Journal of Applied And Basic Machine Intelligence Research (ABMIR)

پژوهش‌های نظری و کاربردی هوش ماشینی

سال چهارم، شماره دوم، پاییز و زمستان ۱۴۰۵، صفحات ۱۱۲ - ۸۹

شناسایی گروه‌های سخت‌قانون در پیام‌رسان تلگرام با استفاده از شبکه عصبی گرافی گراف‌سیج و داده‌های شبیه‌سازی‌شده دارای ویژگی‌های زمانی

مائده عرب بافرانی[۱]، محمدعلی زارع چاهوکی[۲*]

[۱] دانشجوی دکتری، دانشکده مهندسی کامپیوتر، گروه مهندسی نرم‌افزار، دانشگاه یزد، یزد، ایران

[۲] دانشیار، دانشکده مهندسی کامپیوتر، گروه مهندسی نرم‌افزار، دانشگاه یزد، یزد، ایران

مقاله پژوهشی

چکیده

این پژوهش به شناسایی «گروه‌های سخت‌قانون» در پیام‌رسان تلگرام می‌پردازد، گروه‌هایی که کاربران را برای عضویت یا ادامه فعالیت به انجام اقدامات اجباری مانند دعوت از دیگران یا عضویت در کانال‌های خاص ملزم کرده و روند رشد طبیعی گروه‌ها را دست‌کاری می‌کنند. تشخیص این گروه‌ها برای حفظ یکپارچگی پلتفرم و جلوگیری از سوءاستفاده اهمیت بالایی دارد. کمبود داده‌های برچسب‌دار واقعی، ماهیت پویا و زمانی رفتارها و ترکیب هم‌زمان سیگنال‌های ساختاری و رفتاری از چالش‌های این حوزه هستند. برای غلبه بر این چالش‌ها، مجموعه‌داده‌ای شبیه‌سازی‌شده در مقیاس بزرگ شامل ۱۰۰٬۰۰۰ کاربر، ۱۰٬۰۰۰ گروه و ۱۲ تصویر زمانی تولید شد که ویژگی‌های ایستا و زمانی کاربران و گروه‌ها را در برمی‌گیرد. برچسب‌ها با استفاده از یک امتیاز ریسک مبتنی بر تحلیل مؤلفه‌های اصلی تعیین و یک گراف دوبخشی کاربر–گروه ساخته شد. سپس یک مدل شبکه عصبی گرافی مبتنی بر گراف‌سیج با پروتکل ارزیابی کاملاً استقرایی، تابع زیان آنتروپی متقاطع دودویی وزن‌دار و بهینه‌سازی آستانه تصمیم آموزش داده و در کنار شش مدل پایه ارزیابی گردید. مدل پیشنهادی بر روی مجموعه آزمون مستقل به صحت ۹۸/۲ درصد، دقت کلاس مثبت ۹۷/۲ درصد، فراخوانی ۹۳/۷ درصد، امتیاز F1 برابر ۹۵/٤ درصد، AUC برابر ۹۹/۸۷ درصد و میانگین دقت برابر ۹۹/۵۱ درصد دست‌یافت و در تمام بذرها پایدار بود. نتایج نشان داد به دلیل وابستگی ساختاری برچسب‌ها به ویژگی‌های ایستای گروه، مدل‌های جدولی با هزینه محاسباتی کمتر عملکرد اندکی بهتر دارند، درحالی‌که گراف‌سیج در گراف‌های بزرگ و پردرجه مقاوم می‌ماند. این یافته‌ها صرفاً بر روی داده‌های شبیه‌سازی‌شده تأیید شده‌اند و تعمیم آن‌ها به داده‌های واقعی تلگرام، به دلیل نبود داده‌های برچسب‌دار واقعی، همچنان چالشی باز است.

تاریخ دریافت:
۱۴۰۵/۰۱/۱۲
تاریخ پذیرش:
۱۴۰۵/۰۵/۲۷

کلیدواژه‌ها:
گروه‌های سخت‌قانون، پیام‌رسان تلگرام، شبکه‌های عصبی گرافی، گراف‌سیج، داده‌های شبیه‌سازی‌شده، ویژگی‌های زمانی، تشخیص رفتار اجباری، یادگیری عمیق گراف‌محور

نویسنده‌مسئول:
chahooki@yazd.ac.ir

: 10.22034/ABMIR.2026.24432.1231

E-ISSN: 2821-2037

۱- مقدمه

شبکه‌های اجتماعی به یکی از پیچیده‌ترین بسترهای تعامل انسانی تبدیل شده‌اند، ساختاری پویا که در آن گروه‌ها، کاربران و الگوهای رفتاری به‌صورت پیوسته شکل می‌گیرند و تغییر می‌کنند. در میان ساختارهای موجود در شبکه‌های اجتماعی، بخشی از گروه‌ها به‌صورت طبیعی رشد می‌کنند و تعداد اعضای آن‌ها به‌مرور زمان افزایش می‌یابد، درحالی‌که برخی دیگر از الگوهای رفتاری غیرعادی، هدایت‌شده و یا اصطلاحاً اجباری پیروی می‌کنند [۱]، [۲]، [۳]، [٤]. شناسایی این نوع ساختارها از اهمیت بالایی در تحلیل دقیق تعاملات اجتماعی، قابل‌اطمینان بودن پیام‌رسان، شناسایی فعالیت‌های هماهنگ شده و غیرواقعی و پیشگیری از سوءاستفاده‌های هدفمند برخوردار است [۱]، [۲]. بااین‌حال، پیچیدگی ساختاری شبکه‌های اجتماعی، وابستگی زمانی رفتارها و وجود نویز و عدم قطعیت، تشخیص چنین الگوهایی را به یک مسئله چالش‌برانگیز در حوزه یادگیری ماشین و تحلیل گراف تبدیل کرده است [۲]، [٤].

۱-۲ رویکرد پیشنهادی و چارچوب تحقیق

در این پژوهش، چارچوبی برای شبیه‌سازی و یادگیری مبتنی بر گراف برای بررسی رفتارهای اجباری در گروه‌های اجتماعی ارائه می‌شود. ابتدا یک مدل تولید داده مصنوعی طراحی‌شده که ویژگی‌های آماری و رفتاری کاربران و گروه‌ها را با استفاده از توزیع‌های واقع‌گرایانه و شاخص‌های تعامل اجتماعی شبیه‌سازی می‌کند [۳]، [۵]، [٦]، [۷]. سپس، با ساخت یک گراف دوبخشی شامل کاربران و گروه‌ها، ساختار ارتباطی شبکه اجتماعی بازسازی می‌شود. برای تحلیل این ساختار، یک مدل یادگیری عمیق مبتنی بر شبکه‌های عصبی گرافی از نوع گراف‌سیج با مکانیزم توازن کلاس و آموزش بانظارت توسعه داده‌شده است [۸]. این روش امکان استخراج الگوهای ساختاری، رفتاری و تعاملات پنهان در گراف را فراهم می‌سازد [۳]، [۵]، [٦]، [۷].

۱-۳ هدف تحقیق و کاربردپذیری

هدف اصلی این پژوهش، ارزیابی دقت و کارایی روشی گراف‌محور برای شناسایی رفتارهای اجباری در پیام‌رسان‌ها و

۱-۱ شکاف علمی و چالش‌ها

روش‌های سنتی مبتنی بر ویژگی‌های آماری یا تحلیل ایستا، توانایی محدودی در استخراج الگوهای پیچیده رفتاری دارند. علت اصلی این محدودیت این است که رفتارهای اجباری اغلب در تعامل هم‌زمان میان ساختار گراف و روابط همسایگی کاربران، ویژگی‌ها و الگوهای گروهی و تغییرات زمانی رفتارها رخ می‌دهند [۲]. علاوه بر این، در بسیاری از کاربردهای واقعی، دسترسی به داده‌های برچسب‌دار و قابل‌اعتماد بسیار محدود است. نبود داده‌های واقعی با برچسب معتبر برای رفتارهای غیرعادی، یکی از موانع اصلی توسعه مدل‌های یادگیری عمیق در این حوزه محسوب می‌شود. این مسئله ضرورت استفاده از داده‌های مصنوعی کنترل‌شده را برجسته می‌سازد، داده‌هایی که بتوانند ویژگی‌های آماری، ساختاری و رفتاری شبکه‌های اجتماعی موجود را شبیه‌سازی کرده و امکان ارزیابی اصولی روش‌های یادگیری را فراهم سازند [۳]، [۵]، [٦]، [۷].

سپس به‌کارگیری آن در داده‌های واقعی است. استفاده از داده مصنوعی، نه تنها محدودیت دسترسی به داده‌های واقعی را برطرف می‌کند، بلکه امکان کنترل دقیق پارامترهای تولید، تحلیل حساسیت و بررسی اصولی رفتار مدل را نیز فراهم می‌آورد. درنهایت، این پژوهش تلاش می‌کند با تحلیل نتایج، پیشنهادهایی برای توسعه مجموعه داده‌های واقعی ارائه دهد و مشخص کند چه نوع ویژگی‌ها و اطلاعاتی باید به داده‌های اجتماعی افزوده شوند تا تشخیص رفتارهای ساختاری و ناهنجار در مقیاس واقعی امکان‌پذیر گردد.

