J is the number of training data samples.

In iterative training, the output of the previous iteration was used as the input data for the subsequent iteration, and served as a soft estimate. In particular, kth DNN of qth neural network block produces $\hat{\boldsymbol{p}}_k^{(q)}$ from the given $\hat{\boldsymbol{y}}$ and $\left\{\hat{\boldsymbol{p}}_l^{(q-1)}\right\}_{l\in\mathcal{K}\setminus k}$, which is depicted in Figure 4. The resulting procedure is summarized in Algorithm 3.

Algorithm 3 Deep soft interference cancellation (DSIC)

1: Input: Received signals $\boldsymbol{y}$.
2: Initialization: Initialize the probabilities $\{\hat{\boldsymbol{p}}_k^{(0)}\}_{k=1}^{K}$.
3: **for** $q = 1, 2, \dots, Q$ **do**
4: **for** $k = 1, 2, \dots, K$ **do**
5: Estimate $\hat{\boldsymbol{p}}_k^{(q)}$ from $\boldsymbol{y}$ and $\{\hat{\boldsymbol{p}}_l^{(q-1)}\}_{l\in\mathcal{K}\setminus k}$ using the (q,k)th DNN.
6: **end for**
7: **end for**
8: Output: A set of hard-detected symbols $\hat{\boldsymbol{s}}$, obtained by (8).

As mentioned in [25], the DSIC detector is substantially more robust to channel uncertainty than the ISIC and MAP detectors; however, in the case of perfect CSI, a considerable performance gap was found with ISIC as well as the MAP detectors.

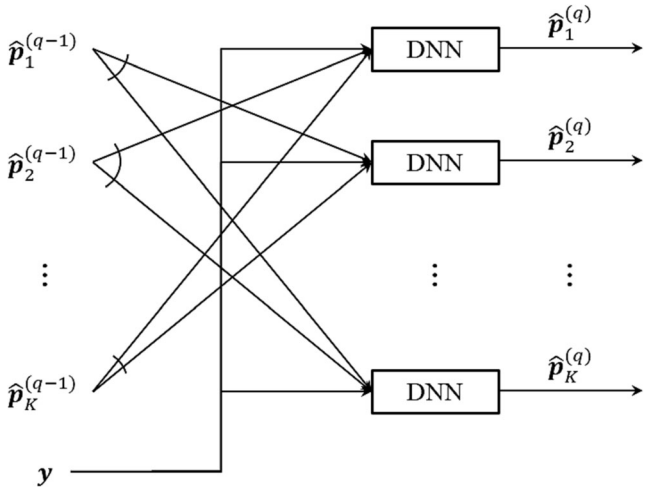

FIGURE 4 Detailed structure of the qth neural network block used for DSIC.

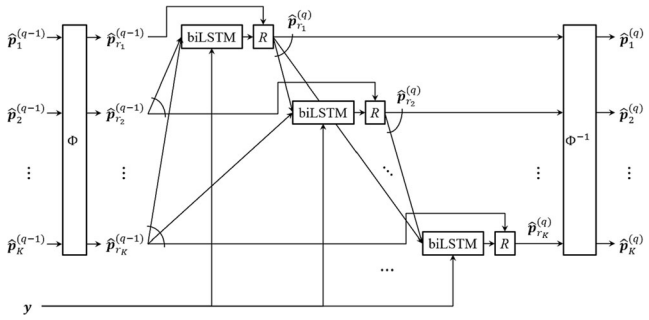

FIGURE 5 Detailed structure of the qth neural network block for the enhanced DSIC (EDSIC)

3 | EDSIC

We propose herein an enhanced scheme called EDSIC that improves the DSIC performance. Instead of DNN, we used biLSTM in DSIC for more effective learning. Instead of predicting soft estimates in parallel, as in DSIC, our EDSIC predicts soft estimates in a sequential manner. Using this concept, the convergence speed was accelerated to obtain accurate estimates. In other words, the performance can be optimized within a limited processing time. The basic structure of EDSIC is similar to that of DSIC in Figure 3. However, as shown in Figures 4 and 5, significant differences exist in the detailed structures.

