

تخمین و بازخورد مشترک کانال با استفاده از ترنسفورمرهای مبتنی بر توکن‌های ماسک‌شده در سیستم‌های مایمو (Massive MIMO) گسترده

Mingming Zhao*, Lin Liu*, Lifu Liu, Mengke Li†, Qi Tian

چکیده- تخمین اطلاعات وضعیت کانال (CSI) در لینک پایین‌دست و دریافت آن با سریار کم، از چالش‌های اصلی در سیستم‌های مایمو گسترده (Massive MIMO) در حالت اف‌دی‌دی (FDD) برای دستیابی به بهره‌ی بالای MIMO به شمار می‌رود. در سال‌های اخیر، پژوهش‌های متعددی برای بهره‌گیری از توان شبکه‌های عصبی عمیق جهت بهبود تخمین کانال و بازخورد آن انجام شده است. با این حال، روش‌های موجود هنوز نتوانسته‌اند از ویژگی‌های همبستگی درونی موجود در CSI به‌طور کامل بهره‌برداری کنند. در نتیجه، معمولاً برای انجام این دو وظیفه از ساختارهای شبکه‌ای متفاوتی به‌صورت جداگانه استفاده می‌شود. برای دستیابی به تخمین و بازخورد مشترک کانال، این مقاله یک شبکه‌ی مبتنی بر رمزگذار-رمزگشا (Encoder-Decoder) پیشنهاد می‌کند که همبستگی‌های ذاتی موجود در حوزه فرکانس ماتریس CSI را آشکار می‌سازد. کل شبکه رمزگذار-رمزگشا برای فشرده‌سازی کانال به‌کار گرفته شده است. برای استخراج و بازآرایی مؤثر ویژگی‌های همبسته، یک روش کدگذاری مبتنی بر Self-Mask Attention همراه با یک استراتژی ماسک‌گذاری فعال به‌منظور افزایش کارایی ارائه شده است. تخمین کانال از طریق بخش رمزگشا انجام می‌شود، جایی که یک مازول حذف نویز سبک مبتنی بر پرسپترون چندلایه (MLP) برای تخمین دقیق‌تر استفاده می‌گردد. نتایج آزمایش‌های گسترده نشان می‌دهد که روش پیشنهادی ما نه تنها در وظایف مشترک تخمین و بازخورد کانال عملکردی بهتر از روش‌های پیشرفته موجود دارد، بلکه در وظایف مستقل نیز نتایج قابل‌توجهی ارائه می‌دهد.

کلمات کلیدی- تخمین کانال، بازخورد کانال، روش ماسک‌توکن، ترنسفورمر

1. مقدمه

فناوری مایمو گسترده (Massive MIMO) یکی از مؤلفه‌های کلیدی شبکه‌های 5G-Advanced و 6G به شمار می‌رود. این فناوری با افزایش تعداد آنتن‌های فرستنده، کیفیت انتقال سلولی و ظرفیت سیستم را بهبود می‌بخشد [1] و [2]. استفاده از آرایه‌های بزرگ آنتن‌های آنالوگ در سیستم‌های MIMO باعث افزایش سریار ارسال و ایجاد تداخل نویزی می‌شود که این موضوع منجر به کاهش چشمگیر عملکرد، به‌ویژه در سیستم‌های FDD، می‌گردد. در سیستم‌های FDD هیچ نوع تقارنی (Reciprocity) وجود ندارد [3]؛ بنابراین ایستگاه پایه مجبور است برای ایجاد لینک ارتباطی مناسب و به‌دست آوردن بهره MIMO، اطلاعات وضعیت کانال پایین‌دست (DL-CSI) را از کاربر دریافت کند. از یک طرف، اندازه بزرگ ماتریس CSI در لینک پایین دست نیازمند سریار قابل‌توجهی است، در حالی که پهنای باند موجود

محدود است؛ از این رو، فشرده‌سازی و بازخورد کانال ضروری است. از طرف دیگر، برای دستیابی به بازخورد دقیق، نیاز به تخمین کانال مبتنی بر پایلوت وجود دارد. پژوهش‌های پیشین روش‌های مختلفی را برای بهینه‌سازی تخمین کانال و/یا بازخورد کانال با هدف بهبود دقت کانال پایین‌دست بررسی کرده‌اند. برای تخمین کانال، روش حداقل مربعات (LS) [4] و روش حداقل میانگین مربعات خطا (MMSE) [5] در روش کلاسیک و رایج در تخمین کانال هستند. با این حال، تخمین‌های LS به‌شدت تحت تأثیر نویز قرار می‌گیرند و به‌کارگیری روش MMSE در عمل با چالش‌هایی همراه است، زیرا پیش‌پیش دسترسی به اطلاعات دقیق کانال لازم است. در سال‌های اخیر، روش‌های مبتنی بر یادگیری ماشینی که شامل شبکه‌های مبتنی بر CNN [6]، [7] و شبکه‌های مبتنی بر Attention [8] هستند، توانسته‌اند دقت تخمین کانال را به‌طور قابل‌توجهی افزایش دهند. برای فشرده‌سازی کانال، استانداردهای 3GPP در نسخه 16 (Rel-16) استفاده از کدبوک‌های نوع I و نوع II بهبود یافته (eType II) را برای سیستم بازخورد CSI مشخص کرده‌اند [9]. ساختارهای مبتنی بر رمزگذار-رمزگشا [10]–[14] نیز می‌توانند عملکرد بازخورد CSI را بیش از پیش بهبود دهند. برای تخمین و بازخورد مشترک کانال، روش‌های مبتنی بر یادگیری عمیق [15]، [16] پیشرفت قابل‌توجهی ارائه کرده‌اند. روش‌های مبتنی بر Attention مشهورترین دسته روش‌های یادگیری عمیق هستند که با تقویت روابط میان زیرکانال‌ها، عملکرد تخمین کانال و بازخورد را بهبود می‌بخشند [17]–[20] با این وجود، روش‌های پیشین معمولاً به انجام تنها یک وظیفه (تخمین یا بازخورد) محدود بوده‌اند یا برای انجام دو وظیفه مشترک، از دو شبکه یا دو روش مستقل استفاده کرده‌اند.

برای آن که تخمین کانال و بازخورد کانال را به‌صورت صریح در یک ساختار یکپارچه قرار دهیم، روشی مبتنی بر Attention به نام FlowMat پیشنهاد می‌کنیم. این روش بر پایه معماری رمزگذار-رمزگشا (Encoder-Decoder) طراحی شده و از ویژگی‌های ماتریس CSI و اجزای مختلف شبکه به‌طور کامل بهره می‌برد. بر اساس این مشاهده که میان ماتریس‌های کانال، همبستگی‌های ذاتی در حوزه فرکانس وجود دارد، امکان جایگزینی متقابل داده‌های حوزه فرکانس فراهم می‌شود. ما از این ویژگی ماتریس کانال استفاده کرده و یک روش کدگذاری مبتنی بر Self-Attention ارائه می‌دهیم تا استخراج ویژگی‌های مرتبط آسان‌تر گردد. سپس، فشرده‌سازی کانال از طریق استخراج ویژگی‌های به‌شدت همبسته محقق می‌شود. ما یک تکنیک ماسک‌گذاری فعال (Active Masking) نیز برای استخراج این ویژگی‌های بسیار همبسته پیشنهاد می‌کنیم. توکن‌های ماسک قابل یادگیری (Learnable Mask Tokens) به رمزگذار کمک می‌کنند تا نداشت بهتری میان ویژگی‌های عمیق و اطلاعات کانال بیاموزد. سپس، تکمیل و بازسازی کانال از طریق رمزگشایی این توکن‌های ماسک انجام

* Mingming Zhao و Lin Liu به‌صورت برابر در این مقاله مشارکت داشته‌اند و باید به‌عنوان هم‌نویسندگان اول در نظر گرفته شوند.
† Mengke Li نویسنده مسئول (Corresponding Author) است.

زیرشبکه حذف نویز غیرکور Non-blind Denoising He و همکاران [23] از یک شبکه پیام‌رسانی تقریبی مبتنی بر حذف نویز یادگرفته‌شده برای سیستم‌های مایمو موج‌میلیمتری (mmWave) مبتنی بر Beamspace بهره‌بردند. روش‌های مبتنی بر Attention در میان روش‌های مبتنی بر Attention، شبکه AttentionNet [8] از بلوک‌های ترنسفورمر برای توجه کانال استفاده می‌کند Gao. و همکاران [20] و همچنین Mashhadi و همکاران [24] به ترتیب ماژول توجه کانال CNN و بلوک‌های غیرمحل CNN را برای تخمین کانال پیشنهاد کرده‌اند.

B. بازخورد کانال

علاوه بر تخمین کانال، بازخورد کانال نیز از فناوری‌های بسیار مهم در لایه فیزیکی سیستم‌های مایمو است Wen. و همکاران [10] روش CsiNet را پیشنهاد کردند که بر پایه ساختار خودرمزگذار (Autoencoder) مبتنی بر CNN بنا شده است. در CsiNet، رمزگذار (Encoder) برای فشرده‌سازی CSI طراحی شده و رمزگشا (Decoder) مسئول بازیابی CSI در سمت تجهیزات کاربر (UE) و ایستگاه پایه (BS) است. سپس، با الهام از CsiNet، مجموعه‌ای از پژوهش‌ها [11]–[14] انواع مختلفی از شبکه‌های CNN را طراحی کردند تا عملکرد بازخورد CSI را بیشتر بهبود دهند Mashhadi. و همکاران [25] یک روش فشرده‌سازی تصویر ارائه دادند که در آن هزینه نخ–عوجاج و کدگذاری آن‌تروپی حسابی به کار گرفته می‌شود تا کمترین سربار بیتی ممکن حاصل شود. شبکه‌های LSTM [26]، [27] نیز در رمزگذار و رمزگشا به کار گرفته شده‌اند تا از همبستگی استخراج‌شده میان زیرحامل‌ها بیشترین استفاده صورت گیرد. بیشتر روش‌های مبتنی بر CNN و LSTM بر بازخورد کامل اطلاعات کانال (F-CSI) تمرکز کرده‌اند، بدون اینکه ویژگی‌های ذاتی کانال—مانند همبستگی—را در نظر بگیرند؛ و این موضوع می‌تواند موجب ویژگی‌های فشرده‌سازی غیرنمایند و دقت پایین در بازسازی شود. در سیستم فعلی G5، بازخورد CSI عمدتاً بر اساس فشرده‌سازی و ارسال بردارهای ویژه (Eigenvectors) ماتریس کانال انجام می‌شود، همان‌طور که در استاندارد 3GPP ذکر شده است Liu. و همکاران [28] روش EVCSiNet را پیشنهاد کردند؛ یک ساختار مبتنی بر CNN که از ویژگی‌های بردار ویژه بهره می‌برد. در سال‌های اخیر، روش‌های مبتنی بر Attention محبوبیت بالایی در حوزه‌های بینایی کامپیوتری (CV) و پردازش زبان طبیعی (NLP) پیدا کرده‌اند Xiao. و همکاران [18] روش EVCSiNet-T را برای بازخورد کانال ارائه کردند که در آن بردار ویژه کانال فشرده شده و مکانیسم Attention برای بازسازی کانال به کار گرفته می‌شود.