۲- پیشینه پژوهش

مرور مطالعات پیشین نشان می‌دهد که باوجود پیشرفت‌های قابل‌توجه در هر یک از حوزه‌های تولید گراف‌های مصنوعی و شبیه‌سازی شبکه‌های اجتماعی، تشخیص رفتارهای غیرعادی و هماهنگ، کاربرد شبکه‌های عصبی گرافی در تحلیل سامانه‌های اجتماعی و شناسایی ناهنجاری‌ها و الگوهای ساختاری در گراف‌ها، ترکیب شبیه‌سازی اجتماعی با یادگیری مبتنی بر گراف برای تحلیل

رفتارهای غیرعادی در سطح گروه‌های اجتماعی، همچنان به‌عنوان یک شکاف پژوهشی مهم مطرح است. بخش عمده پژوهش‌های موجود یا بر داده‌های واقعی و در سطح کاربران تمرکز داشته‌اند، یا صرفاً یکی از ابعاد ساختاری، رفتاری یا زمانی را مدنظر قرار داده‌اند. ازاین‌رو، توسعه چارچوب‌هایی که بتوانند به‌صورت هم‌زمان از داده‌های مصنوعی واقع‌گرایانه، ویژگی‌های رفتاری و زمانی، و یادگیری مبتنی بر گراف برای تحلیل رفتارهای گروهی بهره بگیرند، همچنان نیازمند توجه بیشتر است.

۲-۱ تولید گراف مصنوعی و شبیه‌سازی شبکه‌های اجتماعی

تولید گراف‌های مصنوعی به‌عنوان راهکاری برای غلبه بر محدودیت‌های دسترسی به داده‌های واقعی شبکه‌های اجتماعی، در سال‌های اخیر نقش مهمی در پژوهش‌ها ایفا کرده است. این رویکرد علاوه بر فراهم کردن امکان کنترل دقیق پارامترهای مدل، ریسک‌های اخلاقی مرتبط با حریم خصوصی را نیز کاهش می‌دهد. یکی از کارهای بنیادی در این حوزه، مدل NetGAN [۹] است که با استفاده از مسیرهای تصادفی و مدل‌های مولد تخاصمی، توزیع ساختاری گراف را یاد می‌گیرد و گراف‌هایی با ویژگی‌های توپولوژیکی مشابه نمونه‌های واقعی تولید می‌کند. این روش در حفظ ساختارهای مرتبه‌بالا مانند خوشه‌بندی و الگوهای اتصال موفق عمل می‌کند، اما فاقد مدل‌سازی ویژگی‌های گره و تغییرات زمانی است؛ درنتیجه برای تحلیل رفتارهای اجتماعی پیچیده محدودیت دارد.

مدل GraphRNN [۱۰] با استفاده از شبکه‌های بازگشتی، چارچوبی ترتیبی برای تولید گراف معرفی کرد که بازسازی گراف‌های پیچیده‌تر را ممکن می‌سازد. این مدل در بازتولید ساختارهای پویا عملکرد بهتری نشان می‌دهد، اما همچنان به مفروضات ساده‌سازی‌شده متکی است و قادر به پوشش کامل پیچیدگی رفتار انسانی نیست.

مطالعات جدیدتر نشان داده‌اند که مدل‌های مبتنی بر توجه گرافی می‌توانند ساختارهای اجتماعی واقعی را با دقت بالاتری بازسازی کنند و شباهت آماری بیشتری با شبکه‌های واقعی نظیر فیسبوک و گیت‌هاب داشته باشند [۵]،[۱۶]، درحالی‌که مدل‌های کلاسیک قاعده‌محور مانند R-MAT در حفظ پویایی‌های اجتماعی عملکرد ضعیف‌تری نشان می‌دهند. بااین‌حال، هزینه محاسباتی بالاتر این مدل‌ها و وابستگی آن‌ها به داده‌های آموزشی واقعی، چالش‌هایی برای کاربرد در مقیاس بزرگ ایجاد می‌کند.

در سوی دیگر، پژوهش‌های مبتنی بر مدل‌های عامل‌محور بر تولید جمعیت مصنوعی همراه با شبکه‌های اجتماعی تمرکز دارند [۶]، [۷]. این دسته از مطالعات، جمعیت‌های مصنوعی بزرگ با ساختارهای اجتماعی مشخص تولید می‌کنند تا شبیه‌سازی تعاملات انسانی، مانند انتشار بیماری یا تصمیم‌گیری اجتماعی، امکان‌پذیر شود. نتایج این پژوهش‌ها نشان می‌دهد که در صورت اعتبار آماری مناسب، داده‌های مصنوعی می‌توانند جایگزین قابل‌قبولی برای داده‌های واقعی باشند. مزیت اصلی این رویکرد، توانایی مدل‌سازی رفتار اجتماعی در مقیاس بزرگ است، اما ساده‌سازی بیش‌ازحد رفتار انسانی و وابستگی به قوانین از پیش تعریف‌شده، از محدودیت‌های آن به شمار می‌رود. به‌طور مشابه، چارچوب‌های عامل‌محور جدید نیز نشان داده‌اند که بازسازی پویایی تعاملات اجتماعی می‌تواند برای تحلیل سامانه‌های اجتماعی پیچیده سودمند باشد [۳].

علاوه بر این، مطالعات اخیر تأکید کرده‌اند که کیفیت داده‌های مصنوعی صرفاً به حفظ ویژگی‌های توپولوژیکی محدود نمی‌شود، بلکه میزان شباهت توزیع ویژگی‌های رفتاری و پویایی‌های زمانی نیز در اعتبار این داده‌ها نقش اساسی دارد. درواقع، ارزیابی داده‌های مصنوعی باید علاوه بر معیارهای ساختاری، بر اساس قابلیت استفاده در وظایف پایین‌دستی و میزان انطباق آماری با داده‌های واقعی صورت گیرد [۵]، [۲۸].

۲-۲ تشخیص رفتارهای غیرعادی و اجباری در شبکه‌های اجتماعی

تشخیص رفتارهای هماهنگ و فعالیت‌های از پیش تعیین‌شده، یکی از مسائل مهم در حوزه امنیت و تحلیل شبکه‌های اجتماعی به شمار می‌رود. در پژوهش‌های اولیه نشان داده شد که رفتارهای مصنوعی، ازجمله فعالیت بات‌ها، معمولاً در الگوهای زمانی، تراکم تعاملات و هماهنگی رفتاری پنهان می‌شوند. این مطالعات تأکید می‌کنند که اتکا بر ویژگی‌های آماری ساده مانند نرخ فعالیت برای

شناسایی چنین رفتارهایی کافی نیست و باید ساختار تعاملات نیز در تحلیل لحاظ شود. بااین‌حال، محدودیت اصلی این رویکردها در ناتوانی آن‌ها در مدل‌سازی روابط غیرخطی و پیچیده رفتاری است.

روش CopyCatch که در [۲] ارائه شد، یکی از نخستین سامانه‌های مبتنی بر گراف برای کشف فعالیت‌های هماهنگ محسوب می‌شود. این روش با شناسایی زیرگراف‌های متراکم از رفتارهای مشابه، الگوهای سازمان‌یافته را آشکار می‌کند. نقطه قوت CopyCatch در تحلیل ساختاری شبکه است، اما فاقد مدل‌سازی پویا بوده و به داده‌های واقعی وابسته است. مطالعات بعدی نشان داده‌اند که رفتارهای هماهنگ معمولاً در تعامل هم‌زمان میان ساختار شبکه، بعد زمانی و ویژگی‌های رفتاری نمایان می‌شوند؛ موضوعی که به‌کارگیری مدل‌های پیشرفته‌تر را ضروری می‌سازد.

در پژوهش [۱۱] هماهنگی رفتاری بر اساس انحراف الگوهای اشتراک‌گذاری از رفتار متعارف کاربران موردبررسی قرار گرفت. نتایج این مطالعه نشان داد که رویکردهای مبتنی بر ناهنجاری می‌توانند وابستگی به آستانه‌های ثابت را کاهش داده و کمپین‌های هماهنگ را با انعطاف‌پذیری بیشتری شناسایی کنند. این یافته‌ها بیانگر آن است که تشخیص رفتارهای هماهنگ، بیش از آنکه به یک معیار منفرد متکی باشد، نیازمند ترکیب چندین سیگنال رفتاری و ساختاری است.

مطالعات انجام‌شده در سایر شبکه‌های اجتماعی نیز بر همین موضوع توجه دارند. در پژوهش [۱۲] نشان داده شد که رفتارهای هماهنگ در فیسبوک صرفاً محدود به عملیات نفوذ سازمان‌یافته نیستند و می‌توانند در قالب شبکه‌های رسانه‌ای، تبلیغاتی و سیاسی نیز ظاهر شوند.

همچنین، مطالعه [۱۳] نشان داد که حساب‌های هماهنگ در آبشارهای انتشار اطلاعات، رفتار متفاوتی نسبت به کاربران عادی دارند و نقش پررنگ‌تری در گسترش محتوا ایفا می‌کنند.

افزون بر این، پژوهش [۱٤] بیان کرد که در محیط‌های ویدئومحور، نشانه‌هایی نظیر هم‌زمانی انتشار، بازاستفاده از محتوای چندرسانه‌ای و شباهت الگوهای تولید محتوا می‌توانند در تشخیص رفتارهای هماهنگ مؤثر باشند.

همچنین، مطالعات میان‌پلتفرمی نیز وجود هماهنگی‌های مشترک میان شبکه‌هایی نظیر X، فیسبوک و تلگرام را تأیید کرده‌اند [۱۵].

برای نمونه، پژوهشگران در مطالعات مبتنی بر مدل‌های عامل‌محور و اتوماتای رفتاری نشان داده‌اند که داده‌های مصنوعی، در صورت طراحی مناسب، می‌توانند ساختار شبکه و الگوهای رفتاری کاربران را با درجه‌ای قابل‌قبول از واقع‌گرایی بازسازی کنند [۳]،[۷]. مزیت اصلی این رویکرد، کنترل کامل پارامترها و امکان بررسی سناریوهای غیرقابل‌مشاهده در داده‌های واقعی است؛ بااین‌حال، خطر ساده‌سازی بیش‌ازحد رفتار انسانی همچنان وجود دارد.

