

طراحی زمانبندی و تجمعیع برای یادگیری فدرال ناهمzman از طریق بیاسیم شبکه‌ها

چونگ هسوان هو، ژنگ چن و اریک جی. لارسون

چکیده - یادگیری فدرال (FL) یک چارجوب یادگیری ماشین (ML) مشارکتی است که آموزش روی دستگاه و تجمعیع مبنی بر سور را برای آموزش یک مدل یادگیری ماشین مشترک در بین عامل‌های توزیع‌شده ترکیب می‌کند. در این کار، ما یک طراحی یادگیری ماشین ناهمzman با تجمعیع دوره‌ای را برای مقابله با مشکل سرگردان در سیستم‌های FL پیشنهاد می‌کنیم. با توجه به منابع ارتباطی بیاسیم محدود، ما تأثیر سیاست‌های زمانبندی مختلف و طرح‌های تجمعیع را بر عملکرد همگایین برسیم. با توجه به اهمیت کاهش پایاس و واپسی پیروزرسانی‌های مدل تجمعیع شده، ما یک سیاست زمانبندی پیشنهاد می‌کنیم که به طور مشترک کیفیت کاتال و مسئله پراکنده شناخته می‌آشود. یک رویکرد برای این مسئله، تغییر رویه همزمان به یک رویه ناهمzman است، یعنی سور نیازی ندارد منتظر بماند تا همه دستگاه‌ها آموزش محلی را قبل از انجام تجمعیع پیروزرسانی‌ها به پایان برسانند.

اصطلاحات شاخص - یادگیری فدرال، آموزش ناهمzman، شبکه‌های بیاسیم، زمانبندی، تجمعیع

پخش مدل، آموزش محلی و تجمعیع مدل.
در هر تکرار، فرآیند تجمعیع مدل تنها زمانی می‌تواند شروع شود که همه دستگاه‌ها آموزش محلی را به پایان رسانده باشند. بنابراین، مدت زمان یک تکرار به طور جدی توسط کندترین دستگاه محدود می‌شود [3]. این پدیده که معمولاً در روش‌های همزمان FL مشاهده می‌شود، به عنوان مسئله پراکنده شناخته می‌آشود. یک رویکرد برای این مسئله، تغییر رویه همزمان به یک رویه ناهمzman است، یعنی سور نیازی ندارد منتظر بماند تا همه دستگاه‌ها آموزش محلی را قبل از انجام تجمعیع پیروزرسانی‌ها به پایان برسانند.

در مقالات علمی، چنین چارجوب FL ناهمzman در بسیاری از محیط‌های یادگیری عمیق [5]، [4] به کار گرفته شده است. با این حال، FL کاملاً ناهمzman با پیروزرسانی متوازن [5]، [4] می‌تواند به تبادل مکرر مدل، منجر به هزینه‌های بالای ارتباطی شود. از این رو، ما یک چارجوب FL ناهمzman با تجمعیع دوره‌ای پیشنهاد می‌کنیم که اثر سرگردانی را بدون پیروزرسانی بیش از حد مدل و تبادل اطلاعات بین سور و دستگاه‌های شرکت‌کننده از بین می‌برد. در مقایسه با سایر کارهای موجود در FL با پیروزرسانی‌های ناهمzman، [8]-[6] طرح پیشنهادی ما به راحتی قابل پیاده‌سازی است و به مقدار کم اطلاعات جانبی نیاز دارد.

مقدمه

آموزش مدل‌های یادگیری ماشین (ML) معمولاً به حجم عظیمی از داده‌ها نیاز دارد. امروزه، تعداد روزافزون دستگاه‌های متصفح به شکیب، با ارائه مجموعه‌های بزرگ از داده‌ها که می‌توانند برای آموزش مدل استفاده قرار گیرند، به توسعه الگوریتم‌های یادگیری ماشین کمک کرده است. از آنجایی که نگارنده‌های مربوط به حجم خصوصی در جامعه ما حیاتی شدنداشت، استفاده از داده‌های خصوصی از دستگاه‌های کاربر برای آموزش مدل‌های یادگیری ماشین دشوار شده است. بنابراین، یادگیری فدرال (FL) با پذراش اطلاعات روی دستگاه به دلیل مزایای آن در حفظ حجم خصوصی داده‌ها پیشنهاد شده است. FL یک چارجوب یادگیری ماشین مشارکتی است که در آن چندین دستگاه در آموزش یک مدل جهانی مشترک بر اساس داده‌های محلی موجود شرکت می‌کنند. [1] برخلاف معتمدی یادگیری ماشین، فقط پارامترهای مدل بین دستگاه‌های کاربر و یک سور پارامتر به اشتراک گذاشته می‌شوند. به دلیل نامگون دستگاه‌های شرکت‌کننده، داده‌ای آموزشی باید به صورت مرکزی ذخیره شوند، در یک سیستم پیکسان (iid) باشند، که FL را از چارجوب‌های بهینه‌سازی توزیع‌شده مرسوم که در آن‌ها داده‌های همگن و یکنواخت توزیع‌شده فرض می‌شوند، منحرف می‌کنند.

محدودیت منابع ارتباطی یک دیگر از مسائل حیاتی در سیستم‌های بیاسیم FL است. از آنجایی که تبادل مدل از طریق کاتال‌های بیاسیم انجام می‌شود، عملکرد سیستم (هزینه‌های ارتباطی و تأخیر) به طور طبیعی از محدودیت منابع فرکانس/ زمان رنج می‌پردازد، به خصوص زمانی که تعداد دستگاه‌های شرکت‌کننده زیاد باشد. یک راه حل ممکن برای کاهش بار ارتباطی، اجازه دادن به بخشی از دستگاه‌های شرکت‌کننده برای آپلود پیروزرسانی‌های محلی خود برای تجمعیع مدل است.

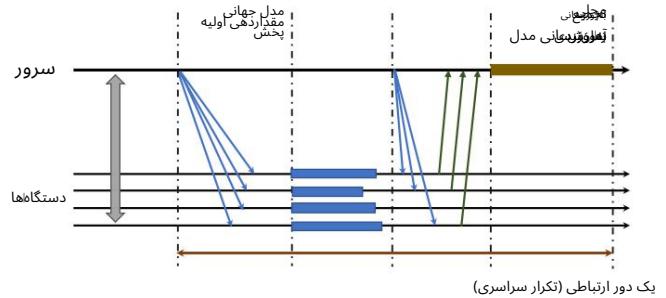
سپس، بسته به منابع ارتباطی اختصاص داده شده و کیفیت لینک بیاسیم، هر دستگاه پیروزرسانی‌های مدل خود را بر این اساس فشرده می‌کند تا پیروزرسانی‌های فشرده شده با توجه به منابع اختصاص داده شده، به طور اعتمادی منتقل شوند. زمانبندی دستگاه و تخصیص منابع ارتباطی برای دستیابی به یادگیری ماشینی کارآمد از نظر ارتباطی در شبکه‌های بیاسیم بسیار مهم هستند. [10]، [9] به طور شهودی، دستگاه‌هایی که تأثیر پیشتری بر عملکرد یادگیری دارند باید در زمانبندی در اولویت قرار گیرند. این هدف ارتباطی مبنی بر یادگیری در تضاد با طراحی مرسوم مبنی بر نرخ است که در شبکه‌های سلولی اتخاذ می‌شود، جایی که هدف دستیابی به کارایی طیفی بالاتر با توان عملیاتی شبکه است. چندین کار موجود، معیارهای مختلفی را برای نشان دادن اهمیت پیروزرسانی‌های محلی، مانند هنجار FL [12]، [11] و سن پیروزرسانی [13] (AOU) در نظر می‌گیرند. با انگیزه دریافت پیروزرسانی‌های مدل با افت فشرده‌سازی کمتر، برخی از کارها کیفیت لینک بیاسیم را در طریق زمانبندی در نظر می‌گیرند. [14]-[12]، [17] از سوی دیگر، برای در نظر گرفتن توزیع داده‌های غیر iid،

میانگین‌گیری فدرال (FedAvg) یکی از معرفت‌آورین و پایه‌ترین الگوریتم‌های [2] است که یک فرآیند تکراری دارد.

نویسنده‌گان از دیارتمان مهندسی برق (ISY)، دانشگاه لینشوبینگ، 58163 لینشوبینگ، سوئد (chung-hsuan.hu@liu.se, zheng.chen@liu.se, erik.g.larsson@liu.se) هستند.

این کار تا حدی توسط Zenith، Lund - Linköping در فناوری اطلاعات (ELLIIT) و بیان Knut و Alice Wallen-berg پشتیبانی شد.

عدم قطعیت توزیع داده‌ها در [19] در نظر گرفته شده است، و در [18] طراحی زمانبندی از اصل ولویت بالاتر به دستگاه‌هایی با تنوع بیشتر در داده‌های محل خود پیروی می‌کند. برخی از کارهای بهینه‌سازی مشترک زمانبندی دستگاه و تخصیص منابع را به سمت حداقل تأثیر [20] (با اتفاق تجربی [21]) در نظر می‌گیرند. با این وجود، همه آنها سیستم‌های FLM را در نظر می‌گیرند. تعداد کمی از کارهای موجود، طراحی را در یک محیط ناهمزنان در نظر گرفته‌اند. در [23] زمانبندی در FLM ناهمزنان بر اساس هدایت رساندن مجموعه مورد انتظار داده‌های آموزش با توجه به عدم قطعیت شرایط کانال، ورود داده‌ها و منابع ارتیاطی محدود در نظر گرفته شده است. با این حال، تأثیر توزیع داده‌های غیر iid در طراحی زمانبندی، در نظر گرفته نشده است.



شكل : افرآیند F و تبادل اطلاعات بین سرور و دستگاه‌های شرکت‌کننده.

در مقایسه با تنظیمات همزمان، FAL ناهمزمان باید با ناهمزمان بهاروزرسانی‌های مدل محلی مقابله کند، زیرا دستگاه‌های مختلف ممکن است آموش محلی را بر اساس ساخته‌های مختلف مدل جهان انجام دهند. برخی از طرح‌های تجمعی اکتشاف در مقالات بررسی شده‌اند. در [23] دستگاه‌هایی که از تذکرایهای گذشته با شکسته‌های انتقال پیشتری مواجه شده‌اند، بهاروزرسانی‌های گرادیان بزرگ‌شده را ارسال می‌کنند. در [24] وزن‌های پیشتری به ردیف‌های کنترل در فرآیند تجمعی داده می‌شود، زیرا ردیف‌های کنترل در مدل جهان مشارکت می‌کنند.

هر دو رویکرد با هدف بکسانسازی سهم مستگاههای مختلف انجام می‌شوند، اگرچه با داده‌های iid، ممکن است ممکن‌گابی کند شود زیرا بهبود رسانی‌های مدل برداست‌آمده از مدل‌های جهانی قدمی‌تر ممکن است حاوی اطلاعات غقید کمی برای نسخه فعلی باشد.

ما خلاصه‌ای از مهمترین دستاوردهای این اثر را ارائه می‌دهیم:
• ما یک چارچوب FL ناهمزمان با تجمع دوره‌ای پیشنهادی برآورده‌ها به همگرایی سریع دست می‌یابد و از بهروزی ترتیب‌دهنده‌ها فواید می‌برد.