C. تخمین و بازخورد کانال نقطه‌ای

تمام روش‌های مبتنی بر یادگیری عمیق برای بازخورد CSI که پیش‌تر ذکر شدند، این فرض را در نظر می‌گیرند که کانال‌های لینک پایین‌دست

می‌شود. برای تخمین کانال، از بخش رمزگشا استفاده می‌کنیم. برای کاهش نویز محیطی و تکمیل ماتریس کانال H، یک ماژول حذف نویز مبتنی بر پرسپترون چندلایه سبک (Lightweight MLP Denoiser) معرفی می‌کنیم. چارچوب پیشنهادی FlowMat امکان تخمین و بازخورد کانال را با عملکرد بالا و سربار محاسباتی کم فراهم می‌کند و از بهترین روش‌های موجود نیز بهتر عمل می‌کند. مهم‌ترین دستاوردهای ما به صورت زیر خلاصه می‌شوند:

- ما همبستگی ذاتی موجود در حوزه فرکانس میان ماتریس‌های کانال را آشکار می‌کنیم. برای استفاده مؤثر از این همبستگی، بردار حوزه فرکانس را به عنوان واحد پایه ویژگی کانال طراحی کرده‌ایم. این رویکرد شباهتی به استفاده از پچ‌های بدون هم‌پوشانی در بینایی کامپیوتری (CV) و توکن‌های کلمات در پردازش زبان طبیعی (NLP) دارد.
- ما یک ترنسفورمر مبتکرانه مبتنی بر توکن به نام FlowMat پیشنهاد می‌کنیم که قادر است تخمین کانال و بازخورد آن را به صورت هم‌زمان در سناریوی لینک پایین‌دست انجام دهد. FlowMat به‌طور منحصربه‌فرد قادر است از همبستگی ذاتی موجود در حوزه فرکانس بهره‌برداری کرده و هر دو وظیفه تکمیل کانال و فشرده‌سازی کانال را با کارایی بالا انجام دهد.
- آزمایش‌های گسترده در وظایف مشترک تخمین و بازخورد کانال نشان می‌دهند که روش پیشنهادی ما از جدیدترین روش‌های موجود عملکرد بهتری دارد. علاوه بر این، به کارگیری روش پیشنهادی در وظایف مستقل تخمین کانال یا بازخورد کانال نیز منجر به بهبود عملکرد می‌شود.

2. کارهای مرتبط

A. تخمین کانال

به تازگی یادگیری عمیق به ابزار اصلی بسیاری از پژوهش‌ها تبدیل شده و به عنوان رویکردی امیدوارکننده برای تخمین کانال مورد استفاده قرار گرفته است. این پژوهش‌ها را می‌توان به دو دسته تقسیم کرد: روش‌های بدون Attention و روش‌های مبتنی بر Attention. روش‌های بدون Attention در میان روش‌های غیرمبتنی بر Attention، ChannelNet [6] نخستین کار مبتنی بر یادگیری عمیق است که برای حل مسئله تخمین کانال ارائه شد. این روش از شبکه SRCNN [21] برای سوپرزولوشن و از شبکه DnCNN برای تکمیل کانال و حذف نویز به ترتیب استفاده می‌کند. Chun و همکاران [22] یک مدل دو فاز ارائه کردند که کانال را در حوزه زمان تخمین می‌زند. فاز اول از یک مدل مبتنی بر پایلوت تشکیل شده که شامل یک MLP دولایه و یک CNN است و فاز دوم از یک مدل مبتنی بر داده بهره می‌برد Jin. و همکاران [7] از یک شبکه حذف نویز تصویر به نام CBDNet برای تخمین کانال استفاده کردند. این شبکه شامل دو زیرشبکه اصلی است: زیرشبکه تخمین سطح نویز-

گیرد. پس از تخمین کانال در سمت کاربر (UE)، ماتریس بردار ویژه متناظر $w \in \mathbb{C}^{N_t \times N_c}$ تشکیل می‌شود که شامل بردارهای $w_k \in \mathbb{C}^{N_t \times 1}$ با شرط نرمال‌سازی $\|w_k\|^2 = 1$ است. در این جا بردار ویژه متناظر با k امین زیرحامل (Subcarrier) در CSI است. ماتریس W می‌تواند به عنوان بردار پیش‌گذرانی لینک پایین‌دست مورد استفاده قرار گیرد و از طریق تجزیه بردارهای ویژه (Eigenvalue Decomposition) محاسبه شود. مدل سیستم مبتنی بر تجزیه بردارهای ویژه و پیش‌گذرانی به صورت زیر محاسبه می‌شود:

$$H_k^H H_k w_k = \lambda_k w_k, \quad (2)$$

$$y_i = w P_i s_i + n_i, \quad (3)$$

در این جا λ_i بیشترین مقدار ویژه ماتریس $H_k^H H_k$ ، $H_k \in \mathbb{C}^{N_r \times N_t}$ را نشان می‌دهد، که در آن H است. بنابراین، برای ایجاد یک لینک ارتباطی دقیق و دستیابی به بهره MIMO، تخمین کانال مبتنی بر پیلوت و بازخورد کانال دو چالش کلیدی در سیستم‌های FDD هستند.

2. تخمین کانال: (Channel Estimation)

پیش از ساخت لینک ارتباطی کامل، ایستگاه پایه (BS) باید پیلوت‌ها را به گیرنده ارسال کند، زیرا کانال‌های لینک پایین‌دست برای فرستنده ناشناخته‌اند. تخمین کانال مبتنی بر سنبل پیلوت معمولاً عملکرد خوبی برای ردیابی تغییرات ناگهانی کانال دارد؛ به ویژه در کانال‌های محوشونده (Fading Channels) در تعداد $(1 \leq N_p \leq N_c)$ از زیرحامل‌های پیلوت با فواصل یکسان. با این حال، چگالی زیرحامل‌های پیلوت و نویز موجود در لینک فضایی مستقیماً بر دقت بازیابی کانال تأثیر می‌گذارد. روش‌های کلاسیک تخمین کانال، مانند روش حداقل مربعات (LS)، به طور گسترده استفاده می‌شوند. در روش LS، با کمینه‌سازی خطای میانگین مربعات (MSE) بین سیگنال دریافتی پیلوت $(y_{i,p}, y_i, p \in \mathbb{C}^{N_p \times N_t})$ یک تخمین به دست می‌آید. تخمین LS در حوزه فرکانس به صورت زیر ارائه می‌شود:

$$\hat{H}_{i,p} = \underset{\hat{H}_{i,p}}{\operatorname{argmin}} \|y_{i,p} - \hat{H}_{i,p} \odot s_{i,p}\|, \quad (4)$$

که در آن $H^H_{i,p}$ ، $S_{i,p} \in \mathbb{C}^{N_p \times N_t}$ به ترتیب ماتریس تخمین زده شده کانال و سیگنال پیلوت ارسال شده را نشان می‌دهند و \odot نیز عملگر ضرب هادامارد (Hadamard Product) را نمایش می‌دهد. پس از آن، عملیات درون‌یابی (Interpolation) انجام می‌شود تا پاسخ‌های کانال در سایر زیرحامل‌ها به دست آمده و کل H_i بازسازی شود. روش LS به دلیل پیاده‌سازی بسیار ساده و پیچیدگی فوق‌العاده کم شناخته شده است.

به طور دقیق در سمت کاربر شناخته شده‌اند؛ در حالی که این موضوع در سیستم‌های مخابراتی واقعی امکان‌پذیر نیست. به دست آوردن اطلاعات دقیق کانال برای ایستگاه پایه اهمیت حیاتی دارد. از این رو، تحقیقات اخیر به بررسی استفاده از شبکه‌های عصبی برای بهینه‌سازی مشترک تخمین کانال و بازخورد کانال پرداخته‌اند [15]. و همکاران [15] یک معماری شبکه عصبی عمیق (DNN) ارائه کرده‌اند که سیگنال‌های پیلوت و مازول تخمین کانال را به صورت انتها به انتها (End-to-End) و به صورت مشترک طراحی می‌کند تا از افت عملکرد ناشی از طراحی جداگانه این بخش‌ها جلوگیری شود [Chen Tong و همکاران] [16]. JCEF را پیشنهاد کرده‌اند، جایی که دو شبکه مجزا ساخته شده است تا به ترتیب تخمین صریح (Explicit) و تخمین ضمنی (Implicit) کانال و همچنین بازخورد آن را انجام دهند.

3. مقدمات

A. مدل سیستم 1: سیستم مایمو گسترده: Massive MIMO برای یک سیستم مایمو گسترده معمول در حالت FDD، فرض می‌کنیم که سیستم دارای $N_t \geq 1$ و $N_r \geq 1$ آنتن فرستنده در ایستگاه پایه (BS) و $N_r \geq 1$ و $N_t \geq 1$ آنتن گیرنده در سمت کاربر (UE) است.

از OFDM با $N_c \geq 1$ و $N_c \geq 1$ زیرفرکانس (Subcarrier) استفاده می‌شود که شامل 4 بلوک منبع (Resource Blocks – RBs) است.

در فاز لینک پایین‌دست (Downlink)، مؤلفه سیگنال دریافتی مربوط به زیرفضای i ام به صورت زیر نشان داده می‌شود:

$$y_i \in \mathbb{C}^{N_c \times 1} \text{ که بیانگر بردار سیگنال دریافتی در زیرباند } i \text{ ام است.}$$

$$y_i = H_i P_i s_i + n_i, \quad (1)$$

که در آن $s_i \in \mathbb{C}^{N_c \times 1}$ ماتریس سیگنال ارسالی به آنتن گیرنده i ام است. $H_i \in \mathbb{C}^{N_r \times N_t}$ بردار پاسخ فرکانسی کانال را نشان می‌دهد که از طریق تخمین کانال مبتنی بر پیلوت به دست می‌آید. همچنین $P_i \in \mathbb{C}^{N_t \times N_c}$ ماتریس پیش‌گذرانی (Precoding) متناظر بوده $n_i \in \mathbb{C}^{N_c \times 1}$ نمایانگر نویز افزوده شده (Additive Noise) است. در ادامه، ماتریس کانال H_i در حوزه فرکانس-فضا از طریق کنار هم قرار دادن (Stacking) عناصر در حوزه فرکانس به صورت زیر به دست می‌آید:

$$H_i = [h_{i,1}, h_{i,2}, \dots, h_{i,N_t}]$$

که در آن هر پاسخ کانال مربوط به آنتن i ام در زیرحامل‌های فرکانسی است. با این حال، اندازه این ماتریس برای بازخورد مستقیم در یک سیستم مایمو گسترده بسیار بزرگ و غیرقابل قبول است. پتانسیل تجزیه بردارهای ویژه مانند روش EZF [29] می‌تواند برای دستیابی به پیش‌گذرانی سبک‌وزن و ایجاد لینک‌های ارتباطی کارآمد مورد استفاده قرار

برای کاهش بیشتر سربار بازخورد CSI لینک پایین دست و همچنین امکان بازسازی دقیق CSI در ایستگاه پایه، ما از روش متداول مبتنی بر شبکه‌های عصبی عمیق (DNN) برای فشرده‌سازی و بازسازی CSI استفاده می‌کنیم.