۲-۳ شبکه‌های عصبی گرافی در تحلیل سامانه‌های اجتماعی

شبکه‌های عصبی گرافی به دلیل توانایی در استخراج هم‌زمان الگوهای ساختاری و رفتاری، پیشرفت‌های چشمگیری در تحلیل شبکه‌های اجتماعی ایجاد کرده‌اند. مدل گراف‌سیج در [۸] با رویکرد یادگیری القایی، امکان پردازش گراف‌های بزرگ و پویا را فراهم می‌کند. این مدل با تجمیع اطلاعات همسایگان، ویژگی‌های محلی گره‌ها را به‌طور مؤثر یاد می‌گیرد، اما فاقد سازوکار توجه برای اولویت‌بندی میزان اهمیت همسایگان است.

در مقابل، مدل شبکه توجه بر گراف[1] در [۱٦] با معرفی مکانیزم توجه، نشان داد که همه همسایگان سهم یکسانی در یادگیری نمایش گره ندارند. این رویکرد دقت مدل را در بسیاری از کاربردها افزایش داد، اما در گراف‌های نویزی یا با تراکم بالا ممکن است با ناپایداری مواجه شود.

پژوهش‌های جدیدتر نیز نشان داده‌اند که استفاده از ساختارهای ناهمگن و بازسازی گراف می‌تواند در شناسایی الگوهای انتشار غیرعادی و شایعات مؤثر باشد [۱۷].

در حوزه کاربردهای اجتماعی، مدل GNN-CARE برای تشخیص تقلب در شبکه‌های اجتماعی توسعه یافت [٤]. این مدل با ترکیب اطلاعات ساختاری گراف و ویژگی‌های رفتاری کاربران، توانایی شناسایی الگوهای پنهان را بهبود می‌بخشد. نتایج گزارش‌شده نشان

[1] Graph Attention Network

می‌دهد که این مدل در داده‌های واقعی عملکرد مطلوبی دارد، اما وابستگی آن به داده‌های برچسب‌خورده، محدودیت‌هایی برای کاربرد در سناریوهای دارای کمبود داده معتبر ایجاد می‌کند.

علاوه بر این، مطالعات اخیر بر توسعه معماری‌های فضایی–زمانی و مقاوم در برابر عدم توازن کلاس متمرکزشده‌اند. پژوهش [۱۸] نشان داد که مدل‌سازی هم‌زمان روابط ساختاری و تغییرات زمانی می‌تواند عملکرد تشخیص گروه‌های متقلب را بهبود بخشد.

همچنین، مدل BalancerGNN باهدف ارتقای عملکرد شبکه‌های عصبی گرافی درمجموعه‌داده‌های نامتوازن ارائه‌شده است [۱۹].

از سوی دیگر، پژوهش‌هایی نظیر [۲۰] و [۲۱] اهمیت تلفیق بعد زمانی با ساختار گراف را در تشخیص الگوهای پیچیده رفتاری برجسته ساخته‌اند. بااین‌وجود، بخش عمده این مطالعات بر تشخیص تقلب در سطح کاربران، تراکنش‌ها یا انتشار محتوا متمرکز بوده و توجه کمتری به تحلیل ناهنجاری در سطح گروه‌های شبکه‌های اجتماعی، به‌ویژه اجتماعات تلگرامی، داشته‌اند.

۲-٤ شکاف پژوهشی

مرور مطالعات پیشین نشان می‌دهد که اگرچه پیشرفت‌های قابل‌توجهی درزمینه تولید داده‌های مصنوعی، تشخیص رفتارهای هماهنگ و بهره‌گیری از شبکه‌های عصبی گرافی حاصل‌شده است، اما همچنان شکاف‌هایی در این حوزه وجود دارد. بخش عمده پژوهش‌های موجود یا مبتنی بر داده‌های واقعی و در سطح کاربران انجام‌شده‌اند، یا تنها یکی از ابعاد ساختاری، رفتاری و زمانی را موردتوجه قرار داده‌اند. همچنین، پژوهش‌های مرتبط با رفتارهای هماهنگ عمدتاً بر پلتفرم‌هایی نظیر فیسبوک، X و تیک‌تاک متمرکز بوده و مطالعات محدودی به تحلیل ناهنجاری در سطح گروه‌های تلگرامی پرداخته‌اند. ازاین‌رو، توسعه چارچوبی مبتنی بر داده‌های مصنوعی که بتواند به‌صورت هم‌زمان ویژگی‌های ایستا، رفتاری و زمانی گروه‌ها را مدل‌سازی کرده و از قابلیت‌های یادگیری مبتنی بر گراف برای طبقه‌بندی الگوهای غیرعادی بهره گیرد، می‌تواند بخشی از این شکاف پژوهشی را پوشش دهد.

۳- روش پیشنهادی

این بخش روش‌شناسی پیاده‌سازی کامل سیستم پیشنهادی را تشریح می‌کند، از ساخت یک معیار[1] مصنوعی کنترل‌شده تا طراحی، آموزش و ارزیابی استقرایی شبکه عصبی گرافی. ما ابتدا مسئله یادگیری را صورت‌بندی می‌کنیم (بخش ۳–۱) و نمای کلی پایپ‌لاین دومرحله‌ای را ارائه می‌دهیم (بخش ۳–۲). سپس الگوریتم‌های تولید داده‌های مصنوعی و توجیه‌های آماری آن‌ها (بخش ۳–۳)، نمایش گراف و کدگذاری ویژگی‌ها (بخش ۳–٤)، پروتکل استقرایی بدون نشت اطلاعات (بخش ۳–۵)، معماری‌های مدل شامل آشکارساز پیشنهادی گراف‌سیج و تمام رقبای آن (بخش ۳–٦) و درنهایت الگوریتم‌های آموزش، بهینه‌سازی و ارزیابی (بخش‌های ۳–۷ تا ۳–۸) را به‌تفصیل بیان می‌کنیم. این طراحی از رویه‌های تثبیت‌شده در یادگیری نمایش گراف [۸]، [۲۲]، [۲۳] و تشخیص ناهنجاری و تقلب مبتنی بر گراف [٤]، [۲٤]، [۲۵] پیروی می‌کند.

۳–۱ صورت‌بندی مسئله و نمادگذاری

در این پژوهش یک پلتفرم آنلاین به‌عنوان یک گراف ناهمگن و در حال تکامل زمانی مدل‌سازی شده است. فرض بر آن بوده است که $U = \{u_1,\dots,u_n\}$ مجموعه n کاربر و $G = \{g_1,\dots,g_m\}$ مجموعه m گروه (جامعه) باشد. عضویت در گام زمانی گسسته t یک رابطه دوبخشی $E_B^t \subseteq U \times G$ است، که در آن $(u,g) \in E_B^t$ اگر و تنها اگر کاربر u در تصویر زمانی t، عضو گروه g باشد. از رابطه عضویت، یک رابطه فرافکنی[2] گروه–گروه $E_P \subseteq G \times G$ القا می‌شود که یال‌های آن گروه‌هایی را که کسر نسبتاً بزرگی از اعضای مشترک دارند به هم متصل می‌کند. هر گروه g دارای یک بردار ویژگی ایستای $x_g^{stat} \in \mathbb{R}^{d_s}$ و برای هر تصویر زمانی $t \in \{0,\dots,T-1\}$ دارای یک بردار ویژگی زمانی $x_g^{temp,(t)} \in \mathbb{R}^{d_t}$ است؛ هر کاربر نیز دارای یک بردار ویژگی $x_u \in \mathbb{R}^{d_u}$ است. با تعداد کاربر ۱۰۰ هزار، تعداد گروه ۱۰ هزار و تعداد تصویر زمانی برابر با ۱۲، مجموعه گره‌های ترکیبی دارای $|V| = n + m$ برابر با ۱۱۰ هزار عنصر است.

[1] Benchmark

[2] Projection

جدول (۱): فهرست نمادها

نماد	تعریف
$\mathbf{U} = \{\mathbf{u_1}, \dots, \mathbf{u_n}\}$	مجموعه n کاربر پلتفرم
$\mathbf{G} = \{\mathbf{g_1}, \dots, \mathbf{g_m}\}$	مجموعه m گروه (جامعه)
$\mathbf{t}$	گام زمانی گسسته (شماره تصویر زمانی)، $t \in \{1,\dots,T-1\}$
$\mathbf{E_B^t} \subseteq \mathbf{U} \times \mathbf{G}$	رابطه عضویت دوبخشی کاربر–گروه در تصویر زمانی t
$\mathbf{E_P} \subseteq \mathbf{G} \times \mathbf{G}$	رابطه فرافکنی گروه–گروه، برخاسته از عضویت مشترک
$\mathbf{x_g^{stat}} \in \mathbb{R}^{\mathbf{d_s}}$	بردار ویژگی ایستای گروه g
$\mathbf{x_g^{temp,(t)}} \in \mathbb{R}^{\mathbf{d_t}}$	بردار ویژگی زمانی گروه g در تصویر زمانی t
$\mathbf{x_u} \in \mathbb{R}^{\mathbf{d_u}}$	بردار ویژگی کاربر u
$\mathbf{n, m, T}$	به‌ترتیب تعداد کاربران، تعداد گروه‌ها و تعداد تصاویر زمانی
$\mathbf{\|V\| = n + m}$	تعداد کل گره‌های گراف ترکیبی
$\mathbf{f_\theta: G \to \{0,1\}}$	تابع طبقه‌بند گروه، پارامترمند با θ
$\mathbf{y_g} \in \{\mathbf{0,1}\}$	برچسب گروه
$\boldsymbol{\pi}$	نرخ شیوع هدف گروه‌های سخت‌قانون
$\boldsymbol{\tau}$	آستانه صدکی امتیاز ریسک برای برچسب‌گذاری
$\mathbf{R_1, \dots, R_5}$	پنج مؤلفه خام امتیاز ریسک پیش از استانداردسازی
$\mathbf{w}$	بردار وزن مشتق‌شده از بارهای مؤلفه اصلی اول (PC1)