The EDSIC detector involves three steps: sorting, estimation, and relaxation. In the first step, the user indices are sorted in descending order according to their channel quality. This preferentially predicts the symbols of users that are expected to be more accurate. In the second step, a soft estimate for each user is sequentially predicted using biLSTM. Figure 6 shows a detailed depiction of the biLSTM block. Unlike LSTM which has only forward

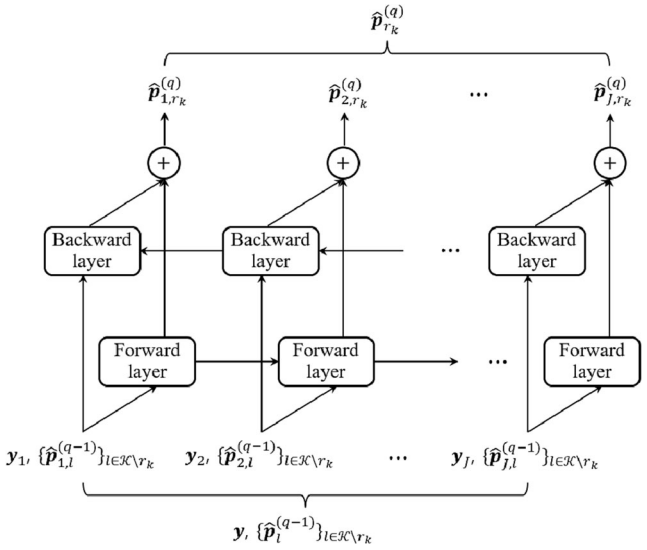

FIGURE 6 Detailed structure of biLSTM block in the qth neural network block for EDSIC.

layers, biLSTM performs the training simultaneously on both the forward and backward layers. During the testing phase, the same structure generates an output from a given input without updating any parameters and predicts the symbol values of each user. Finally, in the relaxation step, the soft estimate is updated by calculating the weighted sum of the predicted values from the current and previous iterations. Herein, we used a relaxation coefficient γ with a value between zero and one. The relaxation block is denoted by R in Figure 5. In Figure 5, Φ denotes the sorting function and Φ^{-1} is a function of the inverse Φ. Our approach is summarized in Algorithm 4.

Algorithm 4 Enhanced DSIC (EDSIC)

1: Input: Received signals $\boldsymbol{y}$.
2: Initialization: Initialize the probabilities $\left\{\hat{\boldsymbol{p}}_k^{(0)}\right\}_{k=1}^{K}$.
3: **for** $q = 1, 2, \ldots, Q$ **do**
4: Sort $\left\{\hat{\boldsymbol{p}}_k^{(q-1)}\right\}_{k=1}^{K}$ in descending order of channel gains to $\left\{\hat{\boldsymbol{p}}_{r_k}^{(q-1)}\right\}_{k=1}^{K}$ and set $\left\{\hat{\boldsymbol{p}}_{r_k}^{(q)}\right\}_{k=1}^{K}$ to $\left\{\hat{\boldsymbol{p}}_{r_k}^{(q-1)}\right\}_{k=1}^{K}$.
5: **for** $k = 1, 2, \ldots, K$ **do**
6: Estimate $\hat{\boldsymbol{p}}_{r_k}^{(q)}$ from $\boldsymbol{y}$ and $\left\{\hat{\boldsymbol{p}}_l^{(q)}\right\}_{l\in\mathcal{K}\setminus r_k}$ using the (q,k)th bidirectional long short-term memory (biLSTM).
7: Update $\hat{\boldsymbol{p}}_{r_k}^{(q)}$ to $\left(\gamma\cdot\hat{\boldsymbol{p}}_{r_k}^{(q-1)} + (1-\gamma)\cdot\hat{\boldsymbol{p}}_{r_k}^{(q)}\right)$.
8: **end for**
9: Sort $\left\{\hat{\boldsymbol{p}}_{r_k}^{(q)}\right\}_{k=1}^{K}$ in the reverse order to $\left\{\hat{\boldsymbol{p}}_k^{(q)}\right\}_{k=1}^{K}$.
10: **end for**
11: Output: A set of hard-detected symbols $\hat{\boldsymbol{s}}$, obtained by (8).

4 | EXPERIMENTAL RESULTS

We evaluated the performances of the MAP, ISIC, DHIC, DSIC, and those of the proposed EDSIC detectors. We considered two types of linear channels with AWGN: a 4×4 channel (i.e., $K = 4$ users and $N = 4$ receiving antennas) and an 8×8 channel. For the transmitted symbol set, we used a constellation of binary phase-shift keying modulations, namely $\mathcal{S} = \{-1, 1\}$. To model the spatial exponential decay, we used a channel matrix $\boldsymbol{h}$ in which the elements are represented as follows,

$$h_{i,j} = e^{-|i-j|}, \quad i \in \{1, ..., N\}, \quad j \in \{1, ..., K\}. \tag{11}$$

We compared the symbol error rates (SERs) of the symbol detectors considered for both the perfect CSI and imperfect CSI cases. In the case of imperfect CSI, we assumed that the detectors can obtain an estimate of $\boldsymbol{h}$ with elements contaminated by independent and identically distributed additive Gaussian noise with variance σ_e^2. Herein, σ_e^2 refers to the variance in channel estimation errors that may occur in practical communication systems.

For both the DHIC and DSIC, we used a neural network structure consisting of three fully connected layers, as proposed in [25]. In this structure, each layer has $100 \cdot (N + (K-1) \cdot (M-1))$, 5000, and $50M$ nodes. where M is the number of possible labels, which is the size of the constellation set. For the proposed EDSIC, we used biLSTM for the first layer and two fully connected layers. Each layer yielded $(500 + 6 \cdot N \cdot K \cdot M)$, $100M$, and $10M$ nodes. Table 1 summarizes the comparison between the number of nodes for DSIC and EDSIC. As shown in Table 1, the proposed EDSIC has relatively few parameters, which enables low complexity, fast convergence during training, and model stability.

For DHIC, DSIC, and EDSIC, we used the Adam optimizer with a learning rate of 0.01 and performed 100 epochs of training. We set the number of training samples J to 5000. For DSIC and EDSIC, we set the number of iterations Q to 6. For the activation function of the EDSIC, sigmoidal and rectified linear units were used, as in [25]. In addition, we set the value of the relaxation factor γ to 0.1 for EDSIC.

Figure 7 shows the SER performances of the detectors over the 4×4 linear channel with AWGN in the perfect CSI case. In Figure 7, the performance of EDSIC approaches that of the ISIC, which is close to the optimal MAP performance. In particular, the optimal MAP detector achieved an SER of 10^{-3} at a signal-to-noise ratio (SNR) of 10.25 dB, whereas ISIC, DHIC, DSIC, and the proposed EDSIC required SNR values equal to 10.6 dB, 12.6 dB, 11.4 dB, and 11 dB, respectively. Figure 8 shows the detectors' SER performances over the 4×4 linear channel with AWGN for the imperfect CSI case. In this case, the proposed EDSIC demonstrates the best performance among all the detectors, including the MAP

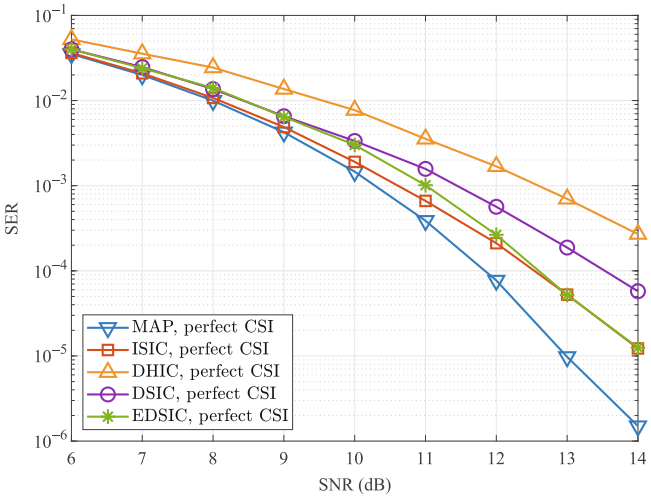

FIGURE 7 Performance comparison of maximum a posteriori (MAP), soft interference cancellation (ISIC), DHIC, DSIC, and that of the proposed EDSIC over the 4×4 linear channel with additive white Gaussian noise (AWGN) in the perfect channel-state information (CSI) case.

TABLE 1 Comparison of the number of nodes for deep hard interference cancellation (DHIC), deep soft interference cancellation (DSIC), and enhanced DSIC (EDSIC).