ما بک سیاست زمان‌بندی پیشنهاد می‌کنیم که به طور مشترک کیفیت کانال و توزیع داده‌های آموزش را در
نظر بگیرد. با هدف کاهش واریانس و یا باس پهلوزرسانی‌های مدل تجمعی. اثربخشی روش پیشنهادی هم
توسط تحلیل همگابی نظری و هم تابع شیوه‌سازی پیشنهادی می‌شود.

ما یک طرح وزن‌دهی آگاه از سن برای تجمعی مدل پیشنهاد می‌کنیم تا اثرات ناهمزنمانی به روزرسانی را کاهش دهیم. **ما تأثیر فشرده‌سازی به روزرسانی، ناهمگونی داده‌ها و ناهمزنمانی درون تکرار از تحلیل همگرایی بر جسته می‌کنیم.** که همچنین یک انگیزه نظری در باره طراحی، زمان‌ابتدی، ما فراهم می‌کنند.

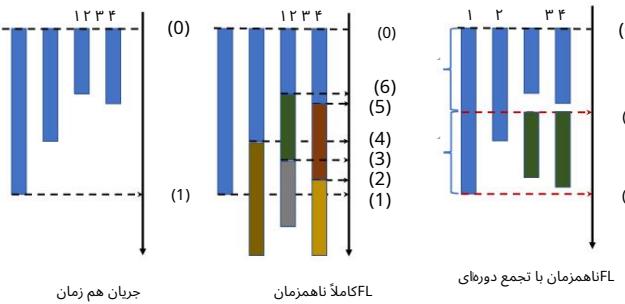
دوم. مدل سیستم

ما یک سیستم $F\cup L$ با دستگاه‌هایی که در آموزش یک مدل یادگیری جهانی مشترک شرکت می‌کنند، که توسط بردار پارامتری یک بعدی R پارامتری شده است را در نظر می‌گیریم. $\{1, \dots, n\}$ نرها عنوان مجموعه دستگاه در سیستم نشان می‌دهیم، هر دستگاه N مجموعه‌ای از داده‌های آموزش محلی $\{S_i\}_{i=1}^n$ را در خود جای داده است، فرض کنید $NS = \{N_1, \dots, N_n\}$ نشان‌دهنده کا، مجموعه داده‌ها در سیستم با اندازه $|S|$ باشد، که در آن، $S = S_1 \cup \dots \cup S_n$.

دستگاه N را تنظیم کرد.

هزار کتابخانه زبان اصلی اعماقی اشپلیو (SG) هدف سیستم پاکن و بدار امنیتی است که تابع زیان تجربی تعریف شده توسط را به حداقل برساند.

- ایشان دهنده شاخص تکرار محلی، (\cdot) انشان دهنده نرخ $\frac{1}{n}$ بوده است.



قطعات و م محل منظم در دستگاه در تکرار سراسری-ام و تکرار محل-ام به صورت زیر تعریف عدم تعادل بالقوه مدل ناشی از آموزش ناهمزنان.

$$(\cdot) = ((\cdot)) + \nexists \quad (\cdot) \sqsupseteq (\cdot \mid \emptyset)$$

که در آن ۰ ضریب منظم‌سازی است. برای هر دستگاه، Δ به روزرسانی محل آن به صورت زیر محاسبه می‌شود

$$+ 1) = (\cdot, \cdot) \square (\cdot, \cdot) \square (\cdot, \cdot) \quad (6)$$

پلیده‌اند..، 0 = و (،) که روی یک مینی-دسته تصادفی انتخاب شده S (B ارزیابی

نیازمندی: جهاد اسلامی همچنانی با همین نیازمندی، FL کاملاً ناهمزمان و FL ناهمزمان، اولین تکرار محلی به صورت زیر است:

$(\cdot, 0) = (\emptyset)$,

كجا

$$\{ \mid \square K(\) \} + 1 \text{ حداکثر} (\) =$$

(((B,))) اشنان دهنده گردایان محاسبه شده بر اساس یک مینی-دسته
اصدافي انتخاب شده S ((B,)) است.
مس از تکمیل آموزش محلی، هر دستگاه به روزرسانی مدل را به صورت تفاوت
بین بردار پارامتر مدل قبل و بعد از آموزش دریافت می‌کند، یعنی

$$() = (,) \sqcup (, 0) \quad (4)$$

(۳) پس از آموزش محلی، زیرمجموعه‌ای از دستگاه‌های N (برای آپلود اطلاعات رسانی) مدل خود به سرور برنامه‌اریزی شده‌اند.

اطلاعات دریافتی را تجمیع کرده و مدل سراسری را مطابق با آن بهاروزرسانی می‌کند.

$$(+1) = \square$$

آخرین تکرار سراسری را نشان می‌داده که در آن دستگاه یک مدل سراسری بهاروزرسانی‌اشده دریافت کرده است. پس از دریافت بهاروزرسانی‌ها از همه دستگاه‌های موجود در Δ سور تجمعیم مدل را به صورت زیر انجام می‌دهد:

که در آن (در ۴) تعریف شده است، و (انسان دهنده وزن ۱) = (است. سپس مدل بهاروزرسان شوچوچوچوپری) کلشودمه هدوای الهای ادله قوه مصیرای الاماعن آمده [۲۶] (همچو رستگان ها) محل (ALU) را به صورت زیر تعریف می کنیم:

$$() = \square (), \quad (7)$$

که نشان دهنده زمان سپری شده از آخرین دریافت یک مدل سراسری به وزارت اسناد است. وزن (ام) اتواند به انفراد مبلغ اتفاقی شیوه $\text{۱۶} \times \text{۰}.\text{۷} \times \text{۰}.\text{۲}$ (هزار و پانصد) مخاطب معمولی-بلاتکلی (با توجه اینجا طبقه ایمودیوکا، بلطفه ایقحطامیازه) (اکتفال).^{*} نشان داده ام. از طریق کانال های بسیم منتقل خواهد شد. طرح فشرده سازی در زیربخش بعدی به تفصیل نشان داده خواهد شد.

ب. Fناهمزمان با تجمع دورهای

ای پرداختن به مشکل پراکنده‌ی در $\text{FL}_{\text{همگام}}^{\text{بعد}}$ بدن ایجاد بار ارتاطی بین از حد، ما بک چارجوب هم زمان با تجمع دوره‌های پیشنهاد می‌کنیم. ایده کل این است که آموزش ناهمازمان در دستگاه‌های مختلف مکان‌باذیر شود، به طوری که صورت دوره‌ای پلاروزرسانی‌ها را از دستگاه‌هایی که محاسبات خود را تکمیل می‌دانند مجاوز نکند، در حالی که یقین آموزش محل خود را بدون وقفه با حذف ادامه می‌دهد. شکل ۷نمودهای جدول زمانی آموزش و پلاروزرسانی $\text{FL}_{\text{همگام}}^{\text{اصل}}$ ، $\text{FL}_{\text{کامل‌ناهامازمان}}^{\text{[5]}}$ و طرح پیشنهادی ما را نشان می‌لاده. از آنجایی که دستگاه‌های مختلف قابلیت‌های محاسباتی متفاوتند، هر زمان که یک دستگاه آموزش می‌لاده، باید این را برای گزارش پلاروزرسانی نشان بخلي خود را به بابان برساند، سیگالی به سورور ارسال می‌کند که آمادگی این را برای گزارش پلاروزرسانی نشان می‌لاده. در هر صد زمان، سورور زیرمجموعه‌ای از دستگاه‌های آماده پلاروزرسانی را برنامه‌ریزی می‌کند. ما (K) را عنوان مجموعه‌ای از دستگاه‌های آماده پلاروزرسانی در تکرار سراسری -ام تعریف می‌کنیم. فرض کنید (K_i) (I) مجموعه دستگاه‌های برنامه‌ریزی شده باشد، با $\min\{|K_i|\} = |K|$ که در آن داکتر تعداد دستگاه‌هایی است

ج. مدل لایه فیزیکی (PHY)

در یک فرآیند FL، پخش ارتباط در دو مرحله انجام می‌شود، سرور مدل سراسری را به دستگاه‌ها ارسال می‌کند (انتقال downlink) و دستگاه‌ها بازروزسانی‌های مدل را به سرور گزارش می‌دهند (انتقال uplink). سرور قدرت انتقال بسیار بالاتری نسبت به دستگاه‌ها دارد و انتقال downlink به دلیل کانال پخش نیازی به تقسیم منابع ارتباطی ندارد. بنابراین، فرض می‌کنیم هیچ خطای فشرده‌سازی در downlink وجود ندارد. با این حال، انتقال uplink از گلوبال‌گاه ارتباطی رنج می‌پردازد، که باعث می‌شود طراحی زمان‌بندی و تخصیص منابع از اهمیت ویژه‌ای، برخوردار باشد.

این مرحله از زمان بندی آپلینک به ویژه برای لایزر طریق بی اسیم اهمیت دارد. شبکه ها به عنوان منابع ارتباطی باید بین دستگاه ها به اشتراک گذاشته شوند.

2 توجه داشته باشد که این تعریف مبتنی بر سن با تعریف سن بهاروزسانی (A05) که در [13] ارائه شده است، متفاوت است، که زمان سپری شده در هر دستگاه را از آخرین مشارکت آن در تجمعی مدل اندازه‌گیری می‌کند.

اما در نظر می‌گیریم که در هر تکار سراسری، مدل بهاروزرسانی می‌اشود.
از همه دستگاه‌ها از طریق یک کانال محوشگی بلوک با بلوک همدوس نماد θ منتقل
می‌شوند. ما معمعاد فرض می‌کنیم
(۱) اصلیهای مبالغه‌بینی دستگاه‌ها، یعنی منحصراً به دستگاه اختصاص داده شده و = ()
اپن () = ۱

برنامه‌ها برای این اتفاذه ایشوند
آن دیس‌های (۷) جمله دوم $3^{\log_2 n}$ برای نرم برداری هستند
مقدار، عبارت سوم به این معنی است که $1 + (1 + 1)^{\log_2 n}$ بایت هستند
برای انتقال هر عنصر غیر صفر از (*) مورد نیاز است.

بنابراین، انتقال از چندین دستگاه بدون تداخل، فرض کنید و (۸) نشان دهنده مقیاس بزرگ و محوشده‌گی کانال در مقیاس کوچک از دستگاه $-th$ به ترتیب سور. با فرض اینکه دستگاه $-th$ -ارسال را انجام داده است حداکثر توان \square و نویز افزایش در کانال دنبال ملاحظه شود. کانال لینک ام برابر است با

$$(\text{log}_2 1 + \frac{(\text{log}_2 h)(\text{log}_2 12)}{2})$$

د. انگیزه طراحی زمانبندی و تجمعیع
تحت تنظیمات FNL ناهمزن پیشنهادی، طراحی کلید
سوالات عبارتند از:
(ابا) توجه به محدودیتهای منابع ارتباطی، چگونه
آیا باید زیرمجموعه‌ای از (KA) را برای تجمعیع مدل تحت سناریوی داده‌های آموزشی
ناهمگن برنامه‌اریزی کنیم؟

اما در نظر داریم که دستگاه‌ها فشرده‌سازی داده‌ها و طرح کدگذاری کانال مناسب را مطابق با کانال اعمال می‌کنند.

ظرفیت در (8) به گونه‌ای که انتقال‌های مدل به بازرسانی شوند بدون خطأ هستند. در نتیجه، سرور می‌تواند به طور قابل اعتمادی دریافت کند تا خود را بین‌المللی دستگاه.

۱۰) (۱) و طرح فشرده‌سازی داده‌ها در زیر مشخص شده‌اند.

(اتخیص منابع ارتباطن: برای دستیابی به سطح یکسانی از افت فشرده‌سازی در بازرسانی‌های مدل از هر کدام دستگاه، ما منابع نماد را به گونه‌ای تخصیص می‌دهیم که

(۱۲) دستگاه‌های مختلف ممکن است واحدهای محاسبه و منطق (ALU) متفاوتی داشته باشند، یا جدیدتر یا قدیمی‌تر، چگونه می‌توان یک سیاست وزن‌دهی مناسب با در نظر گرفتن تازگی طراحی کرد؟ از بهاروزسازی‌های مدل؟

به یاد داشته باشید که هدف، به حداقل رساندن زیان تجربی جهانی است.

$\tilde{\mathbf{x}}^*$ که در (2) داده شده است. ما مدل بھینه را به صورت $\text{argmin}(\mathbf{x})$ تعریف می‌کنیم.

با فرض اینکه \mathbf{x} شدت محدب است، به این معنی که $\mathbf{x} = \mathbf{x}^*$ و $\mathbf{x} - \mathbf{x}^*$ smooth است، یعنی برای هر $R \geq 0$ ،

که نشان می‌دهد دستگاه‌هایی با کانال‌های بهتر با نمادهای کمتر اختصاص داده‌اند. طراحی مشابه در [11] در نظر گرفته شده است.

(۳) پراکندگی و کمیت‌سازی: هر دستگاه نیاز دارد فضیده‌هایی (با توجه به بودجه (۱) بیت‌ها).

ما از پراکندگی تصادفی و به دنبال آن دقت پایین استفاده می‌کنیم. کوانتیزاسیون تصادفی برای فشرده‌سازی مدل. به طور خاص، ما پراکندگی و انتخاب مدل پراکنده تعريف کنید R (۲) ~

جزء ۱ = ۰، (۱) به صورت زیر تعريف می‌شود

که در آن [.] نشان دهنده امید به کل تصادفی بودن است
در تکرارهای گذشته، 1، توجه داشته باشید که (12) کمیت را تعیین می‌کند
اثربخشی آموزش مدل با بررسی شکاف بین
E | E [(+1)] و یک کران بالایی کوچکتر، به طور بالقوه شکاف را کاهش می‌دهد. به ویژه، in (+1) | E

۱۱۲

(12) را می‌توان به عنوان واریانس مدل تجمعیانشده تفسیر کرد.
تولید شده از بهاروزرسانی‌های مدل دستگاه‌های برنامه‌بریزی‌اشده.
ما برخی از عوامل سیستمی که بر آن تأثیر می‌گذارند را به شرح زیر فهرست می‌کنیم.
(اتوزیع داده‌های آموزش: اگر همه دستگاه‌ها idداشته باشند
با توجه به داده‌های آموزشی، انتظار داریم که 0 باشد، همانطور که در بخش دوم بحث شد.
مستقیماً تابعه می‌گرد که

$\text{س}(\cdot) = \sim(\text{علامت}(\cdot)) = \sim(\cdot) \mid 2$

در دستگاههای مختلف زیاد است، که منجر به واریانس زیادی می‌اشود.
از این تجربه می‌توان برای آزمون H_0 از میانگین μ استفاده کرد.

و کوانتیزاسیون است. ۲) | | متفاوت تصادفی است. ۳) بردار فشرده حاصل ($\sim Q \cdot V$) نتیجه ایجاد ترکیب \square , $((\square))$ است. سطح $E[\square]$ کوانتیز بدو پایان است، یعنی، $= E[\square]$ ننتیس ایمن تصادفی است. ۴) دارای کمیت‌های $| |$ است.

۵) در این فرآیند و استسه به مقدم $\square \square \square$ متفاوت تصادفی کوانتیزاسیون ثابت، مازکیم (\square) را پیدا می‌کنیم.

(۳) پهلویان از مجموعه این اهداف، این است که با توجه به این اهداف، می‌توان از این اهداف برای ایجاد مدلی برای پیش‌بینی این اهداف استفاده کرد.

(۴) هم‌زمان با انتقال بهاروزرسانی‌های مدل از طریق کانال‌های پر اسیم با نزد محدود، سورو فقط دریافت می‌کند اطلاعات نویزی به دلیل فشرده‌سازی داده‌ها. تلفات فشرده‌سازی بیشتر، واریانس بیشتری از مدل تجمعی‌شده ایجاد می‌کند.

(۵) پهلویان از مجموعه این اهداف، این است که با توجه به این اهداف، می‌توان از این اهداف برای ایجاد مدلی برای پیش‌بینی این اهداف استفاده کرد.

$$+ 32 + () \log_2 (+ 1) + 1 \square \left\{ \right\} \quad (10)$$

کوچکتر - (1+) || E با نظر گرفتن موارد زیر م جنبه‌ها: دستگاه‌های برنامه‌ریزی اشده باید پک ساختار هم زمانی کار کنند و در اینجا آنها را می‌نماییم.

• اتلاف فیشره‌سازی باید پایین نگه داشته شود، که این امر ما را ترجیب می‌کند تا دستگاه‌های با شرایط کanal بهتر را در اولویت قرار دهیم.
از منظر تجمعی مدل، می‌توانیم با در نظر گرفتن طراحی وزن‌دهی آگاه از سن در فرآیند تجمعی، تأثیر نامطلوب بهاروزرسانی‌های ناهمزمان را کاهش دهیم.

III. طراحی زمان‌بندی و تجمعی برای FL ناهمزمان

به طور خلاصه، طراحی زمان‌بندی مبتنی بر اهمیت داده و آگاه از کanal با سیاست تجمعی ملکی سیاست زمان‌بندی پیشنهاد می‌کنیم که با در نظر گرفتن توزیع داده‌های آموخته و شرایط کanal دستگاه‌های آماده بهاروزرسانی، به دنبال دستیابی به شکاف بهینگی کوچک‌تری است.⁴ می‌بینیم، وزن‌های تجمعی بر این اساس تنظیم می‌شوند تا تأثیر مضر آموخته ناهمزمان کاهش دهد.

مشکل اعمال بهاروزرسانی‌های قدیمی بر روی مدل تکامل‌یافته، از سوی دیگر، ترجیح بهاروزرسانی‌های محلی جدیدتر به مدل کمک می‌کند تا با گذشت زمان سریع‌تر و روان‌تر همگرا شود، هرچند که این خطر وجود دارد که به یک مدل نامتعادل که به سمت دستگاه‌های با قدرت محاسباتی بتر متمایل است، همگرا شود. از آنجایی که سیاست زمان‌بندی پیشنهادی به گونه‌ای انجام می‌شود که همگن توزیع داده‌ها را بهبود می‌بخشد، مزیت استراتژی «ترجیح مدل‌های جدیدتر» قانع‌کننده‌تر می‌شود.

V. تحلیل همگرای

ما بجز از نمادگزاری‌ها و تعاریف را برای تبدیل معرفی می‌کنیم.
تحلیل ماهیت سیستم پیشنهادی.

تعریف .(اتجمعی ناهمزمان):

*فرض کید (عدد نسخه‌های مختلف مدل سراسری دریافتی در تجمعی مدل -am باشد، یعنی تعداد عناصر منحصر به فرد در مجموعه $\{1, 2, \dots, n\}$ فرض کنید) $M = \{1, 2, \dots, n\}$
یک زیرمجموعه دستگاه با $M' = \{1, 2, \dots, m\}$ باشد که ALU میکسانی دارد، یعنی $M \setminus M'$

الف. زمان‌بندی مبتنی بر اهمیت داده‌ها با آگاهی از کanal

برای روش نشدن این ایده، یک مسئله طبقه‌بندی با داده‌های برچسب‌گذاری شده را در نظر می‌گیریم، مجموعه داده‌های آموخته با $(L, |L|)$ نمایش داده، که در آن $\{1, 2, \dots, n\}$ یک مجموعه متناهی است که

نقطه‌نامه‌های سیستمی می‌باشد و با $L = \{\text{آگاهی} \cup \text{نامه} \cup \text{نامه}\}$ تعریف می‌شوند که آگاهی از کanal توزیع داده‌های ناهمزمان را بهبود می‌بخشد.

،
،
، در س

من توانیم (Π) را برای دستیابی به حداقل واریانس برچسب انتخاب کنیم.

$$\Omega(\Pi) = \min_{=1}^n \Pi(\cdot)$$

که نیاز به اعلویت این زمان‌بندی داشت، اما با توجه به محدودیت کمالی (Π) از آگاهی تغییر نمی‌نماید (مشخصه‌هایی که نیاز به اعلویت این زمان‌بندی داشت، اما با توجه به محدودیت کمالی (Π) از آگاهی تغییر نمی‌نماید) $\Pi(\cdot) = \arg\min_{=1}^n \Omega(\cdot)$ است. سپس، $\Pi(\cdot)$ را بدأ می‌کنیم.

ب. تجمعی مدل آگاه از سن برای مقابله با ناهمزمان در تجمعی مدل، وزن‌ها را نه تنها بر اساس نسبت داده‌ها مانند، (5) بلکه بر اساس واحد محاسبه و منطق آن نیز تعیین می‌کنیم. طراحی وزن‌دهی آگاه از سن به شرح زیر است:

$$(13) \quad \Pi(\cdot) = \frac{(\cdot)}{\sum_{=1}^n \Pi(\cdot)}$$

در اینجا، یک ثابت با مقدار حقیقی است و انتخاب مقدار آن را می‌توان به سه حالت تقسیم کرد:

۱. > که از این تعداد، سیستم بهاروزرسانی‌های محلی قدمی‌تر را ترجیح می‌دهد. $1 = \cdot$ که از این تعداد، سیستم بهاروزرسانی‌های محلی جدیدتر را ترجیح می‌دهد. $1 = \cdot$ که معادل طرح پایه در (5) است.

ترجیح بهاروزرسانی‌های قدیمی‌تر می‌تواند به طور بالقوه فراوانی مشارکت را در بین دستگاه‌ها متعادل کند و خطر سوگیری آموخته مدل به سمت دستگاه‌های با قابلیت محاسباتی قوی‌تر را کاهش دهد. این طراحی زمانی که توزیع داده‌ها بسیار غیرمستقل باشد و برخی از دستگاه‌های با قابلیت محاسباتی پایین‌تر، داده‌های آموخته منحصر به فردی داشته باشند، عملکرد خوبی دارد. با این حال، این امر همچنین باعث ایجاد ...

طرح زمان‌بندی ما نه تنها در محیط یادگیری مашینی ناهمزمان مورد نظر، بلکه برای هر محیط یادگیری مانشینی توزیع شده با داده‌های ناهمگن و نامتعادل نیز قابل اجرا است.

همانطور که در سیستم پیشنهادی، زمان‌بندی دستگاه منوط به مجموعه آماده بهاروزرسانی (K) به جای مجموعه کامل دستگاه N است، معیارهای زیر را برای تعیین کمیت سطح ناهمگونی داده‌ها برای هر زیرمجموعه دستگاه M معرفی می‌کنیم.

تعريف .(سطح ناهمگنی داده‌ها): برای یک زیرمجموعه دستگاه M و تعریف می‌کنیم $\Gamma(M) = \min_{=1}^n \Pi(\cdot)$

$$\Gamma_1(M) = \max_{=1}^n \Pi(\cdot)$$

$$\Gamma_2(M) = \left(\max_{=1}^n \Pi(\cdot) \right)^{-1}$$

الف. فرضیات

برای تسهیل تحلیل، فرضیات پایه زیر را در نظر می‌گیریم.

فرض .1.(همواری): تابع زیان محلی $\Pi(\cdot)$ همانطور که در (11) مشخص شده است، هموار است که معادل آن است
، به شرط پیوسته لیپشیز، $\Gamma_1(M) = \Gamma_2(M)$ $\forall M \subseteq N$

فرض .2.(تحدب قوی): تابع زیان محلی $\Pi(\cdot)$ به شدت محدب است، یعنی R

$$\Gamma_1(M) \leq \Gamma_2(M) \leq \Gamma_1(M)$$

و به دنبال آن فرضیات برای ارزیابی گرادیان تصادفی،

... ، فرض .3.(محدودیت گشتاور دوم): $\Gamma_1(M) \leq \Gamma_2(M) \leq \Gamma_1(M)$ از این راضی شود

$$(14) \quad \Gamma_1(M) \leq \Gamma_2(M) \leq \Gamma_1(M)$$

با $= 1$ ساده می‌شوند. دستگاه‌ها در هر $(M, \Gamma_1(M))$ به حالت $\Gamma_1(M)$ دارای ALU میکسانی هستند، بنابراین

برای برخی ثابت‌های مثبت او

ثبتات. به بخش A VIII مراجعه کنید.

توجه داشته باشید که به نظر مارسد در ادبیات رایج است که فرض شود گردیدن کراندار یکنواخت، یعنی $\sigma = 2$ در (14) و یک گردیدن قویاً $\sigma = 1$ در (15) . این فرضیات مثلاً $[11]$, $[6]$, $[5]$, $[29]$, $[24]$, $[22]$, $[21]$, $[14]$, $[13]$ این حال، این فرضیات اساساً ناسازگار است زیرا هیچ محدب قوی وجود ندارد. لافعی که گردیدن آن به طور یکنواخت روی R کران دارد همچنین در $[30]$ اشاره شده است. کار با یک محدب قوی تابع و گردیدن اهای کراندار یکنواخت نیاز به یک کراندار دارند. فضای پارامتری، که منجر به یک مسئله بهینه‌سازی (محدد) متفاوت می‌اشود، که یک الگوریتم حل برای آن باید داشته باشد.

برای گنجاندن یک مرحله پیش‌بینی، بنابراین فرض منکریم که «نویز» در گردیدن‌ها به طور یکنواخت محدود است (طبق انتظار)، همچنین در $[28]$ اتخاذ شده است، که به ما امکان می‌دادهد با [یک] قوی کار کنیم. شرایط محدب در یک فضای نامحدود.

نکته: ابا تجمعیـ دوره‌ای، ALU هر دستگاهی با زمان آموزش محدود شده توسط $\max(0, \text{برآورده می‌کند})$ و $\max(\text{سپس، از تعریف اداره} \cdot \text{سیس})$.

$$(\) \square - \frac{\text{حداکثر}}{\text{حداقل}} + 1 . \quad (15)$$

نکته ۱۲. اگر فرض کنید Z مجموعه‌ای از N باشد. از آنجایی که N یک مجموعه متناهی است، اعداد او $\geq 2^N$ متفق وجود دارد به طوری که

| Г1 (M) | (16)

| Г2 (M) | (1 7)

یعنی، معیارهای سطح ناهمگنی به طور یکنواخت محدود هستند.

علاوه بر این، برای عدد رزرو شده، (فرض می‌کنیم که یک حداقل حداقل ۰ وجود دارد به طوری که

$$\mathbb{E}[(\cdot)^2] \leq \text{دقيقه}, \quad (18)$$

که در آن، انتظار نسبت به تصادفی بودن در کانال های بن سیم

در قضیه زیر، همگرایی اصلی را ارائه می‌ادهیم
نتیجه مدل سیستم ما در یک مورد خاص با یک مورد محلی
کنکار در هد دو، ارتباط، یعنی .. = ۱.

قضیه اتحت فرضیات ۱-۳-۱۸) اعادلات = نیز یادگیری = (۰)،

دقيقة

و. حطلاکشل} ۱۴ و دستگاه حزئی

مشارکت به طوری که $\Pi(\cdot) = \square = 1, \dots, N$ می‌باشد، این امر ممکن است باشد:

$$\mathbb{E} \left[\cdot \left(\cdot + 1 \right) \right] \square$$

$$\begin{aligned} & \square \quad \text{_____} \quad \\ & (.) + \frac{(+1)}{\cancel{(+1)}} \mid \mid (1) \boxed{+} \quad \frac{+}{\cancel{+}} \\ & = 2(1+2), 3 = 2(1+2) \text{ چنانچه} \\ & (.) = \frac{2}{\cancel{2}} + \frac{3}{\cancel{2}} + \frac{2}{\cancel{2}} + \frac{3}{\cancel{2}}. \end{aligned}$$

نکته: ۳۰ تأثیر تکرارهای محلی متعدد، یعنی $1 > F$ در مورد تحلیل همگرایی اقباً اثبات شده است و در متون موجود در مورد $F \geq 1$ مورد استفاده قرار گرفت. [31] [32]. تمرکز تحلیل ما بر تأثیر ... است. فشیده‌سازی بهاروزسانی، ناهمنگوی داده‌های آمو سیستم، ما معتقد‌دید که تعیین نتایج ما به حالتن که $1 > F$ وجود دارد، اما با دنبال کردن رویکردهایی از، به عنوان مثال، [31] با این حال، این امر، از جمله موارد دیگر، مستلزم واگرایی مدل محلی. 6 برابر واضح و سادگی تحلیل ما حالتی را در نظر می‌گیریم که $1 = F$ باشد.

فشرده‌سازی پاروزسانی، ناهمگونی داده‌های آموزشی و ناهمزمانی درون تکرار بر همگرایی سیستم، ما معتقدیم که تعمیم نتایج ما به حالتی که ۱ وجود دارد، امکان‌پذیر است. با ذنبال کردن رویکردهای از، به عنوان مثال، [31]، [29]؛ یا [32] با این حال، این امر، از جمله موارد دیگر، مستلزم کمی‌اسازی است و اگرایی مدل محلی. ۶ برایوضوح و سادگی تحلیلی، ما حالتی را در نظر می‌گیریم که ۱ = باشد.

$$E[(\cdot + 1)] \leq \frac{e}{c}$$

چه زمانی، که در آن و در (15) تعریف شده‌اند، و در قضیه، این ثابت را می‌توان به عنوان یک بهینگی در نظر گرفت شکاف، که با افزایش در داده‌ها چگونه افزایش می‌یابد سatarویی که در آن = او بنا بر این . پیاشد، سیستم = حق با وجود ناهمزمانی درون تکرار، به یعنی همگرا می‌اشود، یعنی ، و فضیله‌های اسازی را بهاروزرسانی می‌کند، یعنی > با این حال، در یک سatarویی غیر like که در آن ، سطح بالاتری از داده‌ها وجود دارد ناهمگی منجر به شکاف بهینگی بزرگتری می‌اشود. وجود دلالت بر این دارد که ناهمزمانی درون تکرار، به یعنی بودن را ایجاد می‌کند شکاف بزرگتر، علاوه بر این، حضور نشان دهنده آن است که برنامه ریزی دستگاه‌هایی با کیفیت کاتالیز بهتر، عملکرد را بهبود می‌بخشدند، او قدر کمی خداوند شیوه‌گرد و ممتاز است (18) شده کمتری را مجاز می‌کند. تعریف شده است.

نکته ۵. جدا از رفتار مجانی، سطح فشرده‌سازی به روزرسانی‌های مدل نیز بر هر تکرار تأثیر می‌گذارد.

عملکرد. با کیفیت لینک بهتر، پراکنده‌گی من تواند
کمتر تهاجمی باشد، و/یا سطح کوانتیزاسیون بالاتر
در نتیجه، عبارت (۴) امن تواند کوچکتر باشد و بنابراین
شکاف بهینگی در هر تکرار کاهش می‌یابد. علاوه بر این
(۴) با نامه‌گوئی داده‌های کمتر رشد می‌کند و همچنان
عملکرد در هر تکرار.

این دو نکته، اهمیت توجه به این موضوع را تأیید می‌کنند.
هم توزیع داده‌های آموزشی و هم کیفیت کانال در ما
طراحی زمان‌بندی دستگاه‌ها

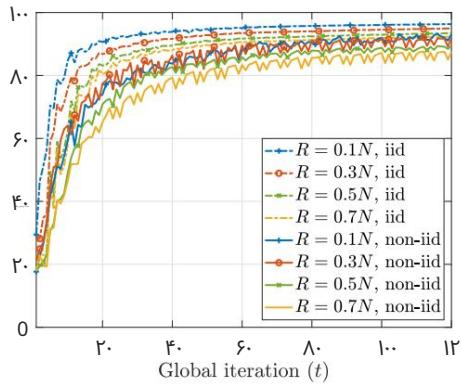
نتایج شبیه‌سازی

ما شبیه‌سازی‌ها را با استفاده از مجموعه داده‌های [33] MNIST حل مسئله طبقه‌بندی ارقام دست‌نویس با استفاده از یک شبکه عصبی کانولوشن با بعد مدل =

۲۱۸۴۰ کانال محو شدن بلوک، نمادهای را برای آپلینک پوشش می‌دهد. انتقال‌ها، ما محوشگر بیلی (CN 0, 1) hرا در نظر می‌گیریم. و کنترل توان آپلینک به گونه‌ای که نسبت سیگنال به نویز دریافتی نسبت ۳۰dB ایل است. پارامترهای سیستم به شرح زیر تعیین شده‌اند.

- (توزيع داده‌های آموختن) $= 60000 | \text{نمونه وجود دارد}$

انتظار بر تصادفی بودن تصادفی بودن گرفته می‌شود. معنی خواهد بود که اگر (X_1, X_2, \dots, X_n) گردیدن یک سری از نتایج پیش‌بازی مدل، و دستگاه محدود به گردیدن 9 کالا^۹ و فشرده‌بازی مدل^{۱۰} باشد، نهانه نمایند، تهاجم تکرارهای گذشتۀ شمر.

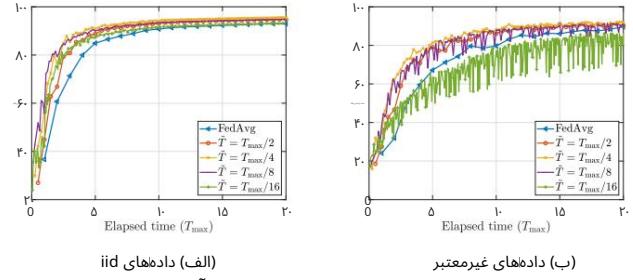


شکل 3: تست دقت طرح پیشنهادی تحت شرایط مختلف نسبت زمانبندی جزئی، که در آن $\alpha = 0.05$ است.

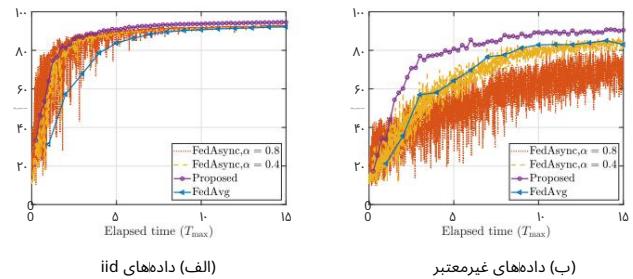
در حالت $\alpha = 0.5$ ، iid نمونه‌ها به طور تصادفی به هر کدام اختصاص داده می‌شوند. دستگاه بدون جایگزین، برای مورد غیر iid، تخصیص داده از تنظیمات [2] اپریوی می‌کند، که در آن هر دستگاه شامل حداقل 10٪ (200) رقم مختلف است. «قابلیت محاسبات ناهمزن» ما به عنوان موتان خطا قابل امتحان مخلقاً دستگاه، تولید شده از توزیع تکفرم U (حداقل، حداکثر)، که

حداقل زمان ممکن برای آموزش دستگاه. «فشنده‌سازی» از کوانتایز تصادفی 4 اسطحی، Q4 استفاده شده است. «نخ یادگیری نزولی» $= 0.01$ در ابتدا است. اتخاذ شده، همراه با ضریب منظم‌سازی $= 0.02$. از آنجایی که متابع ارتباطی جداکننده اشتراک گذاشته می‌شوند دستگاه‌ها، دستگاه‌های برنامه‌ریزی شده‌ای بیشتر به معنای ارتباط کمتر است متابع به ازای هر دستگاه، که منجر به انتلاف فشنده‌سازی بالاتر می‌شود. نتایج دقت آزمایش با انتخاب‌های مختلف در نشان داده شده است شکل 3. اهمانطور که می‌بینیم، در سفاریو، iid نمونه‌گذاری ترجیح داده می‌شود زیرا بهاروزرسانی‌های مدل دریافتی را با کمیتی بهتری ارائه می‌دهد دقت به عنوان نتیجه بیت‌های اختصاص داده شده بیشتر به هر کاربر. روی از سوی دیگر، در سفاریو غیر iid، اینکه بدستان وجود دارد بین انتلاف فشنده‌سازی و پایاس مدل، که باعث می‌شود انتخاب مهم. در این کار، ما بر تأثیر تمرکز می‌گذیم از طراحی زمانبندی برای یک مقدار بهینه ثابت به پارامترهای سیستم زیادی بستگی دارد؛ بهینه‌سازی آن می‌تواند در کارهای آینده مورد مطالعه قرار گیرد.

از نظر تحلیلی بسیار غیر بدینه است زیرا یافتن مقدار بهینه ما قادر یک عبارت قابل حل هستیم که رابطه بین شکاف بهینگ و مقدار را مشخص کند از. برای نتایج عددی گزارش شده در اینجا، ما انتخاب می‌گذیم $\alpha = max/4$ زیرا این بهترین نتایج را در شبیه‌سازی‌ها ارائه می‌دهد. برای مقایسه بین آنها مختلف به زیر مراجعه کنید).



شکل 4: تأثیر بر دقت آزمایش FL ناهمزن-زمانی پیشنهادی با تجمع دوره‌ای در iid و غیر iid سناریوهای که در آنها $\alpha = 0.2$ و $\beta = 0.3$ است.



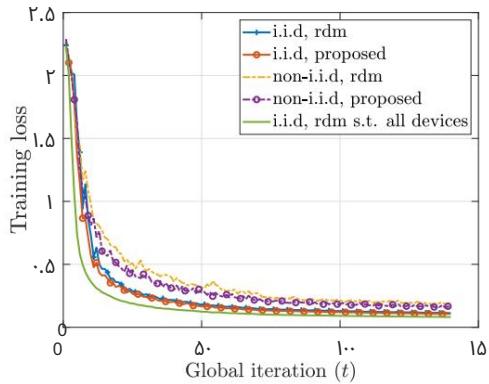
شکل 5: دقت تست FedAvg، الگوریتم ناهمزن پیشنهادی با تجمع دوره‌ای، و FedAsync [4] در iid سناریوهای غیر iid، که در آن $\alpha = 0.2$ و $\beta = 0.3$ است.

مقایسه‌ای منصفانه بین طرح‌های مختلف انجام دهید، ما مشکل را برطرف می‌گذیم میانگین میزان متابع ارتباطی اختصاص داده شده یک‌دوقوله‌ای را می‌گذیم که آنها در دور از این طبقه هستند و تقاضه $\alpha = max/4$ بذیرفته شود، آنکه آنها در دور در دسترس خواهد بود برای FedAsync اگر FA را با گذشت زمان اجرا کند مدت زمان تعداد نمادهای موجود در دور FA باشد / ۰ (حداکثر) در این حال، این نتیجه گیری را نمی‌توان تعمیم داد زیرا نتیجه همانطور که در شکل 4 نشان داده شده است، با این حال، بهترین عملکرد را ارائه می‌دهد.

همانطور که در شکل 4 نشان داده شده است، با این حال، این نتیجه گیری را نمی‌توان تعمیم داد زیرا نتیجه به بسیاری از پارامترهای دیگر سیستم پستگی دارد.

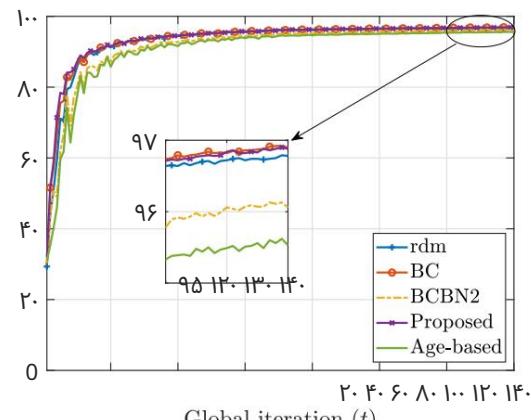
مقایسه بین طرح پیشنهادی ما، FedAvg، FedAsync و FedAvg7 با مقادیر $\alpha = 0.4$ و $\alpha = 0.8$ در شکل 5 نشان داده شده است. شکل 5 مشاهده می‌گذیم که اگرچه FedAsync می‌تواند به کاهش اثر سرگردان، نتیجه دقت آزمایش قوی نشان داده شده است. نوسان، به خصوص با داده‌های آموزش غیر iid، برای هر دو سناریوهای iid و non-iid، FL ناهمزن پیشنهادی ما FedAvg و FedAsync7 از FedAvg بهتر عمل می‌کند، که نشان می‌دهد اثربخشی در از بین اثر سرگردان و دستیابی به FL.

اولاً، در شکل 4، عملکرد روش پیشنهادی خود را ارائه می‌دهیم. طراحی با زمانبندی تصادفی و تجمعی پایه‌گذاری تحلیل همگرایی ۱ در (13) که با مقادیر مختلف $\alpha = max/4$ به دست آمده است، تا ما مقایسه تلفات آموزش را در شکل 5 برای اعتبارسنجی ارائه می‌کنیم. نتیجه دیدیم که جمله خوب در آینه می‌باشد، لذا موارد زیر را مشاهده می‌کنیم عملکرد همگرایی، عملکرد همزنمانی، موارد زیر: نیز برای مقایسه ارائه شده است. سپس، در شکل 5، مقایسه می‌گذاریم که FedAvg و FedAsync7 طور خاص نشان دهنده روش‌های دستگاه اعمال می‌شود + FL (نمادهای پیشنهادی، و FedAsync [4]).

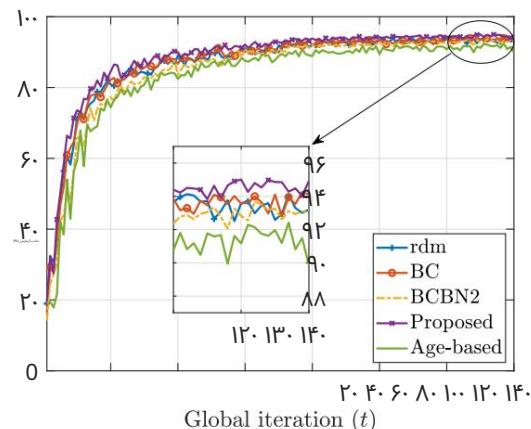


شکل ۶: نتایج آموزش برای روش‌های تصادفی و زمان‌بندی پیشنهادی در سناریوهای iid و غیر iid که در آن $\alpha = 0.2$ و $\beta = 40$ است.

دقت آزمایش (%)



(الف) داده‌های iid



(ب) داده‌های غیرمتغیر

شکل ۷: تست دقتهای سیاست‌های زمان‌بندی مختلف، که در آن $\alpha = 0.2$ ، $\beta = 3$ ، $\gamma = 4$ و $\delta = 3$ است.

با روش‌های زمان‌بندی تصادفی و پیشنهادی، افت آموزش در سناریوی iid کمتر از سناریوی غیر iid است، که نشان می‌دهد هرچه کمتر باشد، افت آموزش کمتر است.

برای هر دو سناریوی iid و non-iid، طرح زمان‌بندی پیشنهادی ما از زمان‌بندی تصادفی بهتر عمل می‌کند، که مزیت انتخاب دستگاه‌های با کیفیت لینک بهتر و در مجموع، نمایش داده‌های همگن‌تر را تأیید می‌کند.

برای نشان دادن تأثیر ناهمزنانی درون تکرار، نتیجه بهروزرسانی‌های همزمان با زمان‌بندی تصادفی (منحنی سبز) را ارائه می‌دهیم. در مقایسه با حالت ناهمزنان (منحنی آبی)، حالت همزمان نتایج آموزش کمتری دارد، که نشان می‌دهد ناهمزنانی درون تکرار برای عملکرد همگرایی مضر است.

علاوه بر این، همانطور که در شکل ۴b با داده‌های غیر iid نشان داده شده است، می‌بینیم که با کوچکتر، که به معنی بزرگتر (۰) و بالاتر است شدنی مذکور پیدا می‌کنیم و ترتیبی داشتنون، دستگاه‌هایی که بالاترین معیار کوئنتیتی دارند. این موضوع، بینش‌های ذکر شده در نکته ۴ را تأیید می‌کند که در سناریوهای داده‌های غیر iid درجه کمتر ناهمزنانی درون تکرار و فشرده‌سازی بهروزرسانی منجر به تحریب بیشتر عملکرد پادگیری می‌شود.

مبتنی بر سن : [13] ابتدا C را با استفاده از همان روشی که در ۲ BCBN2 استفاده شده بودیم می‌کنیم و ترتیبی داشتنون، دستگاه‌هایی که بالاترین معیار کوئنتیتی دارند. این موضوع، بینش‌های ذکر شده در نکته ۴ را تأیید می‌کند که در سناریوهای داده‌های غیر iid درجه کمتر ناهمزنانی درون تکرار و فشرده‌سازی بهروزرسانی منجر به تحریب بیشتر عملکرد پادگیری می‌شود.

$$\frac{1}{\prod(\cdot)} \cdot \frac{\partial}{\partial \cdot} = \frac{\partial}{\partial \cdot} \cdot \frac{1}{\prod(\cdot)}$$

می‌بینیم که سیاست زمان‌بندی پیشنهادی، در هر دو سناریوی iid و غیر iid زمان‌بندی تصادفی یا روش‌های مرتعهتر عمل می‌کند. علاوه بر این، در سناریوی غیر iid، مشاهده می‌کنیم از داده‌ها در روش پیشنهادی، دقتهای از این سیاست با روش‌های صرفه جایگزین می‌شوند. اگاه از کanal BC و [11] BCBN2 دارد.

ج. مقایسه سیاست‌های زمان‌بندی

ابتدا، وزن‌های تجمعی با $=1$ در (13) تنظیم می‌شوند. در شکل ۷، دقتهای آزمایش طرح زمان‌بندی پیشنهادی را با برخی روش‌های مرتعه جایگزین نشان می‌دهیم.

ما تا حداقل دستگاه‌ها را به طور یکنواخت و تصادفی انتخاب می‌کنیم. BCBN2 [11]: ما تا حداقل دستگاه‌ها با بالاترین (۰) انتخاب می‌کنیم. [11] BC: ابتدا یک زیرمجموعه دستگاه (C K C) با $|C| = 0.5$ پیدا می‌کنیم که بالاترین (۰) را دارد. سپس، در میان، C دستگاه‌هایی با بالاترین (۰) را زمان‌بندی می‌کنیم. (۱)

د. مقایسه سیاست‌های تجمعی

در شکل ۸، عملکرد سیاست‌های مختلف تجمعی را تحت روش‌های زمان‌بندی مبتنی بر اهمیت داده و زمان‌بندی تصادفی اگاه از کanal پیشنهادی نشان می‌دهیم. می‌بینیم که طراحی اگاه از سن با $=1$ در خط پایه بهتر عمل می‌کند.

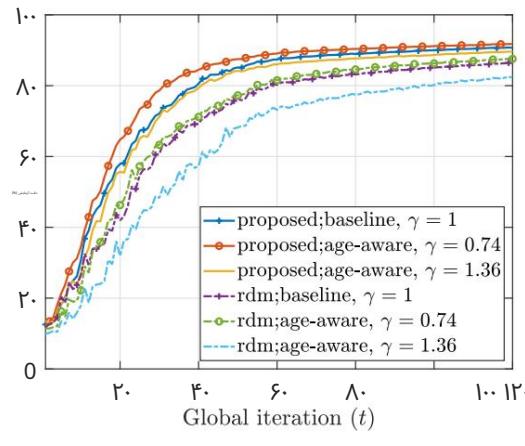
8 برای مقایسه‌ای منصفانه، تخصیص منابع نماد از (9) پیروی می‌کند که متفاوت است از طراحی در [11].

9 این یک نسخه اصلاح‌شده از روش [13] است که به همین ترتیب دستگاه‌های با کیفیت کanal بهتر را پیدا می‌کند و میزان کل خرابی دستگاه را به حداقل می‌رساند.

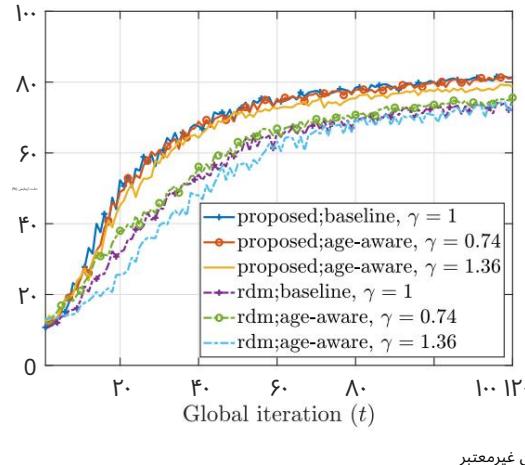
هر دو سناریوی iid و غیر iid برای حالت favoring-fresh و زنده iid افزایش عملکرد فوق العاده دارد. برای حالت غیر iid، افزایش عملکرد طرح تجمعی پیشنهادی زمانی که از زمان بندی تصادفی استفاده می‌شود، آشکارتر است.

هفتم، تقدیر و تشکر

از فردیک جانسون به خاطر مشارکتش در ایده سیاست زمان بندی در طول پروژه پایان نامه کارشناسی ارشدش در بخش سیستم‌های ارتباطی، دانشگاه لینشوبینگ، تشکر می‌کنیم.



(الف) داده‌های iid



(ب) داده‌های غیرiid

شکل ۸: تست دقت سیاست‌های تجمعی پیشنهادی، که در آن $\gamma = 0.74$ و $= 1$ می‌باشد.

در این کار، ما یک چارچوب FL ناهمzman با تجمعی دوره‌ای پیشنهاد کردیم که مزایای آموزش ناهمzman و تجمعی مدل همزمان را با هم ترکیب می‌کند.

برای طرح پیشنهادی، ما یک سیاست زمان بندی مبتنی بر اهمیت داده آگاه از کانال و طرح تجمعی مبتنی بر سن برای FL تحت محدودیت‌های منابع بی‌اسیم را توسعه دادیم.

طراحی زمان بندی و تجمعی پیشنهادی ما، به ویژه با داده‌های آموزشی ناهمگن در بین دستگاه‌های مختلف، عملکرد بهتری نسبت به روش‌های موجود نشان داده است. پیام اصلی این است که اصل طراحی زمان بندی و تخصیص منابع در FL بی‌اسیم باید مبتنی بر کاهش بایاس و واریانس بهاروزرسانی‌های محلی تجمعی شده باشد. برای تنظیمات FL ناهمzman، معنادل کردن تازگی و مفید بودن بهاروزرسانی‌های مدل محلی و تنظیم سهم آنها در مدل جهانی جدید نیز یک جنبه مهم طراحی است.

پیوست هشتم

الف. اثبات قضیه ۱

ما اثبات را با رویکردی مشابه [25] و [29] انجام می‌دهیم، به جز برعی دستکاری‌های اضافی برای مدیریت تأثیر آموزش ناهمzman. همانطور که به طور خلاصه در بخش II نشان داده شده است، تحلیل شکاف بهینگ را از (12) شروع می‌کنیم.

تعریف می‌کنیم.

$$() = \square \quad ()$$

$$\square M ()$$

به عنوان مجموع وزنی در تکرار سراسری و

$$() + 1 = \square \frac{()}{\square M ()} \quad (., 1).$$

و سپس (12) را می‌توان به صورت زیر بازآرایی کرد:

$$\begin{aligned} E [((+1))] &= \frac{\square}{2} E | | \square_{=1}^{\square} \\ &= \frac{1}{2} E () \square_{=1}^{\square} - ((+1)) \frac{\square}{2} \square_{=1}^{\square}. \end{aligned} \quad (19)$$

برای دلیل این انتقال (ایضاً) این را ارزیابی کرده و نمادگاری متغیرهای موجود را ساده می‌کنیم. $E [((+1))] = E ((+1))$ (آنچه در (19) از آن دستمزد نظر نداشت) و $E ((+1)) = 0$ برای سادگی.

به ترتیب، علاوه بر این،

$$\square (., 0) = \square ((., 0); B ()).$$

تخصیص منابع در (9) و محدودیت انتقال در (10) منجر به معادل \square می‌شود. از این رو،

علاوه بر این، ما برخی متغیرها و لمها را معرفی می‌کنیم.

$$() = \square \frac{\square}{\square M} \quad ((., 0); B,) \quad (20)$$

$$\text{که در آن } () \square = \square (., 0); B () = \square (., 0).$$

$$- () = EB () EQ EV () [\bar{()}] = \square \frac{\square}{\square M} \quad ((., 0)). \quad (21)$$

۱۳ در این قسمی کاتچور (کاتچور) شده است، آگاه

$$- \frac{1}{2} \square \square \square \square = \frac{1}{2} \square \square \square \square \quad ((., 0)).$$

لم ۲. افرض کنید \square در قضیه اعریف شده باشند، آگاه واریانس $(.)$ در رابطه زیر صدق می‌کنند:

$$\frac{1}{2} \square \square \square \square = \frac{1}{2} \square \square \square \square \quad (22)$$

اثبات به بخش های VIII-C، VIII-B مراجعه کنید.

$$(20)-(21), \text{ بر اساس } | + \frac{1}{M} \text{ می‌تواند به صورت زیر بازنویسی شود}$$

$$\begin{aligned} & \frac{1}{M} \frac{(+)^r}{2} = \frac{| + \frac{1}{M}}{M} - (0) \frac{()^r}{2} \\ & = \frac{1}{M} \frac{()^r}{2} - \frac{()^r}{2} \frac{| + \frac{1}{M}}{M} - (0) \frac{()^r}{2} + (0) \frac{()^r}{2} \\ & = \frac{1}{M} \frac{()^r}{2} - \frac{()^r}{2} + (0) \frac{()^r}{2} - (0) \frac{()^r}{2} \quad (22) \end{aligned}$$

$$= \frac{1}{M} \frac{()^r}{2} - (0) \frac{()^r}{2} \quad (23)$$

که در آن (22) به دلیل $(0) = 0$ است. بالم ۱۰، (23) کل انتظار (24) برآورده می‌شود.

$$\begin{aligned} & \text{ای } | + \frac{1}{M} \frac{()^r}{2} - (0) \frac{()^r}{2} \quad (\text{دقتیه}) \\ & + (0) \frac{()^r}{2} - (0) \frac{()^r}{2} \quad (24) \end{aligned}$$

طبق تعریف $(0) = 0$. $E[(0)] = 0$ با هم با لم ۲، (24) را می‌توان به صورت زیر بازآرایی کرد:

$$\begin{aligned} & \text{ای } | + \frac{1}{M} \frac{()^r}{2} - (0) \frac{()^r}{2} \quad (\text{دقتیه}) \\ & + (0) \frac{()^r}{2} - (0) \frac{()^r}{2} \quad (25) \\ & \text{لکه‌گذاری شده } (0) \frac{()^r}{2} = \frac{1}{2} \frac{()^r}{2} \text{ که شرط کافی در} \\ & \text{لم، اما ادعا می‌کنیم که} \end{aligned}$$

$$\begin{aligned} & \text{که در آن } = \text{حداکثر } (0) \text{ و } (25) \\ & = \frac{1}{2} \frac{()^r}{2} + \frac{1}{2} \frac{()^r}{2} + \frac{1}{2} \frac{()^r}{2} \\ & = \text{استقرای ثابت می‌کنیم. برای } 1, \\ & \frac{1}{2} \frac{()^r}{2} = \frac{(+)^r}{2} + \frac{(+)^r}{2} + \frac{(+)^r}{2} \quad (26) \end{aligned}$$

آنچه این اثبات می‌کند، فرض می‌کنیم

$$\begin{aligned} & \text{ای } | + \frac{1}{M} \frac{()^r}{2} - (0) \frac{()^r}{2} \quad (\text{دقتیه}) \\ & + \frac{1}{M} \frac{()^r}{2} + \frac{1}{M} \frac{()^r}{2} - (0) \frac{()^r}{2} + \frac{1}{M} \frac{()^r}{2} \\ & = \frac{+ \frac{1}{M}}{(+)^r} + \frac{[+(1) \frac{()^r}{2}]}{(+)^r} + \frac{1}{M} \frac{()^r}{2} \\ & + \frac{1}{M} \frac{()^r}{2} - (0) \frac{()^r}{2} + \frac{1}{M} \frac{()^r}{2} \\ & = \frac{+ \frac{1}{M}}{(+)^r} + \frac{+ \frac{1}{M}}{(+)^r} - (0) \frac{()^r}{2} \quad (27) \\ & \frac{+ \frac{1}{M}}{(+)^r} + \frac{+ \frac{1}{M}}{(+)^r} \end{aligned}$$

که در آن (27) به دلیل منفی بودن جمله دوم در (26) است.

بنایلریلین (Lambert's function) برقرار است و با جایگذاری نتیجه در (19)

$$\begin{aligned} & \frac{1}{2} \frac{()^r}{2} - (0) \frac{()^r}{2} + \frac{1}{M} \frac{()^r}{2} \\ & + \frac{1}{2} \frac{()^r}{2} + \frac{1}{2} \frac{()^r}{2} + \frac{1}{2} \frac{()^r}{2} \\ & = \frac{1}{2} \frac{()^r}{2} + \frac{1}{2} \frac{()^r}{2} + \frac{1}{2} \frac{()^r}{2} \end{aligned}$$

که در آن $[EX]$ انتظار برای تمام تصمیمات زمان‌بندی گذشته است، یعنی $(0) = M = 1, \dots$ و $M = 1$ (15) به دلیل.

ب. اثبات لم ۱

$$\begin{aligned} & \frac{1}{2} \frac{()^r}{2} - (0) \frac{()^r}{2} + \frac{1}{M} \frac{()^r}{2} \\ & = \frac{1}{2} \frac{()^r}{2} + \frac{1}{2} \frac{()^r}{2} + \frac{1}{2} \frac{()^r}{2} \\ & = \frac{1}{2} \frac{()^r}{2} + \frac{1}{2} \frac{()^r}{2} + \frac{1}{2} \frac{()^r}{2} \quad (28) \\ & = \frac{1}{2} \frac{()^r}{2} + \frac{1}{2} \frac{()^r}{2} + \frac{1}{2} \frac{()^r}{2} \quad (29) \\ & = \frac{1}{2} \frac{()^r}{2} + \frac{1}{2} \frac{()^r}{2} + \frac{1}{2} \frac{()^r}{2} \quad (30) \\ & = \frac{1}{2} \frac{()^r}{2} + \frac{1}{2} \frac{()^r}{2} + \frac{1}{2} \frac{()^r}{2} \quad (31) \end{aligned}$$

که در آن (28) و (29) به دلیل تحدب $2 \frac{1}{2} + \frac{1}{2}$ هستند. عبارت سوم (29) عبارت است از محدود شده توسط تحدب (0) .

با اعمال (30) و (31) در بخش (29) به صورت ۱۰ به عقب مرتب می‌شود.

$$\begin{aligned} & \frac{1}{2} \frac{()^r}{2} - (0) \frac{()^r}{2} + \frac{1}{M} \frac{()^r}{2} \\ & = \frac{1}{2} \frac{()^r}{2} + \frac{1}{2} \frac{()^r}{2} + \frac{1}{2} \frac{()^r}{2} \\ & = \frac{1}{2} \frac{()^r}{2} + \frac{1}{2} \frac{()^r}{2} + \frac{1}{2} \frac{()^r}{2} \quad (0) \frac{()^r}{2} - (0) \frac{()^r}{2} + (0) \frac{()^r}{2} \end{aligned}$$

برای جزئیات به بخش D مراجعه کنید.

سپس، بر اساس (16) و (17) داریم

بپس، بر اساس (16) و (17) داریم

(33)

$$\begin{aligned} & \text{نوجه داشته باشید که} \\ & \frac{\partial}{\partial x} \left(\frac{\partial u}{\partial x} \right) = \frac{\partial^2 u}{\partial x^2}, \quad \frac{\partial}{\partial y} \left(\frac{\partial u}{\partial y} \right) = \frac{\partial^2 u}{\partial y^2}, \quad \frac{\partial}{\partial z} \left(\frac{\partial u}{\partial z} \right) = \frac{\partial^2 u}{\partial z^2} \end{aligned}$$

وجه داشته باشید که

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_i} \right) - \frac{\partial \mathcal{L}}{\partial x_i} = 0 \quad (34)$$

جايگذاري (34) در، (33) داريم

$$\text{از آنجایی که } \frac{1}{(4\pi)^2} \text{ است، داریم:}$$

سپس،

$$1 - \frac{1}{(4\pi)^2} < 0.$$

(32)

همانطور که حد در (35) به زمان‌بندی فعلی بستگی ندارد
سیاست، ماینرات کامل شده است.

ج. اثبات لم ۲

با جایگذاری (۳۲) در (۳۱) داریم

$$(1) \square (2)+\square (2)(12) \square \square \square$$

لفرانچی کنید R ۲ ۱، است، هموار (۱) (۲) (۳)

۲. اثبات ۲

ج. اثبات لم

EB ()EQ EV () || () () ~

سمت راست (36) را به عنوان یک تابع جدید تعریف کنید.

$$(1) = (2) + \square(2)(1 \square 2) + \dots$$

از آنجا که (1) یک تابع درجه دوم با حداقل سراسری است
 $\Rightarrow \boxed{0} = \boxed{0}$ نتیجه می‌اشود که

$$\square(2) \frac{d}{dt} \square(2) = \dots = \square(2) + \square(2)$$

Learning curves

$$) = ((2) \square + | | \square (2) | | 2 \quad (37)$$

٣٦، ٣٧ تكib

دقيقة (2) | | (2) | | ٢

14

$$EB()EQM \quad || Q(((, 0) : B())) \quad \text{و از این رو} \\ () || 2 ب ۰ ۰ ۰ () () \quad (2) \square \quad \square | \square (2) || ۲. \quad (38)$$

- منابع
- [1] [ج]. کونکنی، اچین مکاماهان، دی، رامیچ، و بی، ریچاریک، «بهینه‌سازی فدرال: یادگیری ماشین توزیع شده برای هوش روبی دستگاه»، نسخه پیش‌چاپ arXiv arXiv:1610.02527، 2016.
- [2] [آپ. مکماهان، ای، مور، دی، رامیچ، اس، همپسون، و بی، ای، آکاس، «یادگیری کارآمد ارتباطی شبکه‌های عمیق از داده‌های غیرهمترک»، در هوش مصنوعی و آمار، 1282-1273، 2017.
- [3] [جن، چن، آر، موگا، اس، بندجو، و آر، جوزفوچ، «بازنگری SGD توزیع شده هم‌زمان»، در کنفرانس بین‌المللی کارگاه‌های آموزش بازنمایی‌های یادگیری، 2016.
- [4] [آس، شن، اس، کوچو، و آنک، «بهینه‌سازی فدرال ناهم‌زمان»، در کارگاه NeurIPS در مورد بهینه‌سازی برای یادگیری ماشین، 2020.
- [5] [آی، آنک، چن، ی.، نینگ، م.، سلاوسکی، و ه. رانگوالا، «یادگیری فدرال آتلابن ناهم‌زمان برای دستگاه‌های لیهای با داده‌های غیر- IID»، در کنفرانس بین‌المللی IEEE در مورد کلان داده، 15-24، 2020.
- [6] [آر، وانگ، اچ، شو، جن، لیو، اچ، هوانگ، سی، کیانو، و ای، ژانو، «یادگیری فدرال با بهره‌وری منابع و تجمعی سلسه مراتب در محاسبات لیهای»، در IEEE INFOCOM، 2021.
- [7] [جن، گوین، کی، مالک، اچ، زان، ای، یوسف‌پور، ام، ریاض، ام، مالک، و دی، هونا، «یادگیری فدرال با تجمعی ناهم‌زمان با فرآیند مجموعه مقاولات بسته و پنجمین کنفرانس بین‌المللی هوش مصنوعی و آمار، جلد 151، 3607-3581، 2022.
- [8] [سی، زو، اچ، شو، جن، لیو، اچ، زانگ، جن، یانک، ام، دونگ، و جی، چیا، «Tea-fed: یادگیری فدرال ناهم‌زمان با زمان کارآمد برای محاسبات لیهای»، در مجموعه مقاولات هجدهمین کنفرانس بین‌المللی ACM در مورد مزدیهای محاسبات، 2021.
- [9] [آن، گافنی، ان، شلزیگر، کن، کوهن، وای، سی، الدار، و اچ، وی، پور، «سیاست‌های زمان‌بندی برای یادگیری فدرال در شبکه‌های بین‌سیم»، در IEEE Trans. on Communications، 2022.
- [10] [آج، اچ، زد، لیو، تن، اس کونک، و اچ وی پور، «سیاست‌های زمان‌بندی برای یادگیری فدرال در شبکه‌های بین‌سیم»، در IEEE Trans. on Communications، 2020.
- [11] [آی، وینسنت پور، و H. Vincent Poor، MM Amiri، D. Gündüz، SR Kulkarni، «همگایی زمان‌بندی دستگاه آگاه از پیغام‌رسانی برای یادگیری فدرال در لبه بین‌سیم»، در IEEE Trans. on Communications، 2021.
- [12] [آی، لیو، جن، ژو، و کیو کونک، و اچ وی پور، «سیاست زمان‌بندی مبتنی بر سن برای یادگیری فدرال در شبکه‌ای موبایل»، در کنفرانس بین‌المللی IEEE از داده ارتباطات و شبکه‌های شناختی، صفحات 1739-1748، IEEE INFOCOM، 2022.
- [13] [آی، اح، جن، ژانگ، و کی، هوانگ، «یادگیری فدرال در شبکه‌ای موبایل»، در کنفرانس بین‌المللی ICASSP، 2020.
- [14] [آی، صالحی و ای، حسین، «یادگیری فدرال در شبکه‌ای بین‌سیم سلوالی غیرقابل اعتماد و با محدودیت منابع»، IEEE Trans. on Communications، 2021.
- [15] [آی، لیو، جن، ژو، و کی، هوانگ، «زمان‌بندی کاربر با توجه به اهمیت داده‌ها برای یادگیری ماشینی لیهای با کارایی ارتباطی بالا»، در IEEE Trans. on Communications، 2020.
- [16] [آی، دی، گونک، و کی، هوانگ، «یادگیری لیهای سریع فدرال با ارتباطات و محاسبات همپوشانی و برآنماییزی منصفانه کلاینت آگاه از کانال»، در بیست و دومین کارگاه بین‌المللی IEEE در مورد پیشرفت‌های پردازش سیگنال در ارتباطات بین‌سیم، 2021.
- [17] [آی، لیو، جن، ژو، و کی، هوانگ، «یادگیری فدرال ناهم‌گن با کاهش واریانس اسکالنده، و اچ، وی، پور، «یادگیری فدرال برای شبکه‌های بین‌سیم اینترنت اشیا با ارتباطات و شبکه‌های شناختی»، IEEE Trans. on Communications، 2021.
- [18] [آی، تائیک، ه. مودود، و C. Cherkaoui، «Zمن‌بندی مبتنی بر کیفیت داده برای یادگیری لیهای فدرال»، در چهل و ششمین کنفرانس IEEE از داده ارتباطات و شبکه‌های کامپیوتوری مجلن (LCN)، 2021.
- [19] [آی، شن، دی، گانور، ال، یانک، اف، زو، دی، سونگ، دبلیو، لو، و اس، بی، «یادگیری فدرال ناهم‌گن با کاهش واریانس از طریق انتخاب کلیدت طبقه‌بندی شده»، 2022.
- [20] [آی، شن، دی، گانور، ال، یانک، اف، زو، دی، سونگ، دبلیو، لو، و اس، بی، «یادگیری فدرال ناهم‌گن با کاهش واریانس از اس، زو، زد، نیو، ام، جیانگ، و ال، گنگ، «زمان‌بندی مشترک دستگاه و تخصیص منابع برای یادگیری فدرال بین‌سیم با محدودیت تأخیر»، در IEEE Trans. از داده ارتباطات بین‌سیم، جلد 20، شماره 1، 467-453، 2021.
- [21] [آم، ادو، اس، ساماکون، و ام، بنیس، «زمان‌بندی مشترک مشترک و تخصیص منابع تحت عدم قطعیت کانال در یادگیری فدرال»، IEEE Trans. از طریق شبکه‌ای بین‌سیم، 2021.
- [22] [آم، چن، زد، دبلیو، سعد، سی، بن، اچ، وی پور، و اس، کوی، «یک چارچوب مشترک یادگیری و ارتباطات برای یادگیری فدرال از طریق شبکه‌ای بین‌سیم»، در IEEE Trans. از داده ارتباطات بین‌سیم، جلد 20، شماره 1، 283-269، 2021.
- [23] [آس، لی، و جن، دبلیو، لی، «زمان‌بندی انتقال تطبیقی در شبکه‌های بین‌سیم برای یادگیری فدرال ناهم‌زمان»، مجله IEEE در حوزه‌های منتخب در ارتباطات، جلد 39، شماره 12، صفحات 3673-3687، 2021.
- [24] [آی، رانگوالا، H. Rangwala، «FedAT: و یک سیستم یادگیری فدرال با کارایی بالا و کارآمد از نظر ارتباطی با لایه‌های ناهم‌زمان»، در مجموعه مقاولات کنفرانس بین‌المللی محاسبات، شبکه‌سازی، دخیره‌سازی و تحلیل با کارایی بالا، 2021.
- [25] [آی، کوستا، ان، کی، هوانگ، دبلیو، یانگ، اس، وانک، و زد، ژانگ، «دریاره همگرایی FedAvg روی داده‌های غیر-IID در کنفرانس بین‌المللی بازنمایی‌های یادگیری، 2020.
- [26] [آی، کوستا، ان، کی، هوانگ، دبلیو، یانگ، اس، وانک، و زد، ژانگ، «عصر اطلاعات: یک مفهوم، معیار و ابزار جدید»، مبانی و روندها در شبکه‌سازی، جلد 12، شماره 3، صفحات 162-259، 2017.
- [27] [آی، آلسینار، دی، گوبیچ، جن، ل، آر، تومیوکا، و ام، وینوچوچ، «SGD با کارایی ارتباطی از طریق کوئیتویا و دک‌داری گردابیان»، در مجموعه مقاولات سی و پنجمین کنفرانس بین‌المللی سیستم‌های پردازش اطلاعات تصمیم، 2017.
- [28] [آی، بوتو، L. Bottou، FE Curtis، «روش‌های بهینه‌سازی برای یادگیری ماشین در مقایس بزرگ»، SIAM Review، doi.org/10.1137/16M1080173 [29] SU Stich، «SGD: موجود در: چند ایندیبن»، موجود در: SGD، 2018.
- [29] [آی، بوتو، L. Bottou، FE Curtis، «روش‌های بهینه‌سازی برای یادگیری ماشین در مقایس بزرگ»، SIAM Review، doi.org/10.1137/16M1080173 [29] SU Stich، «SGD: موجود در: چند ایندیبن»، موجود در: SGD، 2018.
- [30] [آی، نوچال، M. Nocedal، M. van Dijk، «جنبه‌های گلگایانه‌ای گردابیان تصادفی»، مجله تحقیقات یادگیری ماشین، جلد 60، شماره 2، صفحات 311-223، 2018.
- [31] [آی، دیگاوی، S. Diggavi، «SQuARM-SGD: SGD به سرعت همگرا می‌شود و ارتباط کمی برقرار می‌کند»، در کنفرانس بین‌المللی بازنمایی‌های یادگیری، 2019.
- [32] [آی، کولوسکوا، A. Koloskova، N. Loizou، S. Boreiri، M. Jaggi، «یک نظریه یکاره از SGD غیرهمترک با تغییر توپولوژی و بازروزسانی‌های محلی»، در مجموعه مقاولات سی و هفتمین کنفرانس بین‌المللی یادگیری ماشین، جلد 13-18 PMLR، 2019.
- [33] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [34] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [35] [آی، دیگاوی، S. Diggavi، «SQuARM-SGD: SGD به سرعت همگرا می‌شود و ارتباط کمی برقرار می‌کند»، در کنفرانس بین‌المللی بازنمایی‌های یادگیری، 2019.
- [36] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [37] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [38] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [39] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [40] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [41] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [42] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [43] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [44] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [45] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [46] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [47] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [48] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [49] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [50] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [51] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [52] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [53] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [54] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [55] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [56] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [57] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [58] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [59] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [60] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [61] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [62] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [63] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [64] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [65] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [66] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [67] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [68] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [69] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [70] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [71] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [72] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [73] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [74] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [75] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [76] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [77] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [78] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [79] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [80] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [81] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [82] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [83] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [84] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [85] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [86] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [87] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [88] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [89] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [90] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [91] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [92] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [93] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [94] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [95] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [96] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [97] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [98] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [99] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [100] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [101] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [102] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [103] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [104] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [105] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [106] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [107] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [108] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [109] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [110] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [111] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [112] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [113] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [114] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [115] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [116] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [117] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [118] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [119] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [120] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [121] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [122] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [123] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [124] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [125] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [126] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [127] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [128] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [129] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [130] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [131] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [132] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [133] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [134] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [135] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [136] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [137] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [138] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [139] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [140] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [141] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [142] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [143] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [144] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [145] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [146] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [147] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [148] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [149] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [150] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [151] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [152] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [153] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [154] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [155] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [156] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [157] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [158] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [159] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [160] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [161] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [162] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [163] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [164] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [165] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [166] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [167] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [168] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [169] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [170] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [171] [آی، ای، لکان و سی، کورتس، «یاگاه داده ارقام دست‌نویس»، MNIST، 2010.
- [172] [آی، ای، لکان و سی، کورتس، «یا



کن چن (عضو IEEE) استادیار دانشکده مهندسی برق، دانشگاه لیونزپیونگ، سوئد است. او در کرکا کارشناسی خود را در سال ۱۹۸۱-۱۹۸۲ در دانشگاه علم و فناوری هوازوک (HUST) (جن درافت ۱۹۸۳). سپس، به ترتیب در سال‌های ۱۹۸۴-۱۹۸۵ کارشناسی ارشد و درجه اول خود را در شرکه سنتلر سوپلیک، دانشگاه پاریس-ساکل، فرانسه دریافت کرد. از آن‌ویژه ۱۹۸۷-۱۹۸۸ در شهاب‌الهی کارشناسی لیونزپیونگ و شکمکه‌های پژوهش و تحقیقات انجام داشت. علاوهً اصلی تحقیقات او شامل ارتباطات سیمی، سیستم‌های هوشمند توپوزیشن و شبکه‌های پیچیده است. در دیارکت‌کننده جایزه نوین مقاله نویسنده جوان احتمان ارتباطات IEEE در سال ۲۰۰۲-۲۰۰۳. او در سال ۲۰۱۶-۲۰۱۷ IEEE Communications Letters، ۲۰۱۸-۲۰۱۹ IEEE Transactions on Communications و ۲۰۱۹-۲۰۲۰ IEEE Transactions on Wireless Communications از این مجلات انتشار امداد و در سال ۲۰۱۸-۲۰۱۹ به عنوان مدرس مشترک کارگاه اطلاعات شد. او در سال‌های ۲۰۱۱-۲۰۱۲ به عنوان مدرس مشترک IEEE Transaction on GLOBECOM از مردم رسنی ارتباطات بی‌سیم برای هوش توپوزیشن خدمت کرد. او در حال حاضر سردبیر IEEE Transactions on Green Communications and Networking است.



کج، لارسون (عضو EEE) امادرک دکترا را در سال ۲۰۱۴ دانشگاه اوپسالا، اوپسالا، سوئد
گرفت. او در حال حاضر استاد سیستم‌های اینطاپی در دانشگاه شنیپوینگ (LNU) اند
شوپیوینگ، سوئد است. او با موسسه علمی طبلت فناوری KTH در استکهلم، سوئد، دانشگاه
رجوع و اشتغال، ایالات متحده آمریکا، دانشگاه فلوریدا، ایالات متحده آمریکا و مرکز تحقیقات
کسون، سوئد همکاری داشته است.

دیقه، اصل، حرفهای او در جزو های ارتباطات پیشیم و بدانش سیگنال است.

در نگارش کتابهای «کدگذاری بلوکی فضای-زمان برای ارتباطات بی‌سیم» (انتشارات دانشگاه میریج، ۲۰۰۳) و «مبانی MIMO» (انتشارات دانشگاه کمبریج، ۱۶۰۰-۲۰۱۶) داشته است.

دوبار در سال‌های ۱۴-۱۳ و ۱۲-۱۱ جایزه بهترین ستون مجله پردازش سیگنال IEEE، این‌در سال ۱۵-۱۴ جایزه ComSoc Stephen O. Rice و Leonard G. Abraham. در نظریه ارتباطات، در سال ۱۷-۱۶ جایزه بهترین مقاله آموزش IEEE ComSoc Fred W. Ellersick را دریافت کرد.