وظیفه موردنظر، طبقه‌بندی دودویی گره بر روی گره‌های گروه است: یادگیری $f_\theta: G \to \{0,1\}$ که هر گروه را به $y_g \in \{0 = Normal, 1 = Law\text{-}Hard\}$ نگاشت می‌کند. یک گروه Law–Hard یا همان سخت‌قانون گروهی است که پروفایل رفتاری آن سکوت بالا، مشارکت پایین، فعالیت هدایت‌شده توسط ربات‌ها و رشد ناهنجار، آن را به‌عنوان یک گروه پرخطر و نیازمند نظارت سخت‌گیرانه‌تر نشان می‌دهد. فرضیه مرکزی این است که انتشار اطلاعات در سراسر رابطه دوبخشی کاربر–گروه و رابطه فرافکنی گروه–گروه، تخمین ریسک بهتری نسبت به در نظر گرفتن هر گروه به‌عنوان یک بردار ویژگی مستقل به دست می‌دهد، که با سوگیری استقرایی رابطه‌ای شبکه‌های عصبی مبتنی بر تبادل پیام[1] همخوانی دارد [۸]، [۲۶]، [۲۳].

۳-۲ نمای کلی سیستم

سیستم پیشنهادی یک پایپ‌لاین دومرحله‌ای است. در مرحله اول، یک مولد[2] یک محیط مصنوعی خودسازگار از کاربران، گروه‌ها، عضویت‌ها در طی ۱۲ تصویر زمانی، یک فرافکنی گراف، آمارهای توصیفی و تقسیم‌بندی‌های لایه‌بندی‌شده[3] برای آموزش/اعتبارسنجی/آزمون تولید می‌کند. در مرحله دوم، یک موتور آموزش و ارزیابی این مصنوعات را دریافت کرده، یک شیء گراف Geometric-PyTorch [۲۷] می‌سازد، مدل پیشنهادی گراف‌سیج را به‌همراه شش رقیب آموزش می‌دهد و یک مجموعه ارزیابی کامل (معیارها، منحنی‌های یادگیری، مطالعه حذف، پایداری چندبذری[4]، آزمون معناداری آماری، تحلیل خطا، تشخیص‌های اهمیت ویژگی و اعتبارسنجی توزیع و زمان‌سنجی اجرا) ارائه می‌دهد.

الگوریتم (۱): تولید سرتاسری داده‌های مصنوعی[5]

```
Input: n users, m groups, T snapshots, C communities, target prevalence π=0.20, seed=42
Output: users, groups, memberships{0..T}, temporal features, projection edges, splits
1: seed all RNGs(42)
2: community_id ← SBM partition of users into C blocks (p_intra=0.15, p_inter=0.005)
3: for each user u: type_u ~ Cat(ACTIVE.45, PASSIVE.52, BOT.03)
invite_rate,reply_rate,activity_power ~ type-specific Exp/Beta/Uniform laws
4: for each group g: sample size, activity_ratio, reply_count, age, ... (§3.3.3)
derive silent_ratio, engagement_level, interaction_anomaly, flags
5: R1..R5 ← five raw risk components; standardise; fit 1-component PCA
w ← |PC1 loadings| / Σ|loadings|; risk(g) ← clip(Σ w_i R_i, 0,1)
6: τ ← Percentile_{100(1−π)}(risk); label_g ← 1[risk(g) ≥ τ]
7: for each group g: sample members ∝ homophily weights (Algorithm 2)
8: for t = 0..T−1: AR(1) update of series + burst injection + dynamic join/leave (Algorithm 3)

9: build bipartite graph; build group projection via inverted index (Algorithm 4)
10: drop leakage columns R1..R5, risk_score
11: stratified 70/15/15 split preserving prevalence π
12: return all artefacts
```

[1] Message-Passing
[2] Generator
[3] Stratified
[4] Multi-Seed
[5] End-to-End Synthetic Data Generation

استفاده از یک معیار کاملاً مصنوعی یک انتخاب آگاهانه است: هیچ مجموعه داده واقعی برچسب‌داری از گروه‌های تعدیل‌شده در دسترس نبود و تولید مصنوعی کنترل دقیقی بر فرآیند تولید، دسترسی به واقعیت مبنا و تکرارپذیری کامل را فراهم می‌کند [۲۸]. پایپ‌لاین به‌صورت سراسری با بذر برابر با ۴۲ مقداردهی اولیه شده است. الگوریتم ۱ رویه تولید سرتاسری را که در بخش ۳-۳ بسط داده‌شده است، خلاصه می‌کند.

جدول (۲): مقیاس مجموعه داده‌های مصنوعی

کلاس مثبت	پنجره‌های زمانی	یال‌های گراف(میلیون)	گروه‌ها	کاربران
٪۲۰~	۱۲	۷/۶۴	۱۰۰۰۰	۱۰۰۰۰۰

۳-۳ پایپ‌لاین تولید داده‌های مصنوعی

مولد، چهار ویژگی از اکوسیستم‌های آنلاین واقعی را که محرک یادگیری رابطه‌ای هستند بازتولید می‌کند: ساختار جامعه پنهان، رفتار ناهمگن کاربران شامل حساب‌های خودکار، شکل‌گیری عضویت همسان‌پسندانه[1] و پویایی‌های زمانی پایدار اما دارای فوران[2]. هر مؤلفه با منطق آماری، فرمول‌ها و الگوریتم مربوطه در ادامه شرح داده‌شده است.

۳-۳-۱ ساختار جامعه پنهان (مدل بلوک تصادفی)

کاربران به $c = ۸$ جامعه پنهان با اندازه تقریباً مساوی افراز می‌شوند. ستون فقرات جامعه از یک مدل بلوک تصادفی [۲۹] پیروی می‌کند، که در آن احتمال وجود پیوند بین کاربران i و j تنها به عضویت آن‌ها در بلوک‌های $b(i)$ و $b(j)$ بستگی دارد:

$$\begin{aligned} &((i,j) \in E) = B_{b(i),b(j)}, \\ &B_{c,c} = p_{intra} = 0.15, \\ &B_{c,c'} = p_{inter} = 0.005 \quad (c \neq c') \end{aligned} \qquad (۱)$$

ازآنجاکه $p_{inter} \ll p_{intra}$ است، پیوندها در درون جوامع متمرکز می‌شوند و توپولوژی همسو و ماژولار مشخصه سیستم‌های اجتماعی را بازتولید می‌کنند [۲۹]. همچنین به هر کاربر یک ترجیح موضوعی اختصاص داده می‌شود که تا حدودی با جامعه او هم‌راستا است، تا شکل‌گیری عضویت به‌جای یکنواخت بودن، به‌طور واقع‌گرایانه‌ای همسان‌پسندانه باشد.

۳-۳-۲ مدل‌سازی رفتاری کاربر

هر کاربر یکی از سه گونه رفتاری است که نرخ‌های آن‌ها از توزیع‌های مختص به هرگونه نمونه‌برداری می‌شود و ناهمگنی رفتاری واقعی را تولید می‌کند. نکته مهم این است که ربات‌ها به‌شدت دعوت می‌کنند اما تقریباً هرگز پاسخ نمی‌دهند، که این امضای خودکار و هرزگونه[3] مستندشده در ادبیات ربات‌های اجتماعی است [۱]. به‌طور رسمی، با $t = type(u)$ روابط زیر تعریف می‌شوند:

$$\begin{aligned} &invite_{rate} \sim Exp(\lambda_t), \\ &reply_{rate} \sim Beta(\alpha_t, \beta_t), \\ &activity_{power} \sim Beta/Uniform_t, \end{aligned} \qquad (۲)$$

جدول (۳): گونه‌های رفتاری و توزیع‌های مولد آن‌ها

گونه کاربر	احتمال	$invite_{rate}$	$reply_{rate}$	$activity_{power}$
ACTIVE	۰/۴۵	Exp(۰/۵)	Beta(۵/۳) — high	Beta(۴/۲) — high
PASSIVE	۰/۵۲	Exp(۰/۱)	Beta(۱/۵) — low	Beta(۱/۵) — low
BOT	۰/۰۳	Exp(۵/۰) — very high	Beta(۱/۲۰) — ≈ ۰	Uniform(۰/۸, ۱/۰)

نرخ‌های نمایی، رفتار دعوت دارای دنباله سنگین[4] را به تصویر می‌کشند، درحالی‌که قوانین بتا در بازه [۰,۱] انتخاب طبیعی برای نسبت‌های محدود مانند تمایل به پاسخ و فعالیت هستند.