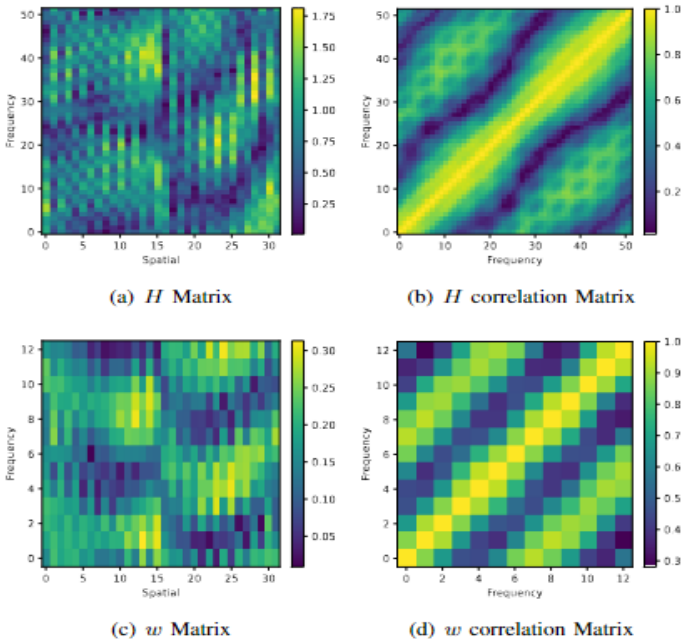
B. مشاهده کلیدی درباره کانال بی‌سیم: ما یک تحلیل جامع از ویژگی‌های کانال بی‌سیم و بردار ویژه انجام دادیم، که ما را به طراحی شبکه‌هایی با دقت و ظرافت بیشتر ترغیب می‌کند. ایستگاه پایه لینک‌های ارتباطی را با بازیابی ماتریس CSI لینک پایین دست (DL CSI) ایجاد می‌کند. ما عمدتاً ماتریس CSI اندازه‌گیری شده DL را $HECNC \times Nt \times Nr$ در حالتی بدون نویز و همچنین بردار ویژه w را در حوزه فضا-فرکانس تحلیل می‌کنیم. ما محاسبات همبستگی حوزه فرکانس را برای H_i , $i=1$ و همچنین برای w به صورت جداگانه انجام دادیم. نتایج در شکل ۱ نشان داده شده است. شکل ۱ (a) و شکل ۱ (b) شباهت قابل توجهی در حوزه فرکانس دارند. در شکل ۱ (c) و شکل ۱ (d)، شباهت در نزدیکی قطر ماتریس به طور قابل توجهی بیشتر است که نشان‌دهنده وجود مقدار زیادی اطلاعات محلی است. همبستگی‌ها حتی در فاصله‌های زیاد نیز بالا باقی می‌مانند. این نشان می‌دهد که CSI لینک پایین دست دارای مقدار مشخصی اطلاعات سراسری (Global Information) است. از این نتایج درمی‌یابیم که: حوزه فرکانس می‌تواند واحد پایه ویژگی‌ها باشد. حوزه فضا به عنوان داده برداری در نظر گرفته می‌شود. بردار ویژگی (Feature Token Vector) مشابه موارد زیر است Patch: های بدون هم‌پوشانی در بینایی کامپیوتری Word Token (CV) در پردازش زبان طبیعی (NLP) بنابراین، ما از واحد توکن فرکانسی (Frequency Token Unit) و ویژگی همبستگی درونی (Self-Correlation) برای ساخت مدل مشترک فشرده‌سازی کانال و تخمین کانال استفاده می‌کنیم.

4. طراحی سیستم

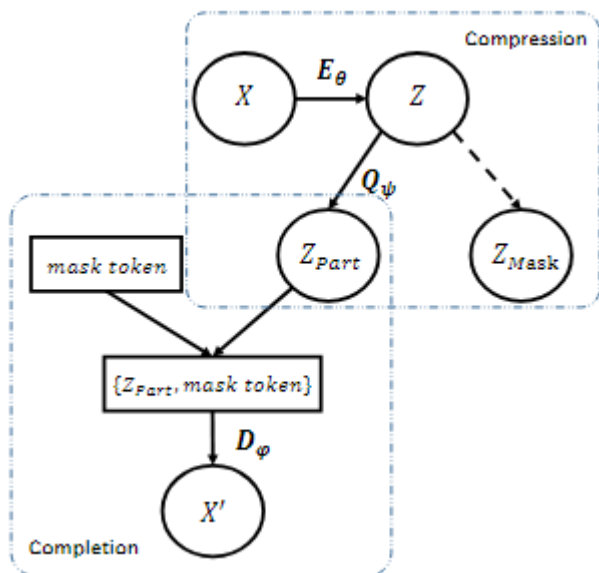
A. روش پیشنهادی FlowMat:

ایده ما درباره توکن ماسک شده از کار Masked AutoEncoders (MAE) متعلق به He و همکاران [30] الهام گرفته شده است؛ در آن کار، یک روش یادگیری خودنظارتی (Self-Supervised Learning) در حوزه بینایی کامپیوتری (CV) پیشنهاد شده است.

3. بازخورد کانال: (Channel Feedback) در سیستم‌های FDD، تجهیزات کاربر (UE) پس از دریافت سمبل‌های پایلوت و تخمین ماتریس کانال لینک پایین دست، CSI تخمینی را به ایستگاه پایه (BS) بازی‌گرداند. سپس BS می‌تواند بردارهای پیش‌گذاری (Precoding Vectors) مناسب را تولید کند تا تداخل بین کاربران کاهش یافته و کیفیت ارتباط افزایش یابد. در سیستم‌های مایمو گسترده، ماتریس CSI به صورت $HEC'Nc \times Nt \times Nr$ در نظر گرفته می‌شود که شامل $2 \times Nc \times Nt \times Nr$ مولفه عددی حقیقی است. این مقدار بزرگ، منجر به سربار بازخورد بسیار زیاد می‌شود. فشرده‌پذیری ماتریس CSI به طور گسترده در مقالات بررسی شده است، زیرا در سیستم‌های واقعی، حداقل سازی پارامترهای بازخورد اهمیت زیادی دارد. با استفاده از ارسال چندجریانی (Multi-stream) لینک پایین دست در سمت UE، بردار ویژه متناظر با زیرحامل‌ها می‌تواند به طور مستقیم به عنوان بردار پیش‌گذاری در BS استفاده شود. به کارگیری بردار ویژه به جای ماتریس کانال موجب می‌شود تا بعد Nr مربوط به آنتن‌های گیرنده کاهش یابد؛ بنابراین اندازه ماتریس کانال به شکل زیر کاهش پیدا می‌کند $2 \times Nc \times Nt$ که به طور قابل توجهی کوچک‌تر از حالت اولیه است و سربار بازخورد را کاهش می‌دهد.



شکل ۱. نمایش ویژگی‌های همبستگی در حوزه فضایی-فرکانسی.



شکل ۲. فرایند فشرده‌سازی و تکمیل در FlowMat. در یک فرایند رمزگذاری سلسله‌مراتبی داده، متغیرهای نهفته مانند Z نشان‌دهنده اطلاعات سطح بالا هستند. با استفاده از ماسک‌گذاری فعال (Active Masking)، بخش ZPZ_PZP یک نمایش فشرده و جمع‌وجور از داده است که می‌تواند داده اصلی را با قرار دادن توکن ماسک و رمزگشایی دوباره بازیابی کند.

آن‌ها یک تصویر را به پچ‌های بدون هم‌پوشانی تقسیم می‌کنند و سپس به صورت تصادفی برخی از این پچ‌ها را ماسک می‌کنند. رمزگذار (Encoder) فقط پچ‌های قابل مشاهده و بدون ماسک را پردازش می‌کند. رمزگشا (Decoder) تصویر اصلی را از نمایش‌های عمیق تولیدشده توسط رمزگذار و توکن‌های ماسک بازسازی می‌کند. هر توکن ماسک یک بردار قابل یادگیری مشترک است که نشان‌دهنده وجود پچ‌های حذف‌شده برای بازسازی است.

تابع هزینه (Loss Function) نیز خطای میانگین مربعات (MSE) بین تصویر بازسازی‌شده و تصویر اصلی است. در مرحله پیش‌آموزش (Pre-training)، بخش بزرگی از تصویر ورودی (حدود ۷۵٪ ماسک) می‌شود و شبکه یاد می‌گیرد که تصویر کامل را بازسازی کند. این روش به شبکه اجازه می‌دهد مدل‌هایی با ظرفیت بالا و تعمیم‌پذیری قوی یاد بگیرد. رمزگذار پیش‌آموزش‌دیده بعدها می‌تواند در وظایف پایین‌دستی مانند طبقه‌بندی و قطعه‌بندی به کار گرفته شود و عملکرد بهتری نسبت به بسیاری از روش‌های پیش‌آموزش دیگر (مانند MoCo v3 [31] و BEiT [32]) ارائه دهد.

1. تعریف مسئله: (Problem Definition)

ما پیشنهاد می‌کنیم که روش ماسک‌گذاری فعال (Active Masking) و کدگذاری Self-Mask-Attention در FlowMat می‌تواند برای

فشرده‌سازی (Compression) و تکمیل (Completion) کانال به کار گرفته شود؛ روشی که کارایی آن در تخمین کانال و بازسازی بازخورد کانال به طور کامل تأیید شده است. ما یک دنباله به صورت $X = \{x_1, \dots, x_N\}, x_n \in \mathbb{C}^{1 \times N_t}$ می‌کنیم تا ماتریس کانال H و ماتریس بردار ویژه V را به طور یکنواخت بیان کند. ما داده‌ها را در دو حوزه فرکانس و فضا به اطلاعات دوبعدی تقسیم می‌کنیم. بردار حوزه فرکانس x_n به عنوان واحد پایه کانال در نظر گرفته می‌شود که مشابه یک توکن (Token) در حوزه‌های CV و NLP عمل می‌کند. در روش‌های سنتی فشرده‌سازی مبتنی بر DNN، معمولاً تمام داده‌های رمزگذار صاف‌سازی (Flatten) شده، سپس از طریق شبکه Downsample و در نهایت با روش‌های Upsampling بازسازی می‌شود. اما این عملیات به وضوح از ویژگی‌های خاص کانال بی‌سیم استفاده کامل نمی‌کند. برای بهره‌برداری کامل از اطلاعات همبستگی میان حوزه‌های فرکانس، ما بخشی از توکن‌ها را به طور فعال ماسک می‌کنیم تا فشرده‌سازی انجام شود. سپس توکن ماسک‌شده را در مجموعه توکن‌های فشرده درج کرده و آن را به عنوان یک ماتریس کامل وارد رمزگشا می‌کنیم تا کل ماتریس کانال اصلی بازیابی شود.