	$K = 4, N = 4,$ $M = 2$		$K = 8, N = 8,$ $M = 2$	
	DHIC/DSIC	EDSIC	DHIC/DSIC	EDSIC
First layer	700	692	2300	2228
Second layer	5000	200	5000	200
Third layer	100	20	100	20
Total	5800	912	7400	2448

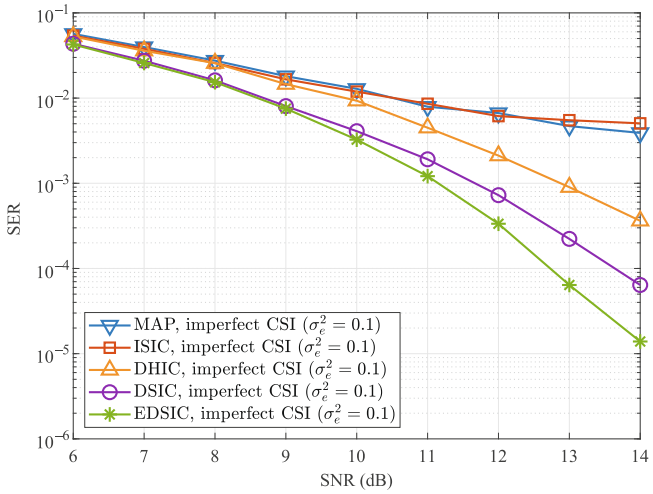

FIGURE 8 Performance comparison of MAP, ISIC, DHIC, DSIC, and that of the proposed EDSIC over the 4×4 linear channel with AWGN in the imperfect CSI case.

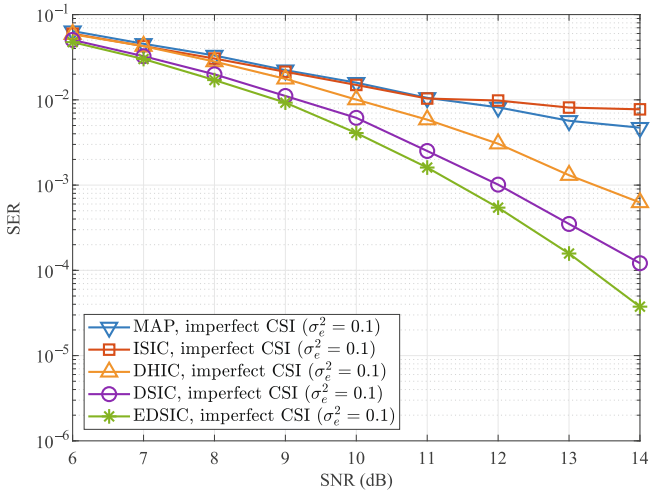

FIGURE 10 Performance comparison of MAP, ISIC, DHIC, DSIC, and that of the proposed EDSIC over the 8×8 linear channel with AWGN in the imperfect CSI case.

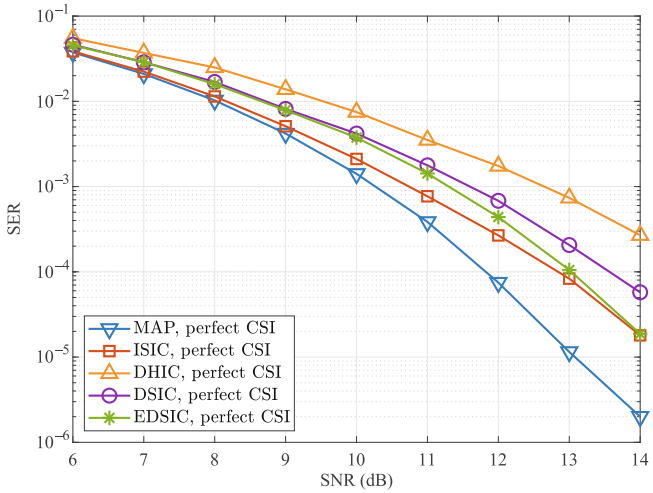

FIGURE 9 Performance comparison of MAP, ISIC, DHIC, DSIC, and that of the proposed EDSIC over the 8×8 linear channel with AWGN in the perfect CSI case.

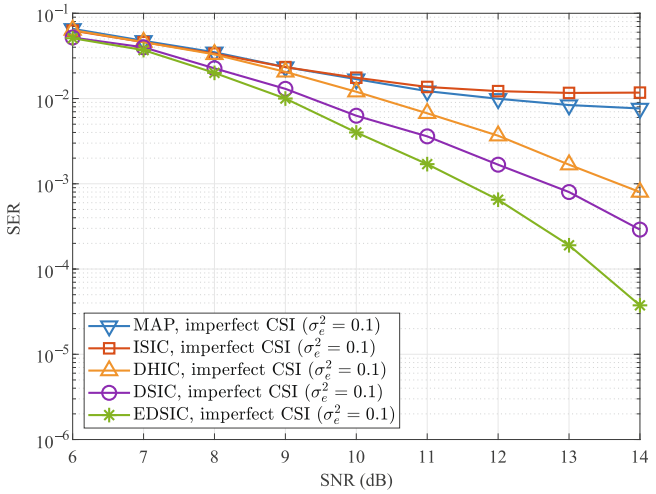

FIGURE 11 Performance comparisons of MAP, ISIC, DHIC, DSIC, and that of the proposed EDSIC over the 4×4 time-varying channel with AWGN in the imperfect CSI case.

detector, in the perfect CSI case. To achieve an SER of 3×10^{-4}, DHIC, DSIC, and the proposed EDSIC require SNR values, which are approximately equal to 14, 12.8, and 12 dB, respectively.

Based on a comparison of Figures 7 and 8, we can observe that model-based methods are designed subject to the assumption of perfect CSI, which leads to significant performance degradation in imperfect CSI environments. By contrast, data-driven methods can maintain their performances even when trained using incomplete data.

Figures 9 and 10 show the performance of the SER when the number of users and the number of receiving antennas are doubled as compared with the scenario in Figures 7 and 8. The results confirm that the performance deteriorates as the number of users increases, especially when channel information is insufficient. Even in this scenario, for the imperfect CSI case, the proposed EDSIC outperforms the DSIC and yields the best performance among all detectors, including the MAP detector, in the perfect CSI case. In particular, to achieve an SER of 6×10^{-4}, DHIC, DSIC, and the proposed EDSIC require SNR values approximately equal to 14, 12.4, and 12 dB, respectively.

The authors of [25] confirmed that the DSIC can effectively track time-varying channels using small training datasets. In this study, by using a BiLSTM structure, which is more suitable for time-series data than the DNN structure used by the DSIC, we found that the proposed

EDSIC further improved this advantage. To verify the performance, we used a set of coded signals generated by the low-density parity-check codes in [28]. The channel observed in the bth frame was modeled as a channel matrix $\boldsymbol{h}(b)$, and its elements are expressed as follows:

$$h(b)_{i,j} = e^{-|i-j|} \cdot |\cos(\Phi_i \cdot b \ (\mathrm{mod}\, 4))|, \tag{12}$$

where $i \in \{1,...,N\}$, $j \in \{1,...,K\}$, $\Phi = [63,66,69,72]$ for the case $N = 4$. It is defined in a form similar to that in [25]. As shown in Figure 11, the proposed EDSIC also yielded the best performance, even in time-varying channels where channel uncertainty exists. Compared with Figure 8, it can be observed that the performance gap with DSIC is larger than that of time-invariant channels. In particular, we confirmed that the SNR performance improvement of approximately 1.4 dB corresponded to a SER of 3×10^{-5}.

5 | CONCLUSIONS

In this study, we proposed an enhanced DSIC scheme (EDSIC) for multiuser symbol detection. The sequential structure and biLSTM neural networks, which are suitable for time-series data, accelerated receiver learning. In addition, to prevent the error propagation that may occur in sequential approaches, we used a relaxation update technique. The simulation results demonstrated that in the case of the time-varying channel, the proposed EDSIC outperformed the DSIC by approximately 1.4 dB. The MAP scheme exhibited the best performance in a perfect CSI environment. However, when the channel estimation was inaccurate, the MAP and ISIC schemes demonstrated a very large performance degradation; thus, the proposed scheme was considered to be the most useful in practical communication environments. In our future work, we will conduct a comprehensive analysis and experiments on the performance of the proposed EDSIC in various real-world wireless channel environments, including satellite-terrestrial or drone-to-drone channels.

ACKNOWLEDGMENTS

This study was supported by an Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korean government (MSIT) (No. 2021-0-00794, Development of 3D Spatial Mobile Communication Technology).

CONFLICT OF INTEREST STATEMENT

The authors declare that there are no conflicts of interest.

ORCID

Jihyung Kim https://orcid.org/0000-0002-3854-3145
Junghyun Kim https://orcid.org/0000-0003-0265-5169

REFERENCES

1. X. Chen, R. Jia, and D. W. K. Ng, *On the design of massive non-orthogonal multiple access with imperfect successive interference cancellation*, IEEE Trans. Commun. **67** (2019), no. 3, 2539–2551.
2. W.-J. Choi, K.-W. Cheong, and J. M. Cioffi, *Iterative soft interference cancellation for multiple antenna systems*, (Proceedings of the Wireless Communications and Networking Conference, Chicago IL, USA), 2000, pp. 304–309.
3. I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*, MIT Press, Cambridge, 2016.
4. D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. V. D. Driessche, T. Graepel, and D. Hassabis, *Mastering the game of Go without human knowledge*, Nature **550** (2017), no. 7676, 354–359.
5. O. Vinyals, I. Babuschkin, W. M. Czarnecki, M. Mathieu, A. Dudzik, J. Chung, D. H. Choi, R. Powell, T. Ewalds, P. Georgiev, J. Oh, D. Horgan, M. Kroiss, I. Danihelka, A. Huang, L. Sifre, T. Cai, J. P. Agapiou, M. Jaderberg, A. S. Vezhnevets, R. Leblond, T. Pohlen, V. Dalibard, D. Budden, Y. Sulsky, J. Molloy, T. L. Paine, C. Gulcehre, Z. Wang, T. Pfaff, Y. Wu, R. Ring, D. Yogatama, D. Wunsch, K. McKinney, O. Smith, T. Schaul, T. Lillicrap, K. Kavukcuoglu, D. Hassabis, C. Apps, and D. Silver, *Grandmaster level in Starcraft II using multi-agent reinforcement learning*, Nature **575** (2019), no. 7782, 350–354.
6. Y. LeCun, Y. Bengio, and G. Hinton, *Deep learning*, Nature **521** (2015), no. 7553, 436–444.
7. D. Gündüz, P. de Kerret, N. D. Sidiropoulos, D. Gesbert, C. R. Murthy, and M. van der Schaar, *Machine learning in the air*, IEEE J. Sel. Areas Commun. **37** (2019), no. 10, 2184–2199.
8. H. Lee, B. Lee, H. Yang, J. Kim, S. Kim, W. Shin, B. Shim, and H. V. Poor, *Towards 6G hyper-connectivity: vision, challenges, and key enabling technologies*, J. Commun. Netw., early access, (2023). DOI 10.23919/JCN.2023.000006.
9. Q. Mao, F. Hu, and Q. Hao, *Deep learning for intelligent wireless networks: a comprehensive survey*, IEEE Commun. Surv. Tutor. **20** (2018), no. 4, 2595–2621.
10. T. O'Shea and J. Hoydis, *An introduction to deep learning for the physical layer*, IEEE Trans. Cognit. Commun. Netw. **3** (2017), no. 4, 563–575.
11. M. Honkala, D. Korpi, and J. M. J. Huttunen, *DeepRx: fully convolutional deep learning receiver*, IEEE Trans. Wirel. Commun. **20** (2021), no. 6, 3925–3940.
12. S. Han, J. Kim, and H.-Y. Song, *A new design of channel denoiser using residual autoencoder*, IET Electron. Lett. **59** (2023), no. 2, 1–3.
13. E. Nachmani, E. Marciano, L. Lugosch, W. J. Gross, D. Burshtein, and Y. Be'ery, *Deep learning methods for improved decoding of linear codes*, IEEE J. Sel. Top. Signal Process. **12** (2018), no. 1, 119–131.
14. S. Zhang, S. Chen, and X. Yang, *Learn codes: inventing low-latency codes via recurrent neural networks*, IEEE J. Sel. Areas Inform. Theory **1** (2020), no. 1, 207–216.
15. S. Dörner, S. Cammerer, J. Hoydis, and S. Ten Brink, *Deep learning based communication over the air*, IEEE J. Sel. Top. Signal Process. **12** (2018), no. 1, 132–143.

16. H. Kim, S. Oh, and P. Viswanath, *Physical layer communication via deep learning*, IEEE J. Sel. Areas Inform. Theory **1** (2020), no. 1, 5–18.
17. J. Kim, B. Lee, H. Lee, Y. Kim, and J. Lee, *Deep learning-assisted multi-dimensional modulation and resource mapping for advanced OFDM systems*, (Proceedings of the IEEE GLobecom Workshops, Abu Dhabi, UAE), 2018, pp. 1–6.
18. H. Ye, L. Liang, G. Y. Li, and B.-H. Juang, *Deep learning-based end-to-end wireless communication systems with conditional gans as unknown channels*, IEEE Trans. Wirel. Commun. **19** (2020), no. 5, 3133–3143.
19. W. Lee, M. Kim, and D.-H. Cho, *Deep power control: transmit power control scheme based on convolutional neural network*, IEEE Commun. Lett. **22** (2018), no. 6, 1276–1279.
20. F. Liang, C. Shen, W. Yu, and F. Wu, *Towards optimal power control via ensembling deep neural networks*, IEEE Trans. Commun. **68** (2020), no. 3, 1760–1776.
21. Y. Shen, Y. Shi, J. Zhang, and K. B. Letaief, *Graph neural networks for scalable radio resource management: architecture design and theoretical analysis*, IEEE J. Sel. Areas Commun. **39** (2021), no. 1, 101–115.
22. H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu, and N. D. Sidiropoulos, *Learning to optimize: training deep neural networks for interference management*, IEEE Trans. Signal Process. **66** (2018), no. 20, 5438–5453.
23. C. Lin, Q. Cheng, and X. Li, *A deep learning approach for MIMO-NOMA downlink signal detection*, Sensors **19** (2021), no. 11, 1–22.
24. T. V. Luong, N. Shlezinger, C. Xu, T. M. Hoang, Y. C. Eldar, and L. Hanjo, *Deep learning based successive interference cancellation for the non-orthogonal downlink*, IEEE Trans. Vehic. Technol. **71** (2022), no. 11, 11876–11888.
25. N. Shlezinger, R. Fu, and Y. C. Eldar, *DeepSIC: deep soft interference cancellation for multiuser MIMO detection*, IEEE Trans. Wirel. Commun. **20** (2021), no. 2, 1349–1362.
26. W. Lee, K. Lee, and T. Q. S. Quek, *Deep-learning-assisted wireless-powered secure communications with imperfect channel state information*, IEEE Int. Things J. **9** (2022), no. 13, 11464–11476.
27. N. Shlezinger, R. Fu, and Y. C. Eldar, *Hybrid beamforming based on an unsupervised deep learning network for downlink channels with imperfect CSI*, IEEE Wirel. Commun. Lett. **11** (2022), no. 7, 1543–1547.
28. 3GPP, Technical Specification Group Radio Access Network; NR; multiplexing and channel coding; (Release 17), Technical Specification (TS) 38.212, 3rd Generation Partnership Project (3GPP), version 17.4.0, December 2022.

AUTHOR BIOGRAPHIES

Jihyung Kim received his PhD in Electrical and Electronic Engineering at the Yonsei University, Seoul, Republic of Korea, in 2007. Since 2007, he has been at the Electronics and Telecommunications Research Institute, Daejeon, Republic of Korea, where he is currently a Principal Researcher. His primary research interests include AI, RIS, and NTN for 5G/6G mobile communications.

Junghyun Kim received his BS, MS, and PhD degrees at the Department of Electrical and Electronic Engineering at the Yonsei University, Republic of Korea, in 2006, 2008, and 2017, respectively. From 2010 to 2013, he worked as an engineer at the Electronics and Telecommunications Research Institute in Daejeon, Republic of Korea. From 2017 to 2019, he worked as a Senior Engineer at Samsung Research, Seoul, Republic of Korea. He is currently an Assistant Professor at the Department of Artificial Intelligence, Sejong University, Seoul, Republic of Korea. Before joining Sejong University, he was a Faculty member at the Soonchunhyang University, Asan, Republic of Korea, from 2019 to 2022. His research interests include intelligent wireless communication systems, postquantum cryptography, drug discovery and development, and graph theory.

Moon-Sik Lee received his PhD degree in Electronic Engineering at the Gwangju Institute of Science and Technology, Gwangju, Republic of Korea, in 2005. From February 2008 to February 2009, he was a Postdoctoral Scholar at the Department of Electrical Engineering, Stanford University, Stanford, CA, USA. Since 2005, he has been a Principal Researcher at Electronics and Telecommunications Research Institute, Daejeon, Republic of Korea, where he is currently the Assistant Vice President of the Satellite Communication Research Division. His research interests include 5G/6G mobile communications, 3D mobile communications, NTN, open RAN, BF/MIMO, and array signal processing. He is a member of the KICS and O-RAN alliances.

How to cite this article: J. Kim, J. Kim, and M.-S. Lee, *Enhanced deep soft interference cancellation for multiuser symbol detection*, ETRI Journal **45** (2023), 929–938. DOI 10.4218/etrij.2022-0462