۳-۳-۳ مهندسی ویژگی‌های گروه

هر گروه یک بردار ویژگی ایستا دریافت می‌کند که از متغیرهای اولیه قابل تفسیر ساخته‌شده است. اندازه گروه از یک قانون لاگ‌نرمال برش‌یافته[5] پیروی می‌کند:

$$participants \sim clip(Lognormal(\mu = 6, \sigma = 1.2), 30, 2 \times 10^5) \qquad (۳)$$

که منعکس‌کننده توزیع‌های دم‌پهن اندازه گروه مشاهده‌شده در پلتفرم‌های آنلاین است؛ علاوه بر این:

$$\begin{aligned} &activity_{ratio} \sim Beta(2,4), \\ &reply_{count} \sim Poisson(2), \\ &group_{age} \sim Exp(900) \end{aligned} \qquad (۴)$$

[1] Homophilous
[2] Bursty
[3] Spam
[4] Heavy-Tailed
[5] Clipped Log-Normal

و یک متغیر رشد گاوسی کوچک:

$$\text{growth}_{\text{proxy}} \sim \mathcal{N}(0.012, 0.015^2) \quad (۵)$$

توصیف‌گرهای مشتق‌شده اصلی عبارت‌اند از:

$$\text{active}_{\text{users}} = \text{activity}_{\text{ratio}} \times \text{participants} \quad (۶)$$

$$\text{silent}_{\text{ratio}} = 1 - \text{activity}_{\text{ratio}} \quad (۷)$$

$$\text{engagement}_{\text{level}} = \frac{\text{reply}_{\text{count}} + \text{active}_{\text{users}}}{\text{participants} + 1} \quad (۸)$$

$$\text{interaction}_{\text{score}} = \log(1 + \text{reply}_{\text{count}}) \times \text{activity}_{\text{ratio}} \quad (۹)$$

$$\text{interaction}_{\text{anomaly}} = (\text{participants}/2 \times 10^5) \times [1 - \text{engagement}_{\text{level}}] \quad (۱۰)$$

$$\text{active}_{\text{density}} = \frac{\text{active}_{\text{users}}}{\log(1 + \text{participants}) + 10^{-5}} \quad (۱۱)$$

پرچم‌های خطر دودویی[1] و کدهای وضعیت/نوع دسته‌بندی‌شده، بردار را کامل می‌کنند و شمارش‌های خام را به سیگنال‌های نرمال‌شده عدم فعالیت، عدم مشارکت و ناهنجاری اندازه تبدیل می‌کنند.

۳-۳-٤ ساخت برچسب از طریق امتیاز ریسک مبتنی بر تحلیل مؤلفه‌های اصلی[2]

برچسب‌ها به‌صورت دستی اختصاص داده نمی‌شوند؛ آن‌ها از یک امتیاز ریسک ترکیبی مشتق می‌شوند که وزن‌های آن به‌جای انتخاب دلخواه، با استفاده از تحلیل مؤلفه‌های اصلی [۳۰] از داده‌ها آموخته می‌شود. پنج مؤلفه ریسک خام استانداردشده محاسبه می‌شود.

جدول (٤): پنج مؤلفه ریسک خام

مؤلفه	تعریف(intuition)
R_1 (silent)	silent_ratio — fraction of inactive members
R_2 (anomaly)	clip(interaction_anomaly, ۰, ۱) — large size × low engagement
R_3 (growth)	۱ − minmax(growth_proxy) — stagnation risk (inverted)
R_4 (engagement)	۱ − engagement_level / max(engagement_level) — low-engagement risk
R_5 (bot)	۰/۷۰bot_activity_flag + ۰/۳۰low_activity_large_group_flag

فرض کنید $R = [R_1,\ldots,R_5]$ به میانگین صفر و واریانس یک استاندارد شود، $\tilde{R} = (R - \mu R)/\sigma R$. یک PCA تک‌مؤلفه‌ای روی $\tilde{R}$ برازش می‌شود، و بارهای مطلق بردار ویژه پیشرو ϕ از ماتریس کوواریانس آن به وزن‌های زیر نرمال می‌شوند:

$$w_i = \frac{|\phi_i|}{\sum_{j=1}^{5}|\phi_j|}, \quad \sum_i w_i = 1 \quad (۱۲)$$

ازآنجاکه مؤلفه اول (PC1) بیشترین میزان از واریانس مشترک را در میان همه مؤلفه‌ها به خود اختصاص می‌دهد، روش مذکور بیشترین ضریب وزنی را برای آن محوری در نظر می‌گیرد که نماینده اصلی تغییرات مشترک است [۳۰]. امتیاز ریسک و برچسب عبارت‌اند از:

$$\text{risk}(g) = \text{clip}\left(\sum_{i=1}^{5} w_i \cdot R_i(g), 0, 1\right) \quad (۱۳)$$

$$y_g = \mathbb{1}[\text{risk}(g) \geq \tau], \quad \tau = \text{Percentile}_{80}(\{\text{risk}(g): g \in G\}) \quad (۱۴)$$

که شیوع دقیق حدود ۲۰٪ گروه‌های مثبت را بدون تنظیم دستی اعمال می‌کند و به نگرانی‌های مربوط به تعادل کلاس که معیارهای ناهنجاری را تحریف می‌کنند، می‌پردازد [۳۱]. الگوریتم ۲ نحوه برچسب‌گذاری را ترسیم می‌کند.

الگوریتم (۲): امتیاز ریسک مبتنی بر PCA و عضویت همسان‌پسندانه

```
Part A — Labelling
 1: build R = [R1..R5] for all groups # Table 3
 2: R̃ ← StandardScaler.fit_transform(R)
 3: ϕ ← PCA(n=1).fit(R̃).components_[0]
 4: w ← |ϕ| / sum(|ϕ|)
 5: risk ← clip(R @ w, 0, 1)
 6: τ ← percentile(risk, 80); label ← (risk ≥ τ)
Part B — Initial membership (t=0), per group g
 7: w(u|g) ← 0.5 + 2·1[com(u)=com(g)] + 1.5·1[topic(u)=topic(g)] + 1·1[bot(u)]
 8: w ← w / sum(w)
 9: M_g ← sample_without_replacement(U, size=participants_g, p=w)
10: add edges {(u,g): u ∈ M_g} to bipartite graph
```

ازآنجاکه برچسب یک تابع قطعی از نرخ سکوت[3]، نرخ مشارکت[4] و ناهنجاری رفتاری در تعامل[5] است و این توصیف‌گرها در معرض مدل‌ها قرار می‌گیرند، این وظیفه ذاتاً برای هر مدلی که آن‌ها را

[1] Bot_activity_flag, Low_activity_large_group_flag, High_silent_ratio_flag, Zero_reply_flag
[2] PCA
[3] Silent_ratio
[4] Engagement_level
[5] Interaction_anomaly

مشاهده می‌کند آسان است. شش ستون ریسک داخلی (R_1–R_6) و امتیاز خام) قبل از ایجاد تقسیم‌بندی‌ها حذف می‌شوند، اما مواد سازنده آن‌ها باقی می‌مانند. این ویژگی، کلید تفسیر تمام نتایج پایین‌دستی (بخش ٤–١٠) است و به‌صورت شفاف و سازگار با ارزیابی آگاه از نشت گزارش می‌شود.

۳-۳-۵ شکل‌گیری عضویت همسان‌پسندانه ($t = 0$)

عضویت‌های اولیه همسان‌پسندی را نمونه‌سازی می‌کنند، تمایل بازیگران مشابه به ارتباط با یکدیگر [۳۲]. برای هر گروه، کاربران نامزد بدون جایگذاری با احتمالی متناسب با وزن شباهت نمونه‌برداری می‌شوند (الگوریتم ۲، بخش B):

$$w(u|g) = 0.5 + 2 \times I[com(u) = com(g)] + 1.5 \times I[topic(u) = topic(g)] + I[bot(u)] \quad (۱۵)$$

که روی تمام کاربران نرمال شده است. تعداد نمونه‌برداری‌شده برابر با تعداد شرکت‌کنندگان واقعی گروه (بدون سقف مصنوعی) است، که یک گراف عضویت دوبخشی تولید می‌کند که آمیختگی آن منعکس‌کننده SBM زیربنایی و ساختار همسان‌پسندانه است [۲۹]، [۳۲].

۳-۳-۶ پویایی‌های زمانی: AR(1)، فوران‌ها و عضویت پویا

دوازده سری زمانی تولید می‌شود. هر سری برای هر یک از گروه‌ها و به‌ازای متغیرهای تعداد پیام، تعداد دعوت، نسبت پاسخ، نرخ سکوت و تعداد اعضا، به‌عنوان یک فرآیند خودهمبسته مرتبه اول تکامل می‌یابد [۳۳]:

$$x_t = \mu + \varphi (x_{t-1} - \mu) + \varepsilon_t, \quad \varphi = 0.85, \varepsilon_t \sim \mathcal{N}(0, \sigma^2), \quad (۱۶)$$

که در آن φ = ۰/۸۵ ماندگاری قوی اما بازگشت به میانگین را تحمیل می‌کند و مدل متعارف برای سیگنال‌های زمانی با حافظه کوتاه [۳۳] دو مکانیزم واقع‌گرایی را می‌افزایند. رویدادهای فوران: با احتمال ۰/۱۵، یک گروه جهش ناگهانی ۳ تا ۸ برابری را در پیام‌ها و دعوت‌ها در یک تصویر تصادفی دریافت می‌کند، درحالی‌که نسبت پاسخ آن سرکوب می‌شود (۰/۳ برابر)، که موج‌های اسپم/حمله را مدل‌سازی کرده و از طریق پرچم is_burst ثبت می‌شود؛ چنین موج‌های کم‌پاسخی منعکس‌کننده تقویت خودکار هستند [۱]. عضویت پویا: پیوستن‌ها با دعوت‌ها و مشارکت مقیاس می‌شوند و خروج‌ها با عدم فعالیت:

$$join_t \sim Poisson(clip(2 \cdot invite\_t + 0.5 \cdot engagement \times 50, 1, 300)), \quad (۱۷)$$

$$leave_t \sim Poisson(clip((1 - activity_{ratio}) \cdot 10 + 2, 1, 50)), \quad (۱۸)$$

که در آن، کاربران پیوسته (مانده در شبکه) با وزن‌دهی مبتنی بر جامعه، موضوع و ربات نمونه‌برداری می‌شوند و کاربران خارج‌شونده (ترک‌کننده) نیز به‌صورت جامعه‌محور انتخاب می‌گردند، به‌گونه‌ای که حذف کاربران نامتناسب و ربات‌ها در اولویت قرار دارد. پس از هر گام، تعداد اعضا دقیقاً از مجموعه عضویت زنده خوانده می‌شود تا شمارش‌های زمانی و گراف عضویت باهم سازگار بمانند. الگوریتم ۳ جزئیات یک تصویر زمانی را نشان می‌دهد.