2. طراحی مدل: (Model Design)

چارچوب رمزگذار-رمزگشا (Encoder-Decoder) برای فشرده‌سازی و تکمیل (Compression & Completion) به صورت زیر تعریف می‌شود:

$$Z = E_{\theta}(X), \quad (5)$$

$$Z_{Part}, Z_{Mask} = Q_{\psi}(Z), \quad (6)$$

که در آن X، Z، Z_Part و Z_Mask به ترتیب بیانگر داده ورودی، داده فشرده، داده مفید و قابل یادگیری، و داده ماسک‌شده هستند. همچنین E_{θ} و Q_{ψ} به ترتیب نشان‌دهنده رمزگذار (Encoder) و Query هستند. برای بازسازی و بازیابی متغیر نهفته Z_Part، ابتدا به جای استفاده از تکنیک Upsampling، از توکن ماسک استفاده می‌کنیم. سپس این متغیر نهفته اصلاح‌شده به رمزگشا (Decoder) ارسال می‌شود. این فرایند به صورت ریاضی به شکل زیر بیان می‌شود:

$$X' = D_{\phi}(Z_{Part}, masktoken). \quad (7)$$

کل فرایند فشرده‌سازی (Compressing) و تکمیل (Completing) در شکل ۲ نشان داده شده است. هدف ما در این مطالعه، شناسایی مجموعه‌ای از توابع رمزگذاری (Encoding)، Query و رمزگشایی (Decoding) است که پارامترهای آن‌ها به ترتیب با θ ، ψ و ϕ نمایش داده می‌شوند. این مجموعه پارامترها از طریق کمینه‌سازی اختلاف بین ماتریس اصلی X و ماتریس بازسازی‌شده X' به دست می‌آید:

$$(\theta, \phi, \psi) = \underset{\theta, \phi, \psi}{\operatorname{argmin}} \|X - X'\|. \quad (8)$$

3. Query، رمزگذار (Encoder) و رمزگشا (Decoder) در FlowMat برای دستیابی مؤثر به فشرده‌سازی و تکمیل از طریق استخراج زیرمجموعه داده‌ها و درج توکن ماسک، طراحی دقیق و حساب شده استراتژی‌های Query، رمزگذار و رمزگشا ضروری است. در این بخش، این سه مؤلفه به‌طور کامل و جزئی معرفی و توضیح داده می‌شوند.

Query کوئری: با توجه به تعداد واحدهای هدف فشرده‌سازی شده در حوزه فرکانس، ما بردارهای حقیقی تصادفی به نام $\text{query} \in \text{ERN}$ تولید می‌کنیم. کوئری یک بردار قابل یادگیری است که در طی آموزش به‌روزرسانی می‌شود. هر بار که زیرمجموعه هدف به‌عنوان ZPart انتخاب می‌شود، مجموعه موقعیت‌های Top-K این کوئری بر اساس نرخ فشرده‌سازی انتخاب شده تا عملیات فشرده‌سازی انجام گیرد.

Encoder رمزگذار: ما عملیات فشرده‌سازی را با انتخاب یک زیرمجموعه از X انجام می‌دهیم. برای دستیابی به بازسازی (Reconstruction) باکیفیت، لازم است که Z فشرده (Compact) و نماینده مناسب (Representative) باشد. ما از مکانیسم Self-Attention در ساختار ترنسفورمر استفاده می‌کنیم تا متغیرهای نهفته کاملاً همبسته به دست آوریم. علاوه بر این، یک مسیر انتقال ویژه با ماسک M' در آخرین بلوک معرفی می‌کنیم تا انتقال اطلاعات سراسری (Global Information) به مجموعه هدف تسهیل گردد. محاسبه رمزگذار (Encoder) به صورت زیر است:

$$\begin{aligned} Y_1 &= XW_1 + \text{Pos}, \\ [Q, K, V] &= [Y_1W_Q, Y_1W_K, Y_1W_V], \\ Y_2 &= \text{Softmax}\left(\frac{QK^T + M'}{\sqrt{d}}\right)V, \\ Z &= W_2Y_2, \end{aligned} \quad (9)$$

که در آن Pos نشان‌دهنده بردار جایگذاری مکانی (Position Embedding) است. پارامترهای W به صورت

(Learnable) W^* ، $(*=1,2,Q,K,V)$ همگی قابل یادگیری (Parameters) هستند. در معادله ۹ (Eq. 9)، ماسک M' به صورت زیر تعریف می‌شود:

$$M'_{r,h} = \begin{cases} 0, & \text{if token } h \text{ is masking} \\ 1, & \text{if token } h \text{ is non-masking} \end{cases} \quad (10)$$

که در آن $r, h \leq N_c$ نشان‌دهنده تعداد زیرحامل‌ها (Subcarriers) هستند. شکل ۴ این مرحله تک‌مرحله‌ای ماسک-آتِنشن (Single Mask-Attention Stage) را نشان می‌دهد.

تصویر بدلیل سائز بالا در صفحه بعد قابل نمایش میباشد

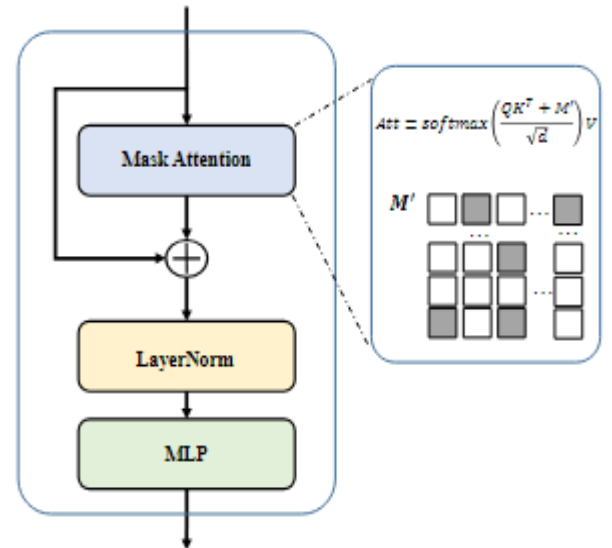
شکل ۳. ساختار روش پیشنهادی ما. در رویکرد ما از چارچوب معمول رمزگذار-رمزگشا (Encoder-Decoder) مبتنی بر DNN استفاده شده و بردار کوئری قابل یادگیری ($Q\phi$) برای انجام ماسک‌گذاری فعال (Active Masking) در آن تعبیه شده است. ابتدا با استفاده از بلوک‌های Attention مشترک چندگانه (ATT) و یک بلوک ماسک آتِنشن (MAT)، متغیرهای نهفته (Latent Variables) استخراج می‌شوند. سپس برای این متغیرهای نهفته، ماسک‌گذاری فعال اعمال شده و یک نمایش فشرده (Compact Representation) به دست می‌آید. در نهایت، رمزگشا (Decoder) ماتریس کامل شده را با جاسازی توکن ماسک و الحاق نمایش فشرده بازیابی می‌کند.

$$\begin{aligned}
 Y_3 &= [Z_{Part}, Mask_Token], \\
 Y_4 &= Y_3 W_3 + Pos, \\
 [Q, K, V] &= [Y_4 W_Q, Y_4 W_K, Y_4 W_V], \\
 Y_5 &= softmax(\frac{QK^T - M'}{\sqrt{d}})V, \\
 X' &= W_1 Y_5.
 \end{aligned} \tag{11}$$

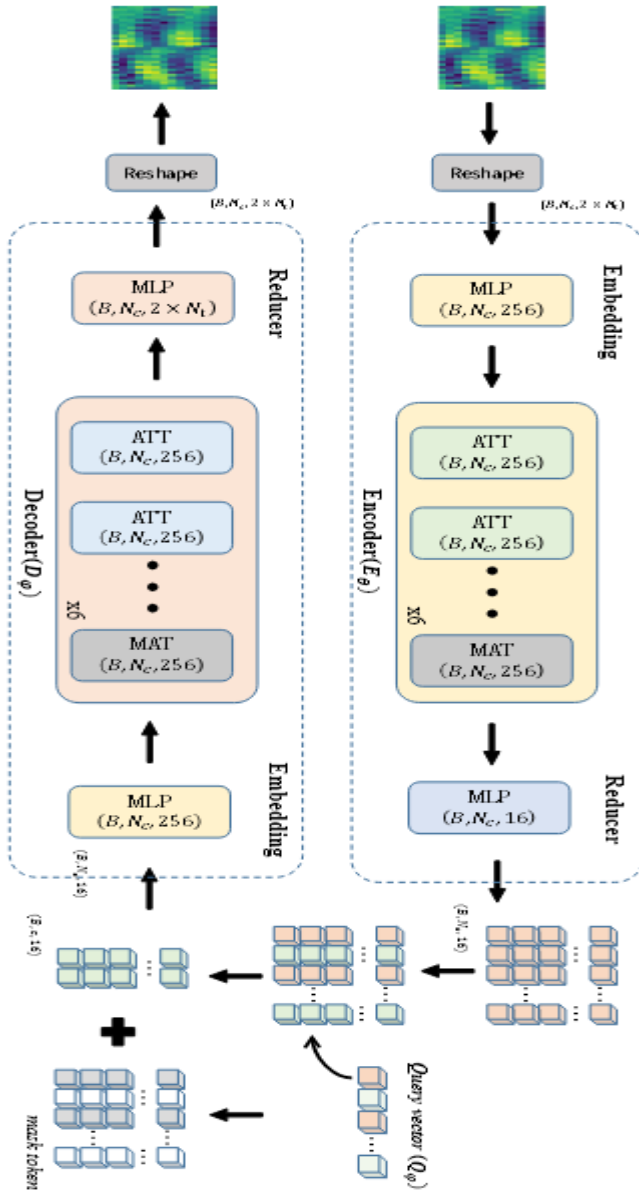
B. تخمین مشترک کانال و بازخورد کانال

برای تخمین کانال: (Channel Estimation)

هدف تخمین کانال، بازیابی ماتریس کانال شامل تمامی N_c زیرحامل‌ها بر اساس اطلاعات پیلوت‌های N_p است. اما پیلوت‌هایی که توسط گیرنده دریافت می‌شوند، دارای نویز محیطی قابل توجهی هستند که این موضوع تخمین دقیق کانال را دشوارتر می‌کند. برای بازیابی CSI لینک پایین‌دست با کیفیت بالا، لازم است که: یک شبکه حذف نویز (Denoising Network) طراحی کنیم تا نویز موجود در پیلوت‌ها را حذف کند. سپس پیلوت‌های تمیزشده را با استفاده از یک شبکه تکمیل (Completion Network) برای بازسازی کامل CSI به کار ببریم. همان‌طور که در شکل ۵ نشان داده شده، شبکه تخمین کانال ما شامل سه مؤلفه اصلی است: بلوک حذف نویز (Denoising Block) بلوک جاسازی توکن ماسک (Mask Token Embedding Block) رمزگشای بازیابی (Recovery Decoder) برای مقادیر مختلط (Complex) ماتریس CSI، ابتدا: ما قسمت حقیقی و موهومی را به هم وصل کرده و به صورت ورودی به شبکه می‌دهیم $R'N_c \times N_t \times 2$:



Decoder رمزگشا: ما ماتریس تکمیل‌شده را با قرار دادن توکن‌های ماسک مقداردهی اولیه می‌کنیم. رمزگشا نیز از مکانیسم Self-Attention استفاده می‌کند تا اطلاعات مؤثر موجود در ZPart را به سمت توکن ماسک منتقل کند. بازسازی داده از طریق چندین لایه انتقال پیام (Multi-layer Message Passing) انجام می‌شود. ما یک مسیر انتقال (Transfer Path) در لایه اول اضافه می‌کنیم تا نقشه ماسک معکوس M' در آن تعبیه شود. این مسیر انتقال به‌طور ویژه طراحی شده تا: توان انتقال اطلاعات مرتبط به توکن ماسک را تقویت کند، و در عین حال از آلوده شدن اطلاعات معتبر به وسیله داده‌های نامعتبر اولیه جلوگیری کند. فرایند رمزگشا به صورت زیر مدل‌سازی می‌شود:

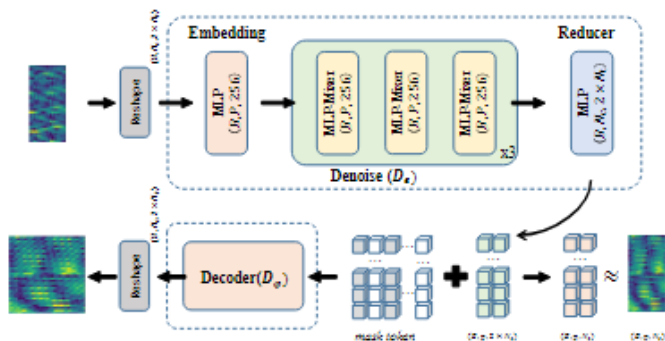


شکل ۴. ساختار یک مرحله تک‌ماسک-آنتنشن در MAT را نشان می‌دهد. در این شکل، علامت M' نشان‌دهنده ماتریس توجه ماسک پیشنهادی است.

تخمین کانال در سیستم ما در سمت تجهیزات کاربر (UE) انجام می‌شود، زیرا محدودیت منابع محاسباتی در UE باید در نظر گرفته شود. طراحی ما بر دستیابی به پیچیدگی کم تمرکز دارد و برای انجام حذف نویز، از MLP-Mixer سبک‌وزن استفاده می‌کنیم. شبکه مورد استفاده در سیستم ما، بعد اطلاعات پایلوت را از بعد ویژگی دوباره به بعد اصلی کاهش می‌دهد. علاوه بر این، ما از تابع زیان پیکسل‌به‌پیکسل (Pixel-wise Loss) استفاده می‌کنیم تا اطمینان حاصل شود که اطلاعات پایلوت با اطلاعات کامل بدون نویز که در Ground Truth وجود دارد، هماهنگ باشد. پس از به‌دست‌آوردن پایلوت تمیز، توکن‌های ماسک را جاسازی کرده و رمزگشای بازبازی را دقیقاً مشابه رمزگشای Dφ به کار می‌بریم. در نهایت، NMSE بین خروجی و اطلاعات کامل بدون نویز محاسبه می‌شود.

2. بازخورد کانال: (Channel Feedback)

هدف شبکه فشرده‌سازی و بازخورد کانال این است که: اطلاعات بی‌سیم را در سمت تجهیزات کاربر (UE) فشرده کند و آن را در ایستگاه پایه (BS) تا حد امکان دقیق بازبازی کند تحت محدودیت‌های تعیین‌شده سیستم.



شکل ۵. ساختار شبکه تخمین کانال را نشان می‌دهد. در این فرایند، حذف نویز قبل از مرحله تکمیل ضروری است. ما از MLP-Mixer سبک‌وزن برای پیاده‌سازی رمزگذار حذف نویز De استفاده می‌کنیم، و رمزگشا با استفاده از رمزگشای Dφ ماتریس کامل کانال را بازبازی می‌کند.

پهنای باند بیت. (Bit Bandwidth)

در کار ما، یک معماری مبتنی بر ترنسفورمر در رمزگذار و رمزگشا پیشنهاد می‌کنیم. همان‌طور که در شکل ۳ نشان داده شده، ورودی یک ماتریس بردار ویژه به ابعاد $N_c \times N_t$ است. از آنجا که مقادیر بردار ویژه مختلط هستند، ما بخش حقیقی و بخش موهومی را به هم

وصل می‌کنیم تا ورودی $N_c \times 2 \times N_t$ به‌دست آید. سپس این ورودی به رمزگذار فرستاده می‌شود. ما زیرحامل‌ها را به‌عنوان توکن‌های رمزگذار (FlowMat) با نماد θ_{θ_E} در نظر می‌گیریم. پس از استخراج ویژگی‌های عمیق Z شبکه یک بردار Query یاد می‌گیرد تا ویژگی‌های عمیق زیرحامل‌های خاص را انتخاب کند و تنها m مورد از آن‌ها را نگه می‌دارد. m یک آپرپارامتر استاین زیرمجموعه با نماد ZPart نشان داده می‌شود. کمی‌سازی: (Quantization) دو روش وجود دارد: کمی‌سازی یکنواخت (Uniform Quantization) کمی‌سازی برداری (VQ Quantization) براساس [33] در بخش V-B2 آن‌ها را به‌طور تجربی مقایسه می‌کنیم. در کمی‌سازی یکنواخت، جریان بیت، دنباله کمی‌سازی شده ZPart_{ZPart} است. در کمی‌سازی VQ، تنها اندیس‌های بردارهای انتخاب‌شده در کدبوک ارسال می‌شوند.

بازسازی (Recovery)

پس از کمی‌سازی، توکن‌های ماسک شده در فرایند رمزگشا استفاده می‌شوند تا اطلاعات بازبازی شود؛ مشابه همان چیزی که در بخش تخمین کانال توضیح داده شد. در نهایت، رمزگشا کل اطلاعات را بازبازی می‌کند.

3. برای وظایف مشترک و مستقل

ابتدا، مدل تخمین کانال برای برآورد کانال‌های از دست‌رفته استفاده می‌شود و بنابراین: درون‌یابی حوزه فرکانس حذف نویز انجام می‌گیرد. سپس، بر اساس کانال تخمین‌زده‌شده، بردارهای ویژه موردنیاز برای بازخورد در زیرحامل‌های مختلف محاسبه می‌شوند. در نهایت، مدل بازخورد کانال برای انجام: فشرده‌سازی بازخورد بردارهای ویژه کانال به کار گرفته می‌شود. ما یک مدل End-to-End می‌سازیم که بدون تغییر دادن ساختار شبکه، تخمین کانال + بازخورد کانال را هم‌زمان انجام می‌دهد. ورودی مدل: پایلوت‌ها خروجی مدل: بردارهای ویژه (Eigenvectors) ما همچنین به‌صورت جداگانه: یک شبکه تخمین کانال و یک شبکه بازخورد کانال را آموزش می‌دهیم که آن‌ها را splitted می‌نامیم و سپس عملکرد ترکیبی آن‌ها را ارزیابی می‌کنیم.

C. استراتژی آموزش (Training Strategy)

شبکه پیشنهادی ما از آموزش تدریجی (Progressive Training) و آموزش مشترک (Joint Training) استفاده می‌کند. در استراتژی آموزش تدریجی، ابتدا از زیان NMSE (خطای میانگین مربعات نرمال‌سازی‌شده) برای آموزش هم‌زمان رمزگذار و شبکه محدودشده حذف نویز استفاده می‌شود. این تابع زیان به‌صورت زیر تعریف می‌شود:

$$L_{CE_1} = \sqrt{\frac{\sum_{n=1}^T \sum_{k=1}^{N_p} (D_e(y_{n,k}) - H_{n,k})^2}{\sum_{n=1}^T \sum_{k=1}^{N_p} D_e(y_{n,k})^2}}, \quad (12)$$

که در آن D_e شبکه «حذف‌نویز» در شکل ۵ است، $y_{n,k}$ کانال k ام پیلوت ورودی است، T تعداد نمونه‌های آموزشی است، و $H_{i,k}$ بردار کانال ایده‌آل انتخاب‌شده متناظر با $y_{n,k}$ است. و رمزگشا با زیان NMSE آموزش داده می‌شود در حالی که پارامترهای رمزگذار و شبکه محدودشده حذف‌نویز ثابت نگه داشته می‌شوند. تابع زیان به صورت زیر فرمول‌بندی می‌شود:

$$L_{CE_2} = \sqrt{\frac{\sum_{n=1}^T \sum_{k=1}^{N_c} (D_e(D_\phi(y_{n,k})) - H_{n,k})^2}{\sum_{n=1}^T \sum_{k=1}^{N_c} D_e(D_\phi(y_{n,k}))^2}} \quad (13)$$

که در آن T تعداد نمونه‌های آموزشی است و H_n کل ماتریس کانال ایده‌آل است. آموزش مشترک، تمام بخش‌های شبکه تخمین کانال را به صورت هم‌زمان آموزش می‌دهد که معادله ۱۲ و معادله ۱۳ را به عنوان توابع زیان ترکیب می‌کند. تابع زیان شبکه فشرده‌سازی و بازخورد کانال چنین است:

$$L_{CF} = \frac{1}{T} \sum_{j=1}^T \frac{1}{N_c} \sum_{i=1}^{N_c} \frac{\|w_{i,j}^H w'_{i,j}\|}{\|w_{i,j}\| \|w'_{i,j}\|}, \quad (14)$$

که در آن N_c تعداد زیرحامل‌های هر نمونه است و $w_{i,j}$ و $w'_{i,j}$ به ترتیب بردارهای ویژه برجسب‌ها و بردارهای ویژه پیش‌بینی‌شده هستند.

5. آزمایش‌ها

A. توضیحات داده‌ها

داده‌ای که ما استفاده می‌کنیم «مجموعه داده باز ارتباطات موبایل ۲۰۲۲» است. این مجموعه داده دارای ۳۲ آنتن فرستنده و ۴ آنتن گیرنده است. تخمین کانال و بازخورد ویژگی‌های کانال بر اساس پیلوت‌های مشخص لازم است. این مجموعه داده دو نوع پیلوت دریافتی مختلف فراهم می‌کند: پیلوت با چگالی بالا (HD) و پیلوت با چگالی پایین (LD) به عنوان اطلاعات ورودی. هر یک از آن‌ها شامل ۳۰۰,۰۰۰ نمونه است. پیلوت‌های چگالی بالا ۲۶ بلوک منبع از ۵۲ بلوک منبع (بلوک‌های منبع با شماره‌های فرد) را برای ارسال اطلاعات پیلوت اشغال می‌کنند. ابعاد هر نمونه داده $4 \times 20.8 \times 4$ است. «۴» اول اشاره به چهار آنتن گیرنده دارد. «۲۰۸» اشاره به ۲۰۸ زیرحامل دارد (۲۶ بلوک منبع که هر کدام ۸ زیرحامل با اطلاعات پیلوت دارند)، و «۴» دوم به ۴ نماد OFDM اشاره دارد. برای پیلوت‌های چگالی

پایین، ۶ بلوک از ۵۲ بلوک منبع اشغال شده‌اند (شماره‌های ۷، ۱۵، ۲۳، ۳۱، ۳۹، ۴۷). ابعاد هر نمونه داده $4 \times 4.8 \times 4$ است، که در آن «۴۸» مربوط به ۶ بلوک منبع است که هر کدام دارای ۸ زیرحامل با پیلوت هستند.

اطلاعات کامل کانال در حوزه زمان شامل مجموعاً ۳۰۰,۰۰۰ نمونه است که به عنوان اطلاعات یکپارچه هم برای پیلوت‌های چگالی بالا و هم پیلوت‌های چگالی پایین استفاده می‌شود. ابعاد هر نمونه $4 \times 32 \times 64$ است که متناظر با ۴ آنتن گیرنده، ۳۲ آنتن فرستنده و ۶۴ نمونه برداری تأخیری است. اطلاعات کامل کانال در حوزه زمان می‌تواند برای بهبود عملکرد در فشرده‌سازی و بازخورد کانال به اطلاعات بردار ویژه تبدیل شود. این بردارهای ویژه توسط مجموعه داده ارائه شده‌اند. پهنای باند ارسال به ۱۳ زیرحامل تقسیم شده است که هر کدام شامل بردارهای ویژه ۳۲ بعدی هستند. در این بخش نیز ۳۰۰,۰۰۰ نمونه وجود دارد که به عنوان برجسب‌های یکپارچه برای پیلوت‌های چگالی بالا و چگالی پایین استفاده می‌شوند. ابعاد هر نمونه 13×32 است.

ما ۹۵٪ نخست نمونه‌ها را برای آموزش و ۵٪ آخر را برای آزمایش استفاده می‌کنیم، یعنی مجموعه آزمایش شامل ۱۵,۰۰۰ نمونه است.

B. نتایج مقایسه

ما ابتدا روش خود را با بهترین روش‌های موجود (SOTA) در هر یک از دو وظیفه به صورت جداگانه مقایسه می‌کنیم و سپس آزمایش‌هایی را روی تخمین مشترک کانال و بازخورد انجام می‌دهیم.

برای تخمین کانال، ما از معیار NMSE برای ارزیابی خطا بین کانال مطلوب و کانال تخمینی استفاده می‌کنیم که به صورت زیر محاسبه می‌شود:

$$NMSE = \frac{\sum_{n=1}^T \sum_{k=1}^{N_p} (D_e(y_{n,k}) - H_{n,k})^2}{TN_p \sum_{n=1}^T \sum_{k=1}^{N_p} D_e(y_k)^2}, \quad (15)$$

که در آن y پیلوت‌های ورودی و H' ماتریس کانال ایده‌آل است. برای نمایش واضح‌تر نتایج، NMSE را بر حسب دسی‌بل (dB) بیان می‌کنیم. برای بازخورد کانال، ما از شباهت کسینوسی بین بردار ویژه کانال بازخوردشده و برجسب به عنوان معیار استفاده می‌کنیم که به صورت زیر محاسبه می‌شود:

$$\text{Rho} = \frac{1}{T} \sum_{j=1}^T \frac{1}{N_c} \sum_{i=1}^{N_c} \frac{\|w_{i,j}^H w'_{i,j}\|}{\|w_{i,j}\| \|w'_{i,j}\|}, \quad (16)$$

که در آن T تعداد نمونه‌های آزمون، N_c تعداد زیرحامل‌های هر نمونه، و w_i, z_i و w_i', z_i' به ترتیب بردارهای ویژه برجسته‌ها و بردارهای ویژه پیش‌بینی شده هستند.

تخمین کانال: ما روش خود را با چهار روش SOTA تخمین کانال مقایسه می‌کنیم [6] ChannelNet، [8] AttentionNet، [7] CBDNet و [20] Attention-Aid. مسئله را مشابه «فوق تفکیک‌پذیری تصویر» و «حذف نویز تصویر» در نظر می‌گیرد. در ChannelNet، پایلوت‌ها به عنوان تصاویر با وضوح پایین در نظر گرفته می‌شوند و شبکه فوق تفکیک‌پذیری [34] SRCNN به همراه شبکه حذف نویز [35] DNCNN برای تخمین کانال به کار می‌روند. AttentionNet یک ساختار جدید رمزگذار-رمزگشا ارائه می‌دهد که شامل رمزگذار ترنسفورمر و رمزگشای CNN با باقیمانده است CBDNet. یک شبکه CNN بهبود یافته ارائه می‌دهد که از شبکه تخمین سطح نویز، شبکه حذف نویز غیرکور و تابع زیان نامتقارن مشترک برای تخمین کانال استفاده می‌کند Attention-Aid. یک چارچوب جدید تخمین کانال مبتنی بر یادگیری عمیق همراه با مازول‌های توجه کانالی برای سیستم‌های MIMO بزرگ مقیاس سنتی معرفی می‌کند. ما از معادله ۱۶ برای ارزیابی استفاده می‌کنیم و عملکرد در جدول ۱ نشان داده شده است. نتایج نشان می‌دهد که روش ما در وظیفه صرفاً تخمین کانال بهتر از سایر مدل‌ها عمل می‌کند. روش ما به دست می‌آورد

6.6106- و 5.9438 را در مجموعه آزمون با چگالی بالا و چگالی پایین به دست می‌آورد (کمتر بودن بهتر است)، که نسبت به CBDNet حدود ۱۴٪ و ۱۱٪ بهبود دارد. در مقایسه با روش مبتنی بر توجه، Attention-Aid، روش ما به ترتیب ۲۱٪ و ۳۹٪ بهبود ایجاد می‌کند. طرح Masked Token ما با افزودن توکن‌های قابل یادگیری، نگاشت بهتری میان ویژگی‌های عمیق و اطلاعات کانال آموخته است و به همین دلیل به نتایج تخمین بهتری دست یافته است ChannelNet. یک روش مبتنی بر CNN است که حذف نویز و تخمین را در دو مرحله جداگانه انجام می‌دهد. اگرچه روش آن‌ها در مجموعه داده چگالی بالا نتایج بسیار خوبی به دست می‌آورد (رتبه دوم میان همه روش‌ها)، اما عملکرد آن در داده‌های چگالی پایین رضایت‌بخش نیست (رتبه آخر میان همه روش‌ها). در مقابل، روش ما می‌تواند در هر دو حالت عملکرد خوبی به دست آورد. برای پیچیدگی محاسباتی، نتایج FLOPs (تعداد عملیات ممیز شناور) و Params (تعداد پارامترهای شبکه) در جدول III نشان داده شده‌اند. در میان روش‌های مقایسه شده، FlowMat کمترین FLOPs را با کمی پارامتر بیشتر دارد.

2. فقط فشرده‌سازی و بازخورد: CSI

برای فشرده‌سازی و بازخورد CSI، ما روش خود را با شش روش SOTA مقایسه می‌کنیم [10] CsiNet و [14] CsiNet+ و ImCsiNet [36] و [37] DCRNet و [38] EVCsiNet و EVCsiNet-T [39] و CsiNet+ شبکه‌های بازخورد مبتنی بر یادگیری عمیق هستند و شامل لایه‌های کانولوشن، لایه‌های Fully-Connected و BatchNorm هستند ImCsiNet. یک شبکه بازخورد ضمنی (Implicit Feedback) است که شامل ساختار LSTM، لایه‌های Fully-Connected و BatchNorm است DCRNet. یک شبکه بازخورد CSI مبتنی بر کانولوشن با اتساع (Dilated Convolution) است، که با بزرگ کردن میدان دریافت بدون افزایش اندازه فیلتر، عملکرد را بهبود می‌دهد EVCsiNet. روشی مبتنی بر بردارهای ویژه کانال است که در آن مجموعه بردارهای ویژه به دست آمده از چندین زیرحامل، در رمزگذار فشرده و در رمزگشا بازسازی می‌شوند. ساختار Residual Block [40] در EVCsiNet استفاده شده است. EVCsiNet-T ...

TABLE I
QUANTITATIVE COMPARISON OF CHANNEL ESTIMATION ONLY ON THE DATASET. LOWER IS BETTER.

Method	HD NMSE (dB)	LD NMSE (dB)
ChannelNet	-6.3720	-2.2810
CBDNet	-5.6891	-5.2179
AttentionNet	-5.3929	-2.8207
Attention-Aid	-5.1555	-4.0300
Ours	-6.6106	-5.9438

TABLE II
QUANTITATIVE COMPARISON OF CHANNEL COMPRESSION AND FEEDBACK ONLY ON THE DATASET.

Method	64 bit	128 bit	256 bit
CsiNet	0.7717	0.8140	0.8433
CsiNet+	0.7634	0.7685	0.7894
ImCsiNet	0.8514	0.8825	0.9240
DCRNet	0.7764	0.7965	0.8611
EVCsiNet	0.8308	0.8687	0.9247
EVCsiNet-T	0.8532	0.9292	0.9632
Ours(VQ)	0.8922	0.9283	0.9464
Ours(UQ)	0.9015	0.9333	0.9648

TABLE III
COMPARISON OF COMPUTATIONAL COMPLEXITY OF ESTIMATION METHODS. IN THE 1ST LINE (SETTING) 'HIGH DENSITY' AND 'LOW DENSITY' MEAN HIGH AND LOW DENSITY PILOT DATASET RESPECTIVELY.

Setting	High density	Low density
Method	FLOPs/Params	FLOPs/Params
ChannelNet	2.59G/0.194M	2.59G/0.194M
Attention-Aid	2.25G/0.169M	2.25G/0.169M
AttentionNet	3.95G/5.29M	9.12G/5.28M
CBDNet	7.76G/4.33M	7.76G/4.33M
OursEstimation	1.18G/5.30M	1.04G/5.30M

TABLE IV
COMPARISON OF COMPUTATIONAL COMPLEXITY OF FEEDBACK METHODS. 'N Bit' MEANS THE COMPRESSION BIT.

Setting	64 Bit	128 Bit
Method	FLOPs/Params	FLOPs/Params
CsiNet	1.63M/0.217M	1.87M/0.430M
CsiNet+	9.51M/0.236M	9.78M/0.449M
DCRNet	0.99M/0.215M	1.36M/0.429M
EVCsiNet	1.60G/5.18M	1.61G/5.65M
EVCsiNet_T	0.21G/16.67M	0.21G/17.52M
ImCsiNet	0.83G/821M	0.84G/821M
OursCompress	1.53G/9.51M	1.54G/9.55M
JCEF	2.54G/1.47M	2.54G/2.75M

روش EVCsiNet-T، روش EVCsiNet را گسترش داده است که شامل (i) تحلیل و پیش‌پردازش داده کانال، (ii) طراحی شبکه عصبی، و (iii) بهبود کمی‌سازی است. ما از معادله ۱۶ برای ارزیابی استفاده می‌کنیم و عملکرد در جدول II نشان داده شده است. ما عملکرد فشرده‌سازی و بازسازی بردارهای ویژه اصلی را گزارش می‌کنیم. جدول II همچنین نشان می‌دهد که روش ما در تمام سطوح سربار بیت بهتر از سایر مدل‌ها است. برای مثال، روش ما مقدار 0.9015 در 64 بیت به‌دست می‌آورد که نسبت به CsiNet حدود 14.53٪ بهبود دارد. FlowMat و EVCsiNet-T از یک ساختار ترنسفورمر یکسان استفاده می‌کنند، و FlowMat در 64 بیت حدود 3.59٪ نسبت به آن بهبود دارد. دلیل اصلی این است که لایه آخر رمزگذار در EVCsiNet-T خروجی را به‌طور مستقیم به یک بردار کم‌بعد کاهش می‌دهد. این کار باعث از دست رفتن اطلاعات مهم شده و کار رمزگشا را برای ارزیابی داده اصلی دشوار می‌کند. بردار ویژه نسبت به DFT چگال‌تر است، بنابراین عملکرد روش‌های CNN پایین‌تر از LSTM و Transformer است، چون LSTM و Transformer از همبستگی میان زیرحامل‌ها استفاده می‌کنند. در روش‌های CNN، CsiNet+ از کرنل 7×7 برای به‌دست آوردن ویژگی کلیدی استفاده می‌کند، که در داده بردار ویژه اطلاعات بیشتری نسبت به کرنل 3×3 در CsiNet از دست می‌دهد. مشابه بخش «تخمین تنها کانال» برای مقایسه پیچیدگی محاسباتی، نتایج FLOPs و Params در جدول IV نشان داده شده‌اند.

روش ما کمتر از EVCsiNet و CsiNet است و Params ما نیز از EVCsiNet-T و ImCsiNet کمتر است.

3. تخمین و بازخورد مشترک کانال‌ها آن‌جا که روش‌های مشترک تخمین کانال و بازخورد کانال کم هستند، ما فقط یک روش مشترک یعنی [16] JCEF را مقایسه می‌کنیم. JCEF ابتدا اطلاعات کامل کانال را تکمیل می‌کند، سپس با کمی‌سازی یکنواخت آن را فشرده کرده و در نهایت اطلاعات را بازسازی می‌کند. در همین حال، ما بعضی از روش‌های تخمین کانال را با روش بازخورد کانال ترکیب می‌کنیم

TABLE V
ABLATION STUDY: THE LOSS FUNCTION FOR CHANNEL ESTIMATION.

Method	L1+L1	L1+NMSE	NMSE+NMSE
High density	0.9748	0.9770	0.9774
Low density	0.9330	0.9363	0.9366

TABLE VI
ABLATION STUDY: TOKENS' CONTENT AND WHETHER UPDATING. U REPRESENTS PARAMETER UPDATE. THE BEST AND SECOND BEST RESULTS ARE SHOWN IN **BOLD** AND UNDERLINED.

Method	zero+ U	randn+ U	zero+ w.o. U	randn+ w.o. U
64 bit	0.9015	0.9036	0.9023	<u>0.9030</u>
128 bit	<u>0.9333</u>	0.9315	0.9330	0.9348
256 bit	0.9648	<u>0.9635</u>	0.9618	0.9588

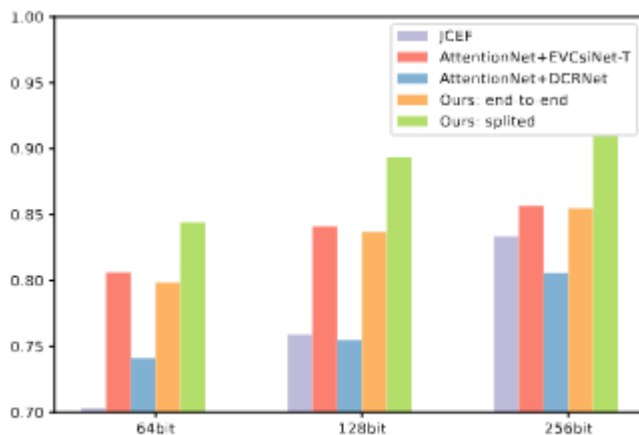
روش بازخورد کانال را با روش پیشنهادی خود مقایسه می‌کنیم. این ترکیب‌ها با نام‌های «AttentionNet+EVCsiNet-T» و «AttentionNet+DCRNet» نشان داده می‌شوند. ما از معادله ۱۶ برای ارزیابی استفاده می‌کنیم و عملکرد در شکل ۶ (a) ارائه شده است. در پایلوت‌های با چگالی پایین شکل ۶ (b) مشاهده می‌شود که مقدار Rho هنگام استفاده از مدل ما به ترتیب 0.0378، 0.0526 و 0.053 در ۶۴ بیت، ۱۲۸ بیت و ۲۵۶ بیت افزایش می‌یابد. در پایلوت‌های با چگالی بالا شکل ۶ (a) نیز مقدار Rho هنگام استفاده از مدل ما به ترتیب 0.0227، 0.0220 و 0.0149 در ۶۴ بیت، ۱۲۸ بیت و ۲۵۶ بیت افزایش می‌یابد. این نتایج نشان‌دهنده اثربخشی FlowMat در تخمین مشترک کانال و فشرده‌سازی کانال است.

C. مطالعه حذف مؤلفه‌ها (Ablation Study)

در این بخش، ما آزمایش حذف مؤلفه‌ها را روی مجموعه داده چگالی پایین انجام می‌دهیم.

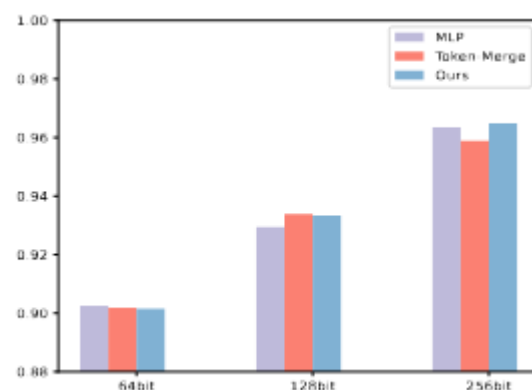
1) راهکار End-to-End در مقابل آموزش جداگانه (splited) برای نشان دادن اثربخشی خط لوله آموزش جداگانه، مدلی با نام «Ours» در شکل ۶ (a) می‌سازیم و شکل ۶ (b) روی پایلوت‌های چگالی بالا آزمایش می‌شود. ما این مدل را با روش End-to-End برای تخمین اطلاعات بدون تغییر ساختار شبکه مقایسه می‌کنیم. ورودی، بازخورد همراه با پایلوت‌ها است و خروجی، بردارهای ویژه است و معادله ۱۴ به عنوان تابع زیان استفاده می‌شود. در پایلوت‌های چگالی پایین دیده می‌شود که مقدار Rho در روش splited در ۶۴ بیت، ۱۲۸ بیت و ۲۵۶ بیت بیشتر از End-to-End است. همین پدیده در پایلوت‌های چگالی بالا نیز مشاهده می‌شود. روش ما در حالت splited بهترین عملکرد را میان روش‌های مقایسه شده دارد. در مقایسه با روش‌های End-to-End دیگر نیز روش ما بهتر عمل می‌کند برای نمونه، JCEF و AttentionNet+DCRNet را ببینید و بهترین عملکرد را در چارچوب End-to-End دارد.

2) تابع زیان برای تخمین کانال برای تخمین کانال، ما همچنین از توابع زیان دیگر برای مطالعه حذف مؤلفه‌ها استفاده می‌کنیم. در مدل اصلی، زیان NMSE برای حذف نویز و تخمین کانال به کار می‌رود. در جدول ۷ با نام «NMSE+NMSE» مشخص شده است. ما چند مدل دیگر نیز برای مقایسه می‌سازیم «L1+L1»: و «L1+NMSE» عبارت قبل از «+» نشان‌دهنده زیان مورد استفاده برای حذف نویز و عبارت بعد از «+» مربوط به زیان تخمین است. مشاهده می‌شود که میان این روش‌ها، «NMSE+NMSE» بهترین نتیجه را دارد. برای مثال، در مجموعه آزمون چگالی پایین به ترتیب حدود 0.0003 و 0.0036 بهتر از «L1+L1» و «L1+NMSE» است.

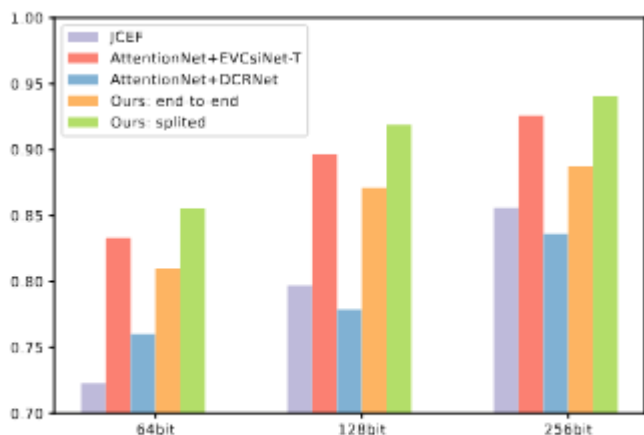


(b) Low Density

شکل ۶. مقایسه عددی تخمین مشترک کانال و فشرده سازی کانال روی مجموعه داده. مقدار بالاتر بهتر است.



شکل ۷. مطالعه حذف مؤلفه‌ها: روش کاهش توکن در فشرده سازی CSI.



(a) High Density

3. روش کاهش توکن در بازخورد CSI در مدل فشرده سازی و بازخورد کانال خود، ما تکنیک انتخاب کانال و جاسازی ماسک را برای کاهش و گسترش توکن‌ها پیشنهاد می‌کنیم. روش‌های دیگری نیز برای کاهش و افزایش توکن وجود دارند. برای مثال، لایه‌های MLP یک روش رایج برای کاهش و افزایش ابعاد هستند [8], [16]. ما از لایه‌های MLP و همچنین روش میانگین گیری توکن‌ها (Token Average Merging) برای مقایسه استفاده می‌کنیم. در روش MLP، یک MLP برای کاهش توکن و یک MLP دیگر برای افزایش توکن استفاده می‌شود. در روش میانگین گیری توکن، همه توکن‌های ورودی را به طور مساوی گروه بندی کرده و مقدار میانگین هر گروه را به عنوان توکن کاهش یافته محاسبه می‌کنیم. برای گسترش توکن‌ها، روش مشابه مدل اصلی است. ما آزمایش‌ها را فقط روی فشرده سازی و بازخورد کانال انجام دادیم و نتایج اثربخشی روش پیشنهادی ما را نشان می‌دهد. در شکل ۷، «Token merging» و «MLP» به ترتیب نشان‌دهنده استفاده از

لایه‌های MLP و میانگین‌گیری توکن‌ها هستند. نتایج نشان می‌دهد که اگرچه روش MLP در فشرده‌سازی با بیت پایین به میزان 0.0011 کمی بهتر از روش ما است، اما روش پیشنهادی ما در فشرده‌سازی با بیت متوسط و بالا به میزان 0.0038 و 0.0014 عملکرد بهتری دارد. در مقایسه با «Token merging»، روش ما در نرخ بیت پایین عملکرد مشابهی دارد، اما در نرخ بیت بالا برتری قابل‌توجهی به میزان 0.006 نشان می‌دهد.

4. محتوای و نحوه به‌روزرسانی آن‌ها: محتوای اولیه توکن‌های ماسک‌شده و این که آیا پارامترهای آن‌ها در طی آموزش به‌روزرسانی شوند یا خیر، بر عملکرد تأثیر می‌گذارد. بنابراین، ما چهار حالت زیر را آزمایش کردیم «zero+update»: مقداردهی اولیه صفر + به‌روزرسانی «randn+update»: مقداردهی اولیه نرمال تصادفی + به‌روزرسانی «zero+without update»: مقداردهی اولیه صفر + بدون به‌روزرسانی «randn+without update»: مقداردهی اولیه نرمال تصادفی + بدون به‌روزرسانی نتایج در جدول VI نشان داده شده است. مشاهده می‌شود که در فشرده‌سازی با بیت بالا و پایین، مقداردهی اولیه تصادفی و به‌روزرسانی تأثیر متفاوتی دارند. وقتی نرخ بیت پایین است (۶۴ بیت)، مقداردهی اولیه با توزیع نرمال ('randn') تأثیر غالب را دارد و به‌روزرسانی تأثیر ضعیفی دارد. وقتی نرخ بیت بالا است (۱۲۸ بیت)، تأثیر به‌روزرسانی غالب می‌شود و به‌روزرسانی پارامترهای توکن عملکرد را تقویت می‌کند.

5. توکن‌ها باید ماسک شوند؟ در نهایت، ما مطالعه حذف مؤلفه‌ها را درباره این که کدام توکن‌ها باید ماسک شوند انجام دادیم. در مدل اصلی، یک بردار Query قابل یادگیری آموزش می‌دهیم تا مشخص کند کدام توکن‌ها باید ماسک شوند. برای مقایسه، یک مدل دیگر ساختیم که بردار Query آن در ابتدا به‌صورت تصادفی انتخاب شده و در طول آموزش و آزمون ثابت نگه داشته می‌شود. نتیجه نشان می‌دهد که عملکرد Query قابل یادگیری و Query ثابت تقریباً یکسان است. اما چون Query قابل یادگیری در برخی موارد کمی بهتر عمل می‌کند، ما در مقاله از همان بردار Query قابل یادگیری استفاده کرده‌ایم.

6. نتیجه‌گیری‌ها

در این مقاله، ما از همبستگی موجود در حوزه فرکانس ذاتی میان ماتریس‌های کانال بهره می‌بریم و یک چارچوب یکپارچه با نام FlowMat برای تخمین و بازخورد مشترک کانال در سیستم‌های مسیو MIMO ارائه می‌کنیم FlowMat. از معماری رمزگذار-رمزگشا بهره می‌برد. کل شبکه رمزگذار-رمزگشا با استفاده از توکن ماسک قابل یادگیری برای بازخورد کانال به کار گرفته می‌شود و عملکرد بسیار عالی‌ای در بازخورد کانال به دست می‌آورد. رمزگشا برای تخمین کانال استفاده

می‌شود. تخمین کانال توسط رمزگذاری با ساختاری مشابه ساختار بخش بازخورد کانال انجام می‌گیرد؛ جایی که یک ماژول حذف نویز مبتنی بر پرسپترون چندلایه سبک‌وزن برای افزایش دقت تخمین به کار گرفته می‌شود. ما آزمایش‌های گسترده‌ای انجام داده‌ایم که عملکرد برتر FlowMat را—چه در وظایف مشترک و چه در وظایف مستقل تخمین کانال و بازخورد کانال نشان می‌دهد. علاوه بر این، FlowMat کاهش قابل‌توجهی در پیچیدگی محاسباتی ارائه می‌کند، به‌ویژه در بخش تخمین کانال.

مراجع

- [1] M. T. L., "Massive MIMO: An Introduction," Bell Labs Technical Journal, 2015.
- [2] F. Tariq, M. R. A. Khandaker, K. K. Wong, et al., "A speculative study on 6G," IEEE Wireless Communications, 2020.
- [3] F. Rusek, D. Persson, B. K. Lau, et al., "Scaling up MIMO: Opportunities and challenges with very large arrays," Signal Processing Magazine, 2012.
- [4] W. G. Song and J. T. Lim, "Channel estimation and signal detection for MIMO-OFDM with time-varying channels," IEEE Transactions on Signal Processing, 2006.
- [5] J. Ma and P. L., "Data-aided channel estimation in large antenna systems," IEEE Transactions on Signal Processing, 2014.
- [6] S. Mehran, P. Vahid, M. Ali, and S. Hamid, "Deep learning-based channel estimation," IEEE Communications Letters, 2019.
- [7] Y. Jin, J. Zhang, B. Ai, and X. Zhang, "Channel estimation for mmWave massive MIMO with convolutional blind denoising network," IEEE Communications Letters, 2020.
- [8] D. Luan and J. Thompson, "Attention-based neural networks for wireless channel estimation," Vehicular Technology Conference, 2022.
- [9] 3GPP, "NR; Physical layer procedures for data (Release 16)," 3GPP TS 38.214 V16.1.0, 2020.
- [10] C.-K. Wen, W.-T. Shih, and S. Jin, "Deep learning for massive MIMO CSI feedback," IEEE Wireless Communications Letters, vol. 7, no. 5, pp. 748–751, 2018.
- [11] Z. Lu, J. Wang, and J. Song, "Multi-resolution CSI feedback with deep learning in massive MIMO system," ICC 2020, IEEE, pp. 1–6.
- [12] Z. Cao, W.-T. Shih, J. Guo, C.-K. Wen, and S. Jin, "Lightweight CNNs for CSI feedback in massive MIMO," IEEE Communications Letters, vol. 25, no. 8, 2021.
- [13] Y. Sun, W. Xu, L. Fan, G. Y. Li, and G. K. Karagiannidis, "Ancinet: An efficient DL approach for CSI feedback compression," IEEE Wireless Communications Letters, vol. 9, no. 12, 2020.
- [14] J. Guo, C.-K. Wen, S. Jin, and G. Y. Li, "CNN-based multi-rate compressive sensing for CSI feedback," IEEE Transactions on Wireless Communications, vol. 19, no. 4, 2020.
- [15] G. Z. Ma and X., "Data-driven deep learning to design pilot and channel estimator for massive MIMO," IEEE Transactions on Vehicular Technology, 2020.

- [16] C. Tong, J. Guo, C.-K. Wen, S. Jin, G. Y. Li, X. Wang, and X. Hou, "Deep learning for joint channel estimation and feedback in massive MIMO," *arXiv: Information Theory*, 2020.
- [17] A. Vaswani, N. Shazeer, N. Parmar, et al., "Attention is all you need," *NeurIPS*, 2017.
- [18] H. Xiao, Z. Wang, D. Li, et al., "A transformer backbone for CSI feedback," *arXiv preprint*, 2022.
- [19] Y. Xu, M. Yuan, and M.-O. Pun, "Transformer empowered CSI feedback for massive MIMO," *Wireless and Optical Communications Conference*, 2021.
- [20] J. Gao, M. Hu, C. Zhong, G. Y. Li, and Z. Zhang, "An attention-aided DL framework for massive MIMO channel estimation," *IEEE Transactions on Wireless Communications*, 2021.
- [21] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep CNNs," *IEEE TPAMI*, vol. 38, no. 2, pp. 295–307, 2015.
- [22] C.-J. Chun, J.-M. Kang, and I.-M. Kim, "DL-based channel estimation for massive MIMO systems," *IEEE Wireless Communications Letters*, vol. 8, no. 4, 2019.
- [23] H. He, C.-K. Wen, S. Jin, and G. Y. Li, "DL-based channel estimation for beam-space mmWave massive MIMO," *IEEE Wireless Communications Letters*, 2018.
- [24] M. B. Mashhadi and D. Gündüz, "Pruning the pilots: DL-based pilot design and channel estimation," *IEEE Transactions on Wireless Communications*, 2021.
- [25] M. B. Mashhadi, D. Gündüz, and Q. Yang, "Distributed deep convolutional compression for massive MIMO CSI feedback," *IEEE Transactions on Wireless Communications*, 2020.
- [26] C. Lu, W. Xu, H. Shen, J. Zhu, and K. Wang, "MIMO channel info feedback using deep recurrent network," *IEEE Communications Letters*, vol. 23, no. 1, 2018.
- [27] M. Chen, J. Guo, C.-K. Wen, S. Jin, G. Y. Li, and A. Yang, "DL-based implicit CSI feedback in massive MIMO," *IEEE Transactions on Communications*, 2021.
- [28] W. Liu, W. Tian, H. Xiao, et al., "EVCsiNet: Eigenvector-based CSI feedback under 3GPP channels," *IEEE Wireless Communications Letters*, 2021.
- [29] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Transactions on Information Theory*, vol. 52, no. 11, 2006.
- [30] K. He, X. Chen, S. Xie, Y. Li, P. Dollár, and R. Girshick, "Masked autoencoders are scalable vision learners," *CVPR*, 2022.
- [31] X. Chen, S. Xie, and K. He, "An empirical study of training self-supervised vision transformers," *ICCV*, 2021.
- [32] H. Bao, L. Dong, and F. Wei, "BEiT: BERT pre-training of image transformers," *arXiv*, 2021.
- [33] A. van den Oord, O. Vinyals, and K. Kavukcuoglu, "Neural discrete representation learning," *NeurIPS*, 2017.
- [34] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep CNNs," *IEEE TPAMI*, 2014.
- [35] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE TIP*, vol. 26, no. 7, 2017.
- [36] M. Chen, J. Guo, C.-K. Wen, et al., "DL-based implicit CSI feedback in massive MIMO," *IEEE Transactions on Communications*, 2022.
- [37] S. Tang, J. Xia, L. Fan, X. Lei, W. Xu, and A. Nallanathan, "Dilated convolution based CSI feedback compression," *IEEE Transactions on Vehicular Technology*, 2022.
- [38] W. Liu, W. Tian, H. Xiao, et al., "EVCsiNet: Eigenvector-based CSI feedback," *IEEE Wireless Communications Letters*, 2021.
- [39] H. Xiao, Z. Wang, W. Tian, et al., "AI enlightens wireless communication: CSI feedback analyses and solutions," *China Communications*, 2021.
- [40] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CVPR*, 2016.