Vol. 5, No. 17, Jan-Mar 2026,(84-93)

Journal of New Achievements in Electrical, Computer and Technology (JNAECT)

مجله دستاوردهای نوین در برق،کامپیوتر و فناوری

دوره ۵، شماره ۱۷، زمستان ۱۴۰۴ (۹۳-۸۴)

journal homepage: https://jnaect.ir

تحلیل فدرال در شبکه‌های لبه: بینش‌های حفظ حریم خصوصی از داده‌های توزیع‌شده

محمدسینا پورشعبانیان[۱]، محمدمهدی شیرمحمدی[۲]

۱ گروه کامپیوتر، واحد همدان، دانشگاه آزاد اسلامی، همدان، ایران Pourshabanian@icloud.com

۲ گروه کامپیوتر، واحد همدان، دانشگاه آزاد اسلامی، همدان، ایران Mmshirmohammadi@iau.ac.ir

چکیده

با گسترش سریع اینترنت اشیا و رایانش لبه، حجم داده‌های توزیع‌شده به‌صورت چشمگیری افزایش‌یافته است؛ موضوعی که ضرورت توسعهٔ چارچوب‌های تحلیلی با حفظ حریم خصوصی را برجسته می‌کند. این مقاله مروری، به بررسی جامع تحلیل فدرال در شبکه‌های لبه با تمرکز بر حفظ حریم خصوصی داده‌ها می‌پردازد.

انگیزهٔ اصلی این پژوهش، نیاز به ایجاد تعادل میان کارایی داده و حفاظت از حریم خصوصی در شرایطی است که مقررات سخت‌گیرانه‌ای مانند مقررات عمومی حفاظت از داده‌ها در حال اجرا هستند. هدف مقاله، تبیین و تحلیل روش‌های کلیدی حفظ حریم خصوصی در چارچوب تحلیل فدرال است که شامل حریم خصوصی تفاضلی محاسبهٔ چندجانبهٔ امن و تجمیع رمزنگاری شده می‌شود.

یافته‌های این مطالعه نشان می‌دهد که ترکیب فشرده‌سازی مدل با رمزنگاری سبک مؤثرترین راهکار برای محیط‌های با پهنای باند محدود است. بااین‌حال، ناهمگنی داده‌ها و سربار ارتباطی بالا همچنان از مهم‌ترین چالش‌های باقی‌مانده در پیاده‌سازی گستردهٔ تحلیل فدرال محسوب می‌شوند. در پایان، مقاله بر چشم‌اندازهای پژوهشی آینده تمرکز دارد که شامل بهینه‌سازی ارتباطات میان‌گره‌های لبه، مدیریت ناهمگنی آماری و سیستمی و ایجاد معیارهای استاندارد برای ارزیابی عملکرد سیستم‌های فدرال است.

واژه‌های کلیدی: تحلیل فدرال، رایانش لبه، حفظ حریم خصوصی، حریم خصوصی تفاضلی، تجمیع امن، ناهمگنی داده‌ها.

۱. مقدمه

با انفجار اینترنت اشیا (IoT[1]) و پیش‌بینی رسیدن تعداد دستگاه‌ها به ۱۲۵ میلیارد تا سال ۲۰۳۰، حجم عظیمی از داده در حال تولید است [۱]. این رشد سریع، نگرانی‌های جدی در مورد حفظ حریم خصوصی داده‌های حساس کاربران ایجاد کرده است که نیاز به رویکردهای نوین تحلیل داده را بیش‌ازپیش ضروری می‌سازد.

تحلیل فدرال[2] در شبکه‌های لبه[3] به‌عنوان یک چارچوب تحلیل داده توزیع‌شده ظهور کرده است که امکان کسب بینش از داده‌های تولید شده در دستگاه‌های لبه متعدد را بدون نیاز به متمرکزسازی اطلاعات حساس فراهم می‌کند. این رویکرد به نگرانی‌های حیاتی پیرامون حریم خصوصی، به‌ویژه باتوجه‌به مقررات سخت‌گیرانه‌ای مانند مقررات عمومی حفاظت از داده‌ها (GDPR[4]) و (CCPA[5]) که حفاظت قوی از داده‌های شخصی را الزامی می‌کنند، پاسخ می‌دهد [۲]. مزیت اصلی تحلیل فدرال، پردازش داده‌ها به‌صورت محلی بر روی دستگاه‌های لبه است که خطرات مرتبط با نقض و دست‌کاری داده‌ها در حین انتقال را به طور قابل‌توجهی کاهش می‌دهد [۳].

اگرچه تحلیل فدرال مزایای آشکاری در حفظ حریم خصوصی و بهینه‌سازی عملکرد مدل‌ها با داده‌های غیر IID[6](مستقل و با توزیع یکسان) [۴، ۵] ارائه می‌دهد، اما اثربخشی این سیستم با چالش‌های قابل‌توجهی مواجه است. مسائلی مانند ناهمگونی داده‌ها (data heterogeneity)، سربار ارتباطی (communication overhead) و محدودیت‌های محاسباتی دستگاه‌های لبه، پیاده‌سازی کارآمد آن را دشوار می‌سازد. علاوه بر این، پتانسیل حملات مخرب، از جمله مسمومیت مدل (model poisoning)، نگرانی‌های جدی را در مورد امنیت و یکپارچگی سیستم‌های فدرال ایجاد کرده است [۶].

تحقیقات فعلی، باوجود پیشرفت‌ها در تکنیک‌های حفظ حریم خصوصی [۷]، هنوز نتوانسته‌اند راه‌حل‌های جامع و کارآمدی برای مقابله با چالش‌های امنیتی و عملکردی در محیط‌های لبه ناهمگن ارائه دهند. همان‌طور که در [۶] اشاره شده است، آسیب‌پذیری در برابر حملات و محدودیت‌های ارتباطی و محاسباتی، از مهم‌ترین موانع موجود هستند. هدف اصلی این مقاله، پر کردن این شکاف تحقیقاتی از طریق ارائه یک چارچوب یا الگوریتم نوین است. این تحقیق تلاش می‌کند با پرداختن به این محدودیت‌ها، امنیت و کارایی تحلیل فدرال را بهبود بخشد و به توسعه بینش‌های مرتبط با حفظ حریم خصوصی در دنیای متصل امروز کمک کند.

۲. پیشینه پژوهش و مبانی نظری

تحلیل فدرال (FA) به‌عنوان یک پیشرفت محوری در حوزه تحلیل داده ظهور کرده است که به طور خاص به نیاز روزافزون برای حفظ حریم خصوصی در بحبوحه رشد تصاعدی دستگاه‌های اینترنت اشیا (IoT) پاسخ می‌دهد. تا سال ۲۰۲۱، تعداد دستگاه‌های IoT به حدود ۱۲,۳ میلیارد افزایش‌یافته است و پیش‌بینی‌ها نشان می‌دهد که این رقم تا سال ۲۰۳۰ می‌تواند به ۱۲۵ میلیارد دستگاه برسد [۱]. یکی از چالش‌های اصلی مرتبط با این سیل داده، خطر ذاتی برای حریم خصوصی کاربر است.

تحلیل فدرال باهدف کاهش این خطرات، چارچوبی را ارائه می‌دهد که «تحلیل داده‌ها در چندین دستگاه لبه را بدون نیاز به انتقال اطلاعات حساس به یک سرور مرکزی» تسهیل می‌کند [۸]. در زمینه محاسبات لبه، FAاز «نزدیکی (proximity)

[1] Internet of Things

[2] Federated Analytics

[3] Edge Network

[4] General Data Protection Regulation

[5] California Consumer Privacy Act

[6] Independent and Identically Distributed

تولید و پردازش داده‌ها» سرمایه‌گذاری می‌کند [۱]. این پارادایم امکان همکاری مؤثر بین دستگاه‌ها را فراهم می‌آورد و به آن‌ها اجازه می‌دهد تا محاسبات را به‌صورت محلی انجام دهند. درعین‌حال، حریم خصوصی از طریق تکنیک‌های نوآورانه‌ای مانند حریم خصوصی تفاضلی[۷]و روش‌های رمزگذاری (encryption) حفظ می‌شود [۳].

علاوه بر این، توانایی مدیریت داده‌های غیر IID(مستقل و با توزیع غیریکسان) در سیستم‌های فدرال، سازگاری و دقت مدل‌های تحلیلی را افزایش می‌دهد. مجموعه‌داده‌های مختلف به ویژگی‌ها و روش‌های خاصی برای تحلیل مؤثر نیاز دارند که FAبا به‌کارگیری رویکردی متناسب با ویژگی‌های هر مجموعه‌داده، به طور مؤثری این نیاز را برآورده می‌سازد [۴، ۵]. اگرچه گره‌های لبه (Edge nodes)ممکن است از نظر سیستم‌عامل و منابع متنوع باشند و با چالش‌هایی در تداوم سرویس مواجه شوند [۱]، تحلیل فدرال به‌عنوان یک چارچوب حیاتی برای بهره‌برداری از داده‌های لبه، ضمن اولویت‌دادن به حریم خصوصی کاربر در دنیایی که به طور فزاینده‌ای به هم متصل است، برجسته می‌شود.

۳. طبقه‌بندی و تاکسونومی روش‌ها

در زمینهٔ تحلیل فدرال در شبکه‌های لبه، روش‌های حفظ حریم خصوصی را می‌توان به سه‌شاخهٔ اصلی تقسیم کرد: روش‌های مبتنی بر رمزنگاری، روش‌های مبتنی بر اختلال نویز، و روش‌های ترکیبی و نوآورانه. این طبقه‌بندی، چارچوبی تحلیلی برای درک بهتر توازن میان حریم خصوصی، دقت و کارایی در سیستم‌های توزیع‌شده ارائه می‌دهد.

۳,۱ .روش‌های مبتنی بر رمزنگاری

این دسته شامل تکنیک‌هایی نظیر Secure Aggregation، Homomorphic Encryption، و SMPC[۸] است. این روش‌ها با استفاده از الگوریتم‌های رمزنگاری پیشرفته، از افشای داده‌های خام در حین تحلیل جلوگیری می‌کنند و سطح بالایی از حریم خصوصی را تضمین می‌نمایند [۷]، باوجود این مزیت، هزینه‌های محاسباتی و ارتباطی بالا یکی از چالش‌های عمدهٔ آن‌هاست که می‌تواند عملکرد کلی سیستم را کاهش دهد [۹]. به‌عنوان نمونه، مطالعهٔ [۷] نشان می‌دهد استفاده از Secure Aggregationمی‌تواند از حملات Model Poisoning جلوگیری کند، اما منجر به افزایش محسوس بار ارتباطی می‌شود.

۳,۲ .روش‌های مبتنی بر اختلال نویز

در این دسته، تکنیک DP[۹]در دو سطح Centralized و Local به کار گرفته می‌شود. این روش‌ها با افزودن نویز به داده‌ها یا پارامترهای مدل، مانع افشای اطلاعات حساس می‌شوند. به دلیل مقیاس‌پذیری و سبک‌وزن بودن، DPگزینه‌ای مناسب برای محیط‌های توزیع‌شده مانند شبکه‌های لبه است [۱۰]. بااین‌حال، کاهش دقت مدل یکی از پیامدهای معمول استفاده از DPاست، به‌ویژه در حوزه‌های حساس مانند سلامت [۱۱].مطالعه [۱۰] نشان می‌دهد که تنظیم دقیق پارامترهای نویز برای دستیابی به تعادل میان حفظ حریم خصوصی و دقت، چالشی کلیدی محسوب می‌شود.

۳,۳.روش‌های ترکیبی و نوآورانه

این دسته، ترکیبی از رویکردهای پیشین است که هدف آن بهینه‌سازی trade-off میان حریم خصوصی، دقت و کارایی است. به‌عنوان‌مثال، ادغام Differential Privacy با Secure Aggregation یا استفاده از فشرده‌سازی مدل‌ها می‌تواند ضمن

[7] Differential Privacy

[8] Secure Multi-Party Computation

[9] Differential Privacy

حفظ حریم خصوصی، هزینه‌های ارتباطی را کاهش دهد [۱۲]. این روش‌ها به دلیل انعطاف‌پذیری و قابلیت تطبیق بالا، برای محیط‌های ناهمگن با دستگاه‌های دارای توان محاسباتی متفاوت بسیار مناسب‌اند [۱۳].

بااین‌حال، پیچیدگی پیاده‌سازی یکی از چالش‌های اصلی آن‌هاست. مطالعه [۱۳] نشان می‌دهد الگوریتم‌های فشرده‌سازی پارامترها قادرند حجم داده‌های ارسالی را کاهش دهند، درحالی‌که سطح معقولی از حریم خصوصی حفظ می‌شود. به‌منظور مقایسهٔ شاخه‌های مختلف روش‌های حفظ حریم خصوصی در تحلیل فدرال لبه، جدول ۱ ارائه شده است.

جدول ۱. طبقه‌بندی روش‌های حفظ حریم خصوصی در تحلیل فدرال لبه

شاخه اصلی	سناریوهای کاربردی	محدودیت‌ها	مزایا	جزئیات فنی	نمونه روش‌ها	منابع
روش‌های مبتنی بر رمزنگاری	تحلیل داده‌های پزشکی یا مالی و IoT حساس؛ Edge	هزینه محاسباتی و پهنای باند بالا؛ تأخیر در یادگیری	حفظ قوی حریم خصوصی؛ عدم افشای داده‌های خام	رمزنگاری انتها به انتها؛ امکان جمع‌بندی مدل‌ها بدون دسترسی به داده‌های خام؛ برای SMPC چندین سرور همکاری	Secure Aggregation, Homomorphic Encryption, SMPC	[۷]
روش‌های مبتنی بر اختلال نویز	تحلیل داده‌های کاربران موبایل یا سنسورها؛ یادگیری شخصی‌سازی شده	کاهش دقت مدل؛ تنظیم حساس پارامتر نویز؛ نیاز به δ و ε مدیریت	سبک‌وزن؛ مقیاس‌پذیر؛ مناسب شبکه‌های لبه با محدودیت منابع	افزودن نویز Laplace یا Gaussian به گرادیان‌ها یا مدل؛ DP-Learning برای انتشار امن نتایج	Differential Privacy (Centralized / Local)	[۱۱]
روش‌های ترکیبی و نوآورانه	تحلیل داده‌های بزرگ و توزیع‌شده در حوزه سلامت یا IoT	پیچیدگی بالای پیاده‌سازی؛ نیاز به هماهنگی دقیق بین سرورها	توازن میان دقت، کارایی و حریم خصوصی؛ انعطاف‌پذیر	ترکیب نویز DP با رمزنگاری تجمعی؛ فشرده‌سازی گرادیان برای کاهش پهنای باند؛ الگوریتم‌های بهینه‌سازی هیبرید	ترکیب DP با Secure Aggregation, فشرده‌سازی مدل‌ها	[۱۳]

این تاکسونومی، نه‌تنها درک روش‌های موجود را تسهیل می‌کند، بلکه مسیرهای پژوهشی آینده برای توسعهٔ راهکارهای نوآورانه با توازن بهتر میان سه معیار کلیدی (Privacy–Accuracy–Efficiency) را نیز روشن می‌سازد [۱۴].

۴. ارزیابی تطبیقی و تحلیل نقادانه

باتوجه‌به گسترش روزافزون تحلیل فدرال در محیط‌های لبه، انتخاب رویکرد مناسب برای حفظ حریم خصوصی به یکی از چالش‌های اصلی پژوهشگران و طراحان سیستم‌های توزیع‌شده تبدیل شده است. اهمیت این مسئله به‌ویژه در کاربردهای حساس مانند سلامت دیجیتال، خدمات مالی و شهرهای هوشمند دوچندان است، زیرا هرگونه افشای داده می‌تواند منجر به نقض اعتماد کاربران، خسارات مالی و تهدید امنیت اطلاعات شود. ازاین‌رو، توجه به طراحی مکانیزم‌هایی که هم امنیت داده را

تضمین کنند و هم از نظر کارایی و مقیاس‌پذیری مناسب باشند، به یک ضرورت پژوهشی و صنعتی تبدیل شده است. در این راستا، رویکردهای مختلفی برای تضمین حریم خصوصی پیشنهاد شده‌اند که از حریم خصوصی تفاضلی محلی گرفته تا رمزنگاری تجمیعی امن و مدل‌های ترکیبی را شامل می‌شوند. هر یک از این روش‌ها دارای نقاط قوت و محدودیت‌های خاص خود هستند و انتخاب میان آن‌ها نیازمند موازنه‌ای دقیق میان سه مؤلفهٔ اصلی یعنی کارایی، دقت و سطح امنیت است، چرا که افزایش امنیت معمولاً به افزایش هزینه‌های محاسباتی و کاهش سرعت منجر می‌شود. درواقع، در طراحی سیستم‌های تحلیل فدرال، دستیابی هم‌زمان به این سه هدف یک چالش کلاسیک اما حیاتی محسوب می‌شود که تاکنون هیچ روش واحدی نتوانسته آن را به‌صورت کامل برطرف کند.

حریم خصوصی تفاضلی محلی (Local Differential Privacy) به دلیل سبک‌وزن بودن و نیاز اندک به منابع محاسباتی، گزینه‌ای مناسب برای دستگاه‌های لبه با توان محدود محسوب می‌شود [۱۵]. این روش با افزودن نویز به داده‌ها، حریم خصوصی کاربران را حفظ می‌کند، اما کاهش دقت ناشی از این نویز می‌تواند در کاربردهای حساس مانند سلامت الکترونیک و تحلیل‌های پیش‌بینی دقیق چالش‌برانگیز باشد [۱۱]. در چنین سناریوهایی، طراحی سازوکارهای تنظیم پویا برای کنترل شدت نویز و افزایش دقت آماری به یکی از محورهای اصلی پژوهش تبدیل شده است؛ بنابراین، این روش بیشتر در سناریوهایی نظیر شهر هوشمند یا حمل‌ونقل هوشمند که داده‌ها در مقیاس کلان تولید می‌شوند و دقت مطلق اهمیت کمتری دارد، عملکرد مناسبی از خود نشان می‌دهد [۱۶]. علاوه بر این، توانایی این روش در ادغام مقادیر آماری بزرگ از کاربران متعدد بدون نیاز به ارسال داده‌های خام، باعث افزایش مقیاس‌پذیری سیستم می‌شود. چنین ویژگی‌ای در محیط‌های پویا که دستگاه‌های لبه به‌صورت متناوب به شبکه متصل می‌شوند، یک مزیت قابل‌توجه به شمار می‌رود.

در مقابل، رمزنگاری تجمیعی امن (Secure Aggregation) بالاترین سطح حفاظت از حریم خصوصی را فراهم می‌کند و برای کاربردهایی مانند خدمات مالی و پزشکی که امنیت داده‌ها حیاتی است، مناسب‌ترین گزینه محسوب می‌شود [۷]. این روش با استفاده از پروتکل‌های رمزنگاری چندطرفه و جمع امن، مانع از دسترسی سرور مرکزی به داده‌های خام می‌شود و تنها نتایج آماری کلی را آشکار می‌سازد. بااین‌حال، هزینهٔ محاسباتی و ارتباطی بالای آن در محیط‌های لبه با محدودیت منابع می‌تواند مانع از کارایی مطلوب شود [۱۷]. در محیط‌هایی که زیرساخت ارتباطی و توان پردازشی کافی وجود دارد، این روش بالاترین سطح دقت و امنیت را تضمین می‌کند و به‌ویژه در محیط‌هایی با داده‌های غیر IID و کاربران متعدد، توانایی محافظت از اطلاعات حساس را دارد. همچنین، در سناریوهای چندبخشی مانند بیمارستان‌های توزیع‌شده یا بانک‌های زنجیره‌ای، استفاده از این روش می‌تواند ریسک نشت داده را تقریباً به صفر برساند.

روش‌های ترکیبی (Hybrid Approaches) نیز با ادغام مزایای حریم خصوصی تفاضلی و رمزنگاری، تلاش می‌کنند تا تعادل میان حریم خصوصی، دقت و کارایی را برقرار کنند [۱۸]. این روش‌ها با بهره‌گیری از فشرده‌سازی مدل‌ها، تنظیم پویا میان سطح نویز و میزان رمزنگاری و سازوکارهای تطبیقی، انعطاف‌پذیری بالایی دارند و برای محیط‌های پیچیده و ناهمگن مانند شبکه‌های لبه چندسطحی، صنعت هوشمند، سلامت دیجیتال و تحلیل داده‌های IoT مناسب‌اند [۱۳]. یکی از دستاوردهای مهم در این حوزه، طراحی معماری‌های تطبیقی است که سطح رمزنگاری و نویز را بر اساس نوع داده، حساسیت اطلاعات و شرایط شبکه به‌صورت خودکار تنظیم می‌کنند. باوجود این، پیاده‌سازی عملی آن‌ها نیازمند هماهنگی دقیق میان اجزای سیستم، طراحی پروتکل‌های ارتباطی کارآمد و مدیریت منابع پردازشی است که خود از چالش‌های کلیدی این دسته محسوب می‌شود [۱۹]. علاوه‌برآن، ارزیابی عملکرد این مدل‌ها در محیط‌های واقعی همچنان یک زمینهٔ باز پژوهشی است که نیازمند توسعه ابزارهای شبیه‌سازی و تست‌دهی استاندارد می‌باشد.

به‌طورکلی، انتخاب روش مناسب حفظ حریم خصوصی در تحلیل فدرال نه‌تنها به نوع داده‌ها و محدودیت‌های سخت‌افزاری بستگی دارد، بلکه با اهداف کاربردی، سطح تحمل خطا و میزان حساسیت اطلاعات مرتبط است. پژوهش‌های اخیر تلاش دارند با طراحی الگوریتم‌های تطبیقی و چندلایه، راهکارهایی ارائه دهند که بتوانند در شرایط مختلف، سطح بهینه‌ای از امنیت، دقت و کارایی را تضمین کنند. چنین الگوریتم‌هایی معمولاً از یادگیری تقویتی یا بهینه‌سازی چندهدفه برای تنظیم خودکار

پارامترهای حریم خصوصی بهره می‌برند. این مسئله اهمیت ویژه‌ای در توسعه سیستم‌های تحلیل داده توزیع‌شده در محیط‌های لبه دارد و انتخاب رویکرد صحیح می‌تواند عملکرد کلی سیستم و رضایت کاربران را به طور قابل‌توجهی بهبود دهد. برای مقایسه و مرور ویژگی‌های هر یک از این رویکردها، جدول ۲ تفاوت‌های کلیدی میان آن‌ها را نشان می‌دهد.

جدول ۲. مقایسهٔ روش‌های حفظ حریم خصوصی در تحلیل فدرال لبه

معیار	روش‌های ترکیبی (Hybrid Approaches)	رمزنگاری تجمیعی امن (Secure Aggregation)	حریم خصوصی تفاضلی محلی (Local DP)
سطح حفاظت از حریم خصوصی	و DP قابل‌تنظیم، ترکیب رمزنگاری حفاظت چندلایه‌ای ایجاد می‌کند [۱۸]	بسیار بالا، داده‌ها بدون آشکارشدن در سرور مرکزی [۷] تجمیع می‌شوند	بالا، از طریق افزودن نویز به داده‌ها [۱۵] پیش از ارسال
هزینه محاسباتی در دستگاه لبه	متوسط، بسته به نحوهٔ ترکیب روش‌ها و نوع [۱۳] فشرده‌سازی	بالا، به دلیل پیچیدگی [۱۸] عملیات رمزنگاری	پایین، اما افزودن نویز منجر به کاهش دقت می‌شود [۱۱]
سربار ارتباطی	متوسط، با بهره‌گیری از تکنیک‌های فشرده‌سازی و [۱۳] بهینه‌سازی ارتباطی	بالا، به علت انتقال داده‌های [۱۸] رمزگذاری شده	پایین، زیرا تنها نتایج نویزی ارسال [۱۵] می‌شوند
تحمل در برابر ناهمگنی داده‌ها	قابل‌تنظیم، با استفاده از الگوریتم‌های تطبیقی [۱۳]	بالا، به دلیل تجمیع امن در [۱۳] سطح سرور	محدود، به دلیل تأثیر نویز بر [۱۵] IID داده‌های غیر
قابلیت تفسیرپذیری نتایج	متوسط، بسته به نسبت نویز و سطح رمزنگاری [۱۳]	بالا، به علت تجمیع دقیق [۷] داده‌ها	محدود، به دلیل اعوجاج ناشی از [۱۵] نویز

۵. چالش‌های باز و مسیرهای پژوهشی آینده در تحلیل فدرال لبه

باوجود پیشرفت‌های قابل‌توجه در حوزهٔ تحلیل فدرال و شبکه‌های لبه، همچنان چالش‌های متعددی وجود دارد که مانع تحقق کامل پتانسیل این فناوری می‌شوند. در این بخش، مهم‌ترین چالش‌های باز همراه با مسیرهای پژوهشی آینده برای هر یک موردبررسی قرار می‌گیرد.

۵،۱. ناهمگنی چندبعدی (داده‌ها، دستگاه‌ها و شبکه)

ناهمگنی چندبعدی از مهم‌ترین موانع در سیستم‌های تحلیل فدرال است و شامل تفاوت در توزیع داده‌ها (Statistical Heterogeneity)، تفاوت در توان محاسباتی و انرژی دستگاه‌ها (Systems Heterogeneity) و ناپایداری شبکه‌های ارتباطی (Network Heterogeneity) می‌شود [۱۳]. این ناهمگنی‌ها باعث کاهش کارایی، کندی همگرایی مدل‌ها و افت دقت نهایی می‌گردند.

در مسیر آینده، طراحی الگوریتم‌های تطبیق‌پذیر و غیرهم‌زمان (Asynchronous Adaptive Algorithms) می‌تواند راهکاری مؤثر باشد. چنین الگوریتم‌هایی با سازگاری خودکار نسبت به شرایط متفاوت دستگاه‌ها و شبکه‌ها، عملکرد بهینه‌تری

در محیط‌های ناهمگن ارائه می‌دهند [۱۳]. علاوه بر این، بهره‌گیری از تکنیک‌های فشرده‌سازی مدل و برنامه‌ریزی ارتباطات غیرهم‌زمان می‌تواند سربار ارتباطی را به طور مؤثری کاهش دهد [۲۰]

۵,۲. حفظ حریم خصوصی در طول زمان (مقابله با حملات مبتنی بر توالی)

یکی از چالش‌های نوظهور در تحلیل فدرال، حفظ حریم خصوصی در برابر حملات زمانی (Temporal Attacks) است که از الگوهای تغییرات داده‌ها در بازه‌های زمانی برای استخراج اطلاعات حساس استفاده می‌کنند [۱۱].

راه‌حل پیشنهادی، به‌کارگیری تفاضل حریم خصوصی پویا (Dynamic Differential Privacy) است که میزان نویز افزوده‌شده را متناسب با تغییرات داده و شرایط زمانی تنظیم می‌کند [۷]. افزون بر این، استفاده از مدل‌های تصادفی پویا (Dynamic Stochastic Models)می‌تواند به کاهش خطر حملات مبتنی بر توالی کمک کرده و پایداری حریم خصوصی را در بلندمدت تضمین نماید [۲۱].

۵,۳.تضمین حریم خصوصی در محیط‌های غیرصادقانه(Byzantine Environments)

در محیط‌های فدرال باز و ناامن، احتمال وجود شرکت‌کنندگان مخرب (Byzantine Clients) که داده‌ها یا مدل‌ها را به‌صورت عمدی دست‌کاری می‌کنند، چالشی جدی است. این رفتارها می‌تواند منجر به حملات Model Poisoning شده و دقت مدل کلی را به‌شدت کاهش دهد [۲۲].

پژوهش‌های آینده باید بر توسعهٔ مکانیسم‌های تشویق و اعتمادسنجی (Incentive and Trust Mechanisms) متمرکز شوند. این مکانیسم‌ها با شناسایی و پاداش‌دهی به گره‌های قابل‌اعتماد و حذف مشارکت‌های مشکوک، امنیت سیستم را تقویت می‌کنند [۱۴]. همچنین، استفاده از تجمیع مقاوم (Robust Aggregation Methods) می‌تواند از اثرگذاری داده‌های آلوده جلوگیری نماید [۶]

۵,۴.فقدان معیارهای استاندارد و دیتاست‌های بنچمارک

نبود معیارهای ارزیابی استاندارد و دیتاست‌های بنچمارک عمومی، یکی از موانع مهم در توسعه و ارزیابی دقیق روش‌های تحلیل فدرال است [۱۳]. در غیاب چنین معیارهایی، مقایسهٔ منصفانهٔ میان روش‌ها دشوار می‌شود و روند پیشرفت پژوهش‌ها کند می‌گردد.

پیشنهاد می‌شود یک پلتفرم ارزیابی یکپارچه (Unified Evaluation Platform) طراحی شود که شامل معیارهای مشترک برای دقت، حریم خصوصی، کارایی و مصرف منابع باشد [۱۶]. افزون بر این، توسعهٔ پروتکل‌های بنچمارک‌سازی (Benchmarking Protocols)می‌تواند امکان مقایسهٔ استاندارد میان روش‌های مختلف را فراهم سازد [۱۳].

مرور چالش‌های فوق نشان می‌دهد که تحلیل فدرال در محیط‌های لبه همچنان در مراحل تکامل قرار دارد. طراحی الگوریتم‌های تطبیق‌پذیر، ادغام تفاضل حریم خصوصی پویا، توسعهٔ مکانیسم‌های اعتمادسنجی و ایجاد پلتفرم‌های ارزیابی استاندارد از مسیرهای کلیدی پژوهش‌های آینده به شمار می‌آیند. تمرکز بر این محورها می‌تواند به ارتقای دقت، مقیاس‌پذیری و امنیت در سیستم‌های تحلیل فدرال لبه کمک کرده و راه را برای پیاده‌سازی مؤثر آن‌ها در کاربردهای واقعی هموار سازد [۱۴].

۶. نتیجه‌گیری

این مرور جامع به‌وضوح نشان می‌دهد که تحلیل فدرال (Federated Analytics) در لبه یک پارادایم امیدوارکننده و متحول‌کننده برای استخراج بینش‌های ارزشمند از داده‌های توزیع‌شده، بدون به خطر انداختن حریم خصوصی حساس کاربران است [۳]. ما در این مقاله خلاصه‌ای از دسته‌بندی روش‌های حفظ حریم خصوصی را ارائه کردیم که در آن روش‌های ترکیبی (مانند ترکیب حریم خصوصی تفاضلی و تجمیع ایمن) برتری بالقوه‌ای در بهینه‌سازی بده‌بستان بین کارایی، دقت و حفظ حریم خصوصی در سناریوهای واقعی نشان می‌دهند. اگرچه تکنیک‌هایی مانند حریم خصوصی تفاضلی (DP) و تجمیع ایمن (Secure Aggregation) به بلوغ نسبی در ارائه تضمین‌های خصوصی‌سازی رسیده‌اند [۷]، اما چالش‌های اساسی مانند

ناهمگنی چندبعدی (Systems Heterogeneity & Statistical) و تضمین حریم خصوصی در محیط‌های غیرصادقانه (Byzantine) همچنان پابرجاست و کارایی و قابلیت اطمینان سیستم‌های فدرال را به‌طورجدی تحت‌تأثیر قرار می‌دهد [۱۳]. حرکت به جلو، جامعه پژوهشی باید بر روی طراحی الگوریتم‌های تطبیق‌پذیر غیرهم‌زمان و توسعه مکانیسم‌های استاندارد ارزیابی و اعتمادسنجی متمرکز شود تا پتانسیل کامل این فناوری برای ایجاد یک اکوسیستم داده‌ای همکارانه، کارآمد و امن محقق شود [۱۶، ۱۴].

۷. منابع و مراجع

1. Bao, G. & Guo, P. Federated learning in cloud-edge collaborative architecture: key technologies, applications and challenges. Journal of Cloud Computing 11, 94 (2022). https://doi.org/10.1186/s13677-022-00377-4
2. Elkordy, A. R., Ezzeldin, Y. H., Han, S., Sharma, S., He, C., Mehrotra, S. & Avestimehr, S. Federated analytics: a survey. APSIPA Transactions on Signal and Information Processing 12(1) (2023). https://doi.org/10.48550/arXiv.2302.01326
3. Wang, D., Shi, S., Zhu, Y. & Han, Z. Federated analytics: opportunities and challenges. IEEE Network 36, 151–158 (2022). https://doi.org/10.1109/MNET.101.2100328
4. Zhu, A. Y., Wang, D., Chen, X. & Zhang, Y. Federated analytics: opportunities and challenges. IEEE Network 35(6), 192–199 (2021). https://doi.org/10.1109/MNET.101.2100328
5. Lu, Y. Evaluation of recent developments in federated learning. Applied and Computational Engineering 184, 1–6 (2025). https://doi.org/10.54254/2755-2721/2025.LD25894
6. Haripriya, R., Khare, N., Pandey, M. & Biswas, S. A privacy-enhanced framework for collaborative big data analysis in healthcare using adaptive federated learning aggregation. Journal of Big Data 12, 113 (2025). https://doi.org/10.1186/s40537-025-01169-8
7. Wang, Z., Ji, H., Zhu, Y., Wang, D. & Han, Z. A survey on federated analytics: taxonomy, enabling techniques, applications and open issues. arXiv preprint arXiv:2404.12666 (2024).
8. Vankayalapati, R. K., Sudha Rani, P. R., Valiki, S. & Teja Ganti, V. K. A. Federated edge computing for privacy-preserving analytics in healthcare and IoT systems. International Journal of Computer Trends and Technology 73(1), 80–90 (2025). https://doi.org/10.14445/22312803/IJCTT-V73I1P110
9. Guembe, B., Misra, S. & Azetad, A. Privacy issues, attacks, countermeasures and open problems in federated learning: a survey. Applied Artificial Intelligence 38(1), e2410504 (2024). https://doi.org/10.1080/08839514.2024.2410504
10. Bonawitz, K., Kairouz, P., McMahan, B. & Ramage, D. Federated learning and privacy: building privacy-preserving systems for machine learning and data science on decentralized data. Communications of the ACM 65(4), 90–97 (2022). https://doi.org/10.1145/3500240
11. Xia, Q., Ye, W., Tao, Z., Wu, J. & Li, Q. A survey of federated learning for edge computing: research problems and solutions. High-Confidence Computing 1(1), 100008 (2021). https://doi.org/10.1016/j.hcc.2021.100008
12. Chen, W.-N., Choquette-Choo, C. A., Kairouz, P. & Suresh, A. T. The fundamental price of secure aggregation in differentially private federated learning. arXiv preprint arXiv:2203.03000 (2022).
13. Fang, X., Zeng, Q. & Yang, G. Local differential privacy for human-centered computing. EURASIP Journal on Wireless Communications and Networking 2020, 65 (2020). https://doi.org/10.1186/s13638-020-01675-8

14. Mansouri, M., Önen, M., Ben Jaballah, W. & Conti, M. SoK: secure aggregation based on cryptographic schemes for federated learning. Proceedings on Privacy Enhancing Technologies 2023(1), 140–157 (2023). https://doi.org/10.56553/popets-2023-0009
15. Truex, S., Baracaldo, N., Anwar, A., Steinke, T., Ludwig, H., Zhang, R. & Zhou, Y. A hybrid approach to privacy-preserving federated learning. arXiv preprint arXiv:xxxx.xxxxx (2023).
16. Hu, K., Gong, S., Zhang, Q., Seng, C., Xia, M. & Jiang, S. An overview of implementing security and privacy in federated learning. Artificial Intelligence Review 57, 204 (2024). https://doi.org/10.1007/s10462-024-10846-8
17. Song, J., Luo, J., Lu, R., Xie, S., Chen, B. & Wang, Z. A joint approach to local updating and gradient compression for efficient asynchronous federated learning. arXiv preprint arXiv:2407.05125 (2024).
18. Du, J., Li, S., Chen, X., Chen, S. & Hong, M. Dynamic differential-privacy preserving SGD. arXiv preprint arXiv:2111.00173 (2021). https://doi.org/10.48550/arXiv.2111.00173
19. Fang, M., Cao, X., Jia, J. & Gong, N. Z. Local model poisoning attacks to Byzantine-robust federated learning. arXiv preprint arXiv:1911.11815 (2019). https://doi.org/10.48550/arXiv.1911.11815
20. Chen, W.-N., Choquette-Choo, C. A., & Kairouz, P. Communication efficient federated learning with secure aggregation and differential privacy. OpenReview (PRIML 2021 Workshop) (2021).
21. Elkordy, A. R., Ezzeldin, Y. H., Han, S., Sharma, S., He, C., Mehrotra, S. & Avestimehr, S. Federated analytics: a survey. APSIPA Transactions on Signal and Information Processing 12(1) (2023). https://doi.org/10.48550/arXiv.2302.01326
22. Chen, W.-N., Choquette-Choo, C. A., Kairouz, P., & Suresh, A. T. Communication efficient federated learning with secure aggregation and differential privacy. arXiv preprint arXiv:xxxx.xxxxx (2021).

Federated Analytics in Edge Networks: Privacy-Preserving Insights from Distributed Data

Mohammadsina Pourshabanian[1], Mohammad Mehdi Shirmohammadi[2]

[1] Department of Computer Science, Hamedan Branch, Islamic Azad University, Hamedan, IranPourshabanian@icloud.com

[2] Department of Computer Science, Hamedan Branch, Islamic Azad University, Hamedan, IranMmshirmohammadi@iau.ac.ir

Abstract_

With the rapid expansion of the Internet of Things and edge computing, the volume of distributed data has increased dramatically, highlighting the need to develop privacy-preserving analytical frameworks. This review article provides a comprehensive review of federated analytics in edge networks with a focus on data privacy.

The main motivation for this research is the need to strike a balance between data efficiency and privacy protection in a context where stringent regulations such as the General Data Protection Regulation are in effect. The aim of this paper is to explain and analyze the key privacy-preserving methods in the federated analysis framework, which include differential privacy, secure multi-party computation, and cryptographic aggregation.

The findings of this study show that combining model compression with lightweight encryption is the most effective solution for bandwidth-constrained environments. However, data heterogeneity and high communication overhead remain the most important remaining challenges in the large-scale implementation of federated analysis. Finally, the paper focuses on future research perspectives, which include optimizing the inter-edge node communication, managing statistical and system heterogeneity, and establishing standard metrics for evaluating the performance of federated systems.

Keywords: federated analysis, edge computing, privacy, differential privacy, secure aggregation, data heterogeneity.