الگوریتم (۳): یک تصویر زمانی (AR(1) + فوران + عضویت پویا)

```
Input: prev series x_{t-1}, member sets {M_g}, burst schedule
 1: for each series s ∈ {msg,inv,reply,silent,mem}:
 x_t[s] ← clip(μ_s + φ(x_{t-1}[s]−μ_s) + ε, bounds_s)
# AR(1), φ=0.85
 2: for g with burst at t: msg_t,inv_t ← × BURST_MULT(3..8);
 reply_t ← ×0.3
 3: for each group g:
 n_join ← Poisson(clip(2·inv_t + 0.5·engagement×50, 1, 300))
 n_leave ← Poisson(clip((1−activity_ratio)·10 + 2, 1, 50))
 leave_score(u) ← 2·1[com≠] + 1.5·1[topic≠] + 1·bot;
 remove n_leave ∝ leave_score
 sample n_join new users ∝ community/topic/bot weights,
exclude current members
 M_g ← (M_g − leavers) ∪ joiners; member_count_g ← |M_g|
 4: persist 7 temporal features per group:
msg,inv,reply,silent,mem,net_change,is_burst
```

۳-۳-۷ ساخت گراف و فرافکنی گروه

یک گراف دوبخشی کاربر–گروه از عضویت‌های t = ۰ ساخته می‌شود. فرافکنی گروه–گروه از ضریب ژاکارد [۳٤] استفاده می‌کند که برای گروه‌های g_i و g_j با مجموعه‌های اعضای M_i و M_j برابر است با:

$$J(g_i, g_j) = |M_i \cap M_j| / |M_i \cup M_j|, \text{ edge created iff } J \geq 0.05. \quad (۱۹)$$

برای جلوگیری از هزینه $O(m^2)$ مقایسه همه باهم، اشتراک‌ها از طریق یک نمایه معکوس (کاربر ← گروه‌ها) تجمیع می‌شوند، بنابراین تنها جفت‌هایی که واقعاً دارای یک عضو مشترک هستند

ارزیابی می‌گردند. الگوریتم ٤ یک واقعیت ساختاری برای تحلیل‌های بعدی مرکزی است: با ۱۰۰۰۰ گروه که اعضای خود را از ۱۰۰۰۰۰ کاربر جذب می‌کنند، جفت‌های بسیار کمی به $J \geq$ ۰/۰۵ می‌رسند که تنها ٦۹ یال باقی می‌مانند و فرافکنی به ۹۹٦۳ مؤلفه تقسیم می‌شود. بنابراین مدل‌ها از رابطه دوبخشی متراکم یاد می‌گیرند (بخش ٤-۱۰).

الگوریتم (٤): فرافکنی گروه از طریق نمایه معکوس (ژاکارد دقیق)

```
Input: memberships (t=0), threshold θ=0.05
 1: inv ← {}; for (u,g) in memberships: inv[u].append(g)
 2: inter ← defaultdict(int)
 3: for u, groups_u in inv: for each pair (g_i<g_j) in groups_u:
inter[(g_i,g_j)] += 1
 4: for (g_i,g_j), c in inter:
J ← c / (|M_{g_i}| + |M_{g_j}| − c)
if J ≥ θ: add projection edge (g_i,g_j, weight=J)
 5: return projection edges
```

۳-۳-۸ اعتبارسنجی توزیع داخلی

ازآنجاکه هیچ بدنه مرجع واقعی در دسترس نبود، مولد دو کلاس برچسب را با دو معیار مکمل در تقابل قرار می‌دهد: آماره کولموگروف-اسمیرنوف[1] دونمونه‌ای [۳۵]، که بیشترین فاصله بین توابع توزیع تجمعی تجربی است:

$$KS = \sup_x | F_0(x) - F_1(x) |, \quad (۲۰)$$

و دیورژانس جنسن-شنون[2] [۳٦]، که یک نسخه هموارشده، متقارن و کران‌دار از دیورژانس کولبک-لایبلر[3] بر روی هیستوگرام‌های ٦۰ بازه‌ای است:

$$JS(P\|Q) = ½\, KL(P\|M) + ½\, KL(Q\|M),\ M = ½(P+Q). \quad (۲۱)$$

این معیارها تأیید می‌کنند که ویژگی‌های تعریف‌کننده برچسب، کلاس‌ها را از هم جدا می‌کنند درحالی‌که ویژگی‌های تصادفی این کار را نمی‌کنند. به‌طور شفاف گزارش می‌شود که این یک مقایسه داخلی بین برچسب ۰ و برچسب ۱ است، نه یک اعتبارسنجی مصنوعی در برابر واقعی.

۳-۳-۹ کنترل نشت و تخصیص لایه‌بندی‌شده

شش ستون ریسک داخلی حذف می‌شوند، پس‌ازآن یک تخصیص لایه‌بندی‌شده به‌صورت ۷۰ درصد آموزش ۱۵ درصد اعتبارسنجی و ۱۵ آزمون (با بذر ٤۲) منجر به ۷۰۰۰ گروه آموزشی، ۱۵۰۰ گروه اعتبارسنجی و ۱۵۰۰ گروه آزمون می‌شود، که هرکدام نرخ تقریباً برابر با ۲۰٪ مثبت را حفظ می‌کنند. لایه‌بندی از جابجایی شیوع ناشی از تخصیص جلوگیری می‌کند، که یک منبع شناخته‌شده از خوش‌بینی در ارزیابی نامتوازن است [۳۱].

٤-۳ نمایش گراف و کدگذاری ویژگی‌ها

مولد یک شیء واحد همگن داده از PyTorch-Geometric [۲۷] می‌سازد. ویژگی‌های طبقه‌بندی کاربر به‌صورت one-hot کدگذاری شده و با ویژگی‌های عددی کاربر الحاق می‌شوند. برای گروه‌ها، مجموعه ایستا شامل تمام ستون‌های عددی به‌جز شناسه، برچسب، موضوع و ستون‌های نشت حذف‌شده و مجموعه زمانی شامل هفت سیگنال زمانی در یک تصویر زمانی انتخاب شده است. ازآنجاکه فضاهای ویژگی کاربر و گروه در ابعاد متفاوت هستند ($_{g}d \neq {}_{u}d$)، هر دو با صفر به یک عرض مشترک $_{g}d + {}_{u}d = d$ حاشیه‌گذاری[4] شده و به‌صورت عمودی روی هم چیده می‌شوند:

$$X = [[X\_U \mid 0]; [0 \mid X\_G]] \in \mathbb{R}^{(n+m)\times d}, \quad (۲۲)$$

بنابراین یک گره کاربر، ویژگی‌های خود را در بلوک اول و صفرها را در بلوک گروه حمل می‌کند و بالعکس. این کار امکان تبادل پیام همگن را فراهم می‌سازد [۸]، [۲٦]. مقیاس‌بندی ویژگی‌ها بدون نشت است: یک مقیاس‌ساز استاندارد تنها روی گروه‌های آموزشی (در تمام تصاویر زمانی) برازش شده و بر روی گروه‌های اعتبارسنجی/آزمون اعمال می‌شود تا از تأثیر آمارهای آزمون بر آموزش جلوگیری کند.

٥-۳ پروتکل ارزیابی استقرایی

یک پروتکل استقرایی دقیق تضمین می‌کند که گروه‌های آزمون هرگز در طول آموزش مشاهده نمی‌شوند، و از نشت تراگذرایی[5] که بسیاری از نتایج یادگیری گراف را تحت تأثیر قرار می‌دهد،

[1] Kolmogorov-Smirnov Test
[2] Jensen-Shannon Divergence
[3] Kullback-Leibler Divergence
[4] Pad
[5] Transductive

جلوگیری می‌نماید [۸]. سه مجموعه یال تودرتو با فیلتر کردن هر دو رابطه به گروه‌های قابل‌مشاهده در هر فاز ساخته می‌شوند (الگوریتم ۵).

جدول (۵): زیرگراف‌های تفکیک‌شده براساس فاز که یک محیط استقرایی واقعی را اعمال می‌کنند

مرحله	گراف قابل‌مشاهده
آموزش	users + training groups only
اعتبارسنجی	users + training + validation groups
آزمون	full graph (users + all groups)

الگوریتم (۵): ساخت زیرگراف استقرایی

Input: full edge_index E, num_users n, allowed group sets per phase
1: train_groups = train_ids; val_groups = train_ids ∪ val_ids; test_groups = all
2: for phase in {train, val, test}:
keep edge (a,b) iff every group endpoint of (a,b) ∈ allowed(phase)
(bipartite edge: its single group endpoint; projection edge: both endpoints)
3: return E_train ⊆ E_val ⊆ E_test

ازآنجاکه گراف‌سیج توابع تجمیع را به‌جای جاسازی‌های[1] خاص هر گره یاد می‌گیرد، می‌تواند گروه‌های دیده‌نشده در زمان آموزش را نمایش دهد [۸].

۳-۶ معماری مدل‌ها

تمام معماری‌ها در تمام آزمایش‌ها ثابت هستند و تنها پیکربندی آموزش و ابرپارامترها تنظیم می‌شوند.

۳-۶-۱ مدل پیشنهادی گراف‌سیج

در این پژوهش آشکارساز پیشنهادی یک شبکه گراف‌سیج است [۸]، که کانولوشن گراف کامل را با تجمیع همسایگی مبتنی بر نمونه‌برداری قابل یادگیری جایگزین می‌کند و امکان مقیاس‌پذیری روی گراف با ۷/۶۴ میلیون یال و تعمیم استقرایی را فراهم می‌سازد. برای گره v باهمسایگی $\mathcal{N}(v)$، هر لایه k مقدار زیر را محاسبه می‌کند:

$$h_{\mathcal{N}(v)}^{(k)} = \mathrm{AGG}_k\left(\{h_u^{(k-1)} : u \in \mathcal{N}(v)\}\right), \quad (۲۳)$$

$$\mathrm{AGG} = \mathrm{mean(default)}$$

$$h_v^{(k)} = \sigma\left(W_k \cdot \left[h_v^{(k-1)} \,\|\, h_{\mathcal{N}(v)}^{(k)}\right]\right) \quad (۲۴)$$

که در آن $\|$ نشان‌دهنده الحاق و $\sigma = \mathrm{ReLU}$ است. تجمیع‌کننده میانگین[2] به‌صورت زیر پیاده‌سازی می‌شود:

$$h_v^{(k)} = \mathrm{ReLU}\left(W_k \cdot \left[h_v^{(k-1)} \,\|\, \mathrm{mean}_{u\in\mathcal{N}(v)} h_u^{(k-1)}\right]\right) \quad (۲۵)$$

این شبکه دولایه SAGEConv و یک سر پرسپترون چندلایه دوبلوکه را با حذف تصادفی [۳۷] پس از هر بلوک در هم می‌آمیزد:

SAGE → ReLU → Dropout → SAGE → ReLU → Dropout → Linear → ReLU → Dropout → Linear → logit.

دولایه انتشار، به هر گروه یک میدان پذیرش دوهاپی[3] می‌دهد که اعضای آن و سایر گروه‌هایی که آن اعضا به آن‌ها تعلق دارند را شامل می‌شود و سیگنال رابطه‌ای که فرضیه گراف بر آن تکیه دارد، قابل‌دسترس خواهد بود [۸]، [۲۳]. الگوریتم ٦ عبور روبه‌جلو[4] را نشان می‌دهد.

الگوریتم (٦): عبور روبه‌جلو گراف‌سیج (به ازای گره v)

1: h₀(v) ← x_v
2: for k = 1,2:
a ← mean_{u∈N(v)} h_{k-1}(u)
h_k(v) ← ReLU(W_k · concat(h_{k-1}(v), a))
h_k(v) ← dropout(h_k(v), p=0.5)
3: z ← ReLU(W3 · h_2(v)); z ← dropout(z); logit(v) ← W4 · z
4: return logit(v)

۳-۶-۲ شبکه عصبی کانولوشنی گراف

یک خط پایه GCN [۲۲] با عمق و سر یکسان، لایه‌های SAGEConv را با دولایه GCNConv جایگزین می‌کند که انتشار نرمال‌شده متقارن را اعمال می‌نمایند:

$$H^{(k)} = \sigma\left(\tilde{D}^{-1/2}\tilde{A}\tilde{D}^{-1/2}H^{(k-1)}W_k\right), \quad \tilde{A} = A + I \quad (۲۶)$$

که در آن $\tilde{A}$ ماتریس مجاورت با حلقه‌های خودی و $\tilde{D}$ ماتریس درجه آن است [۲۲]. این میانگین‌گیری ثابت و نرمال‌شده با درجه، در برابر هموارسازی بیش‌ازحد[5] روی گراف‌های با درجه بالا آسیب‌پذیر است [۳۸]، [۳۹].

[1] Embeddings
[2] Mean
[3] Two-Hop Receptive Field
[4] Forward-Pass
[5] Over-Smoothing

۳-۶-۳ شبکه CARE-GNN

مدل CARE-GNN [٤]یک آشکارساز تقلب چندرابطه‌ای است. کانولوشن درون‌رابطه‌ای آن یک تجمیع آگاه از برچسب و فیلترشده با شباهت را انجام می‌دهد: برای هدف i و همسایه j، پیام را با شباهت کسینوسی یکسوشده[1] وزن‌دهی می‌کند و در طول آموزش، تنها p درصد برتر از شبیه‌ترین همسایگان را نگه می‌دارد:

$$s_{ij} = \max\left(0, \cos(h_i, h_j)\right), \quad (۲۷)$$

$$m_{ij} = (Wh_j) \cdot s_{ij} \cdot 1[s_{ij} \geq Q_{1-p}], \quad p = 0.7 \quad (۲۸)$$

که در آن، Q_{1-p} برابر با چندک 1-p ام (یا صدک 1-p ام) توزیع شباهت‌ها است. دو کانولوشن مشابه، به‌طور جداگانه بر روی روابط دوبخشی و فرافکنی اعمال می‌شوند. سپس، یک لایه Softmax با وزن‌های قابل یادگیری، خروجی این دو کانولوشن را به‌صورت وزن‌دار با یکدیگر ادغام می‌کند و حاصل، پیش از ورود به بخش پرسپترون چندلایه، ترکیب می‌شود:

$$h = \mathrm{ReLU}(\alpha_1 h_{\text{bipartite}} + \beta_1 h_{\text{projection}}), \quad (۲۹)$$
$$(\alpha_1, \beta_1) = \mathrm{softmax}(\text{inter\_weights})$$

CARE-GNN تنها رقیبی است که به‌طور صریح از رابطه فرافکنی بهره‌برداری می‌کند که با توجه به تقریباً خالی بودن آن (بخش ۳-۳-۷)، در بخش ٤-۱۰ تعیین‌کننده است.

۳-۶-٤ مدل‌های پایه جدولی کلاسیک و عصبی

چهار خط پایه غیرگرافی، مقایسه را کامل می‌کنند. مدل لایه‌های GroupMLP در رابطه ۳۰ درج‌شده است.

$$\text{Linear}(d \to 128) \to \text{ReLU} \to \text{Dropout} \to \text{Linear}(128 \to 64) \to \text{ReLU} \to \text{Linear}(64 \to 1) \quad (۳۰)$$

مدل را بر روی بردار تخت گروه بدون گراف اعمال می‌کند. خطوط پایه کلاسیک عبارت‌اند از رگرسیون لجستیک، جنگل تصادفی [٤۰] و یک گروه تقویت‌شده با گرادیان تنظیم‌شده به‌نام XGBoost [٤۱]، یادگیرنده‌های جدولی قوی و پرکاربردی که به‌عنوان مراجع چالش‌برانگیز عمل می‌کنند [٤۰]، [٤۱].

۳-۷ روند آموزش و بهینه‌سازی

تمام مدل‌های عصبی یک تابع زیان آنتروپی متقاطع دودویی با توازن کلاس را حداقل می‌کنند. با توجه به عدم توازن کلاس‌ها، یک وزن کلاس مثبت از فرکانس‌های آموزشی تنظیم می‌شود:

$$\mathcal{L} = -\frac{1}{N}\sum_{i=1}^{N}[w_+ y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)], \quad (۳۱)$$

$$w_+ = \frac{\text{freq}(0)}{\text{freq}(1)}$$

که بدون نمونه‌برداری مجدد با عدم توازن شیوع مقابله می‌کند [۳۱]. بهینه‌سازی از Adam [٤۲] استفاده می‌کند، یک زمان‌بند ReduceLROnPlateau (ضریب ۰/۵، صبر[2] ۱۵ روی زیان اعتبارسنجی) نرخ را کاهش می‌دهد و گرادیان‌ها در حداکثر نرم ۵ برش داده می‌شوند. Dropout سنگین با پارامتر p برابر ۰/۵ [۳۷] و کاهش وزن، مدل را تنظیم می‌کنند. برای گراف پویا، مدل در این پژوهش «چرخشی» آموزش داده می‌شود، به این صورت که در هر دوره، یک تصویر زمانی (از میان ۱۲ تصویر) انتخاب شده و این تصاویر به ترتیب چرخش می‌یابند، درحالی‌که ارزیابی همواره روی آخرین تصویر زمانی انجام می‌گیرد. تمام مدل‌های عصبی از یک بودجه مشترک ۳۰۰ دوره‌ای بهره می‌برند و دارای صبر برابر با ٤۰ بر اساس معیار F1 بر روی داده‌های اعتبارسنجی هستند و درنهایت وزن‌های مربوط به بهترین مقدار F1 بازیابی می‌شوند. آستانه تصمیم‌گیری در بازه [۰/۹۵, ۰/۰۵] (شامل ۹۱ مقدار) به‌گونه‌ای جستجو می‌شود که بیشترین مقدار F1 را روی داده‌های اعتبارسنجی به دست دهد، سپس همین آستانه، بدون تغییر، روی مجموعه آزمون اعمال می‌گردد. این امر، انتخاب یک نقطه عملیاتی مهم در شرایط عدم توازن داده‌ها محسوب می‌شود [۵۲، ٦٤]. الگوریتم ۷، روند کلی آموزش را خلاصه کرده است.

۳-۷-۱ تنظیم ابرپارامترها

هر مدل توسط جستجوی شبکه‌ای[3] بر روی F1 اعتبارسنجی تنظیم می‌شود. ازآنجاکه گراف فراتر از ۵۰۰۰۰۰ یال است، شبکه‌ها به‌طور خودکار به ۲ تا ۳ پیکربندی برای هر مدل کاهش می‌یابند و حالت

[1] Rectified
[2] Patience
[3] Grid Search

چرخش استفاده می‌شود که جامعیت را فدای زمان اجرای قابل‌مدیریت می‌کند. پیکربندی‌های انتخاب‌شده در جدول ٦ آمده است.

الگوریتم (۷): آموزش GNN استقرایی با بهینه‌سازی آستانه

```
Input: model, E_train/E_val/E_test, snapshots {0..11}, budget=300, patience=40
 1: criterion ← BCEWithLogits(pos_weight=freq0/freq1); opt ← Adam(5e-4, wd=5e-4)
 2: best_f1 ← −1; wait ← 0
 3: for epoch = 1..300:
 snap ← rotate: epoch mod 12
 X_G ← scaler.transform(features(snap)); update group node features
 logits ← model(X, E_train); loss ← criterion(logits[train], y[train])
 backprop; clip_grad_norm(5); opt.step()
 evaluate on E_val (final snapshot); scheduler.step(val_loss)
 if val_f1 > best_f1: best_f1 ← val_f1; save weights; wait ← 0
 else wait += 1
 if wait ≥ 40: break
 4: restore best weights
 5: τ* ← argmax_{τ∈[0.05,0.95]} F1_val(τ); report test metrics at τ*
```

۳–۸ پروتکل ارزیابی

مدل‌ها با مجموعه استاندارد برای طبقه‌بندی دودویی نامتوازن ارزیابی می‌شوند: صحت[1]، دقت[2]، فراخوانی[3]، معیار F1، مساحت زیر منحنی ROC[4] و میانگین دقت (مساحت زیر منحنی دقت–فراخوانی[5]). ازآنجاکه صحت تحت شرایط عدم توازن گمراه‌کننده است، بر F1 و معیارهای مستقل از آستانه AP/AUC تأکید می‌شود؛ نمای دقت–فراخوانی برای مسائل با مثبت‌های نادر بسیار آگاهی‌بخش است [٤٣]، [٤٤]. برای ارزیابی استحکام[6]، مدل پیشنهادی روی پنج بذر تصادفی (٤٢، ٥٢، ٦٢، ٧٢، ٨٢) مجدد آموزش داده‌شده و میانگین ± انحراف معیار گزارش می‌گردد. ادعای اینکه مدل گراف خط پایه جدولی را شکست می‌دهد با استفاده از آزمون رتبه علامت‌دار ویلکاکسون[7] [٤٥] یک‌طرفه بر روی امتیازهای زوجی بذرها آزمایش می‌شود، که روش ناپارامتری توصیه‌شده برای مقایسه طبقه‌بندهاست [٤٦]. زمان‌های آموزش و استنتاج به‌صورت زمان واقعی[8] برای هر مدل ثبت می‌شود.

جدول (٦): ابرپارامترهای انتخاب‌شده برای هر مدل

مدل	ابرپارامترهای تنظیم شده
گراف‌سیج (پیشنهادی)	hidden ۱۲۸, dropout ۰/۵, lr ۵×۱۰$^{-۴}$, weight-decay ۵×۱۰$^{-۴}$, aggr = mean
GCN	hidden ۱۲۸, dropout ۰/۵, lr ۵×۱۰$^{-۴}$, weight-decay ۵×۱۰$^{-۴}$
CARE-GNN	hidden ٦٤, dropout ۰/٤, lr ۱×۱۰$^{-۳}$, weight-decay ۱×۱۰$^{-۴}$, top$_p$ ۰/۷
Random Forest	۲۰۰ trees, unlimited depth, min_split ۲
XGBoost	۲۰۰ trees, max_depth ٤, lr ۰/۰۵, subsample ۰/۸

٤– نتایج و ارزیابی

این بخش یافته‌های تجربی را گزارش می‌دهد. عملکرد مطلق مدل گراف‌سیج پیشنهادی (بخش ٤–۱)، پویایی‌های یادگیری آن (بخش‌های ٤–۲ و ٤–۳)، مقایسه در برابر تمام خطوط پایه ازنظر صحت و هزینه (بخش‌های ٤–٤ و ٤–۵) و تحلیل‌ها از طریق مطالعه حذف، پایداری چندبذری و آزمون معناداری، تحلیل خطا، اعتبارسنجی توزیع، اهمیت ویژگی‌ها و تشخیص‌های ساختار گراف (بخش‌های ٤–٦ تا ٤–۱۰) ارائه خواهد یافت.

٤–۱ عملکرد مدل پیشنهادی

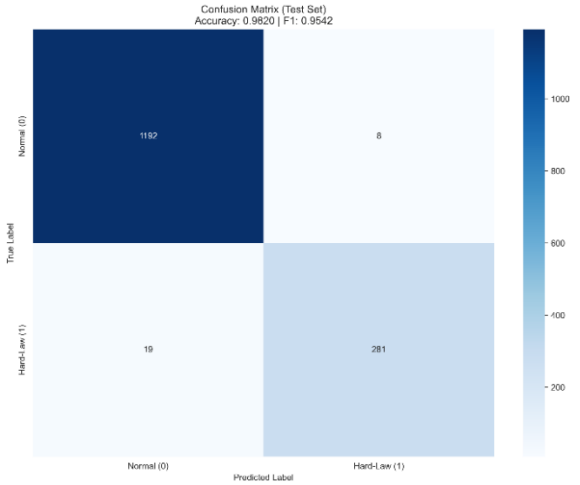
مدل گراف‌سیج تنظیم‌شده برای بودجه کامل آموزش داده شد، و در دوره ۲۵۸ از ۲۹۸ با آستانه تصمیم بهینه‌شده ۰/۸۹ متوقف گردید. بر روی گراف آزمون کاملاً استقرایی و نگه‌داشته شده[9]، این مدل به نتایج جدول ۷ دست می‌یابد.

جدول (۷): عملکرد مجموعه آزمون گراف‌سیج (۱۵۰۰ گروه)

میانگین دقت	AUC	F1	فراخوانی	دقت	صحت
۰/۹۹۵۱	۰/۹۹۸۷	۰/۹۵۴۲	۰/۹۳٦۷	۰/۹۷۲۳	۰/۹۸۲۰

[1] Accuracy
[2] Precision
[3] Recall
[4] AUC
[5] AP
[6] Robustness
[7] WilcoxonS rank-ignedTest
[8] Wall-clock
[9] Held-out

در شرایط مطلق، عملکرد مدل قوی است: بر روی ۱۵۰۰ گروه آزمون (۱۲۰۰ عادی، ۳۰۰ سخت‌قانون) این مدل تنها ۲۷ خطا (۸ مثبت کاذب و ۱۹ منفی کاذب) مرتکب می‌شود که منجر به ۱۱۹۲ منفی صحیح و ۲۸۱ مثبت صحیح می‌گردد (شکل ۱).

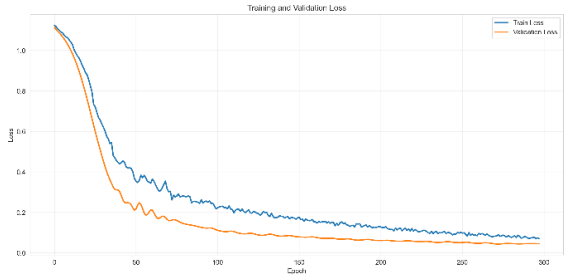


شکل (۱): ماتریس درهم‌ریختگی مجموعه آزمون در مدل گراف‌سیج

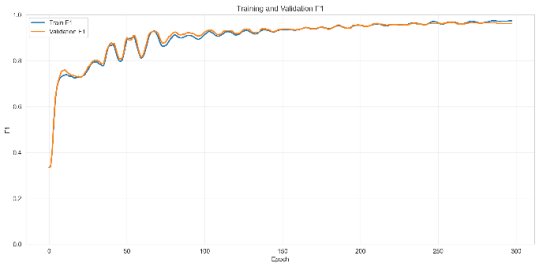
تعداد بسیار پایین مثبت کاذب (دقت ۰/۹۷۲) به این معنی است که تقریباً هر گروهی که پرچم‌گذاری شده، واقعاً پرخطر است، درحالی‌که ۱۹ منفی کاذب (فراخوانی ۰/۹۳۷) خطاهای پرهزینه‌تری برای تعدیل‌سازی[1] هستند و در بخش ٤–٨ توصیف می‌شوند.

٤–۲ پویایی‌های یادگیری و تحلیل بیش‌برازش

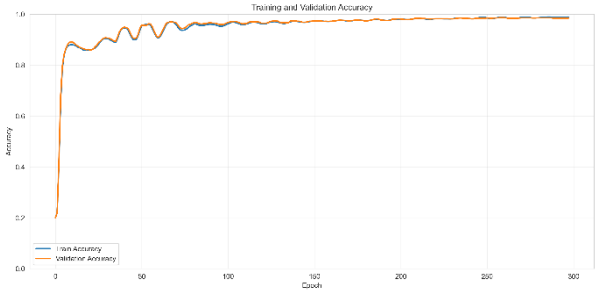
هر دو زیان آموزش و اعتبارسنجی به‌آرامی از ۱/۱۲≈ به ۰/۰۵≈ طی ۲۹۸ دوره کاهش می‌یابند (شکل ۲). زیان اعتبارسنجی به‌طور پیوسته در زیر زیان آموزش قرار دارد که پیامد موردانتظار استفاده از Dropout سنگین (p = ۵/۰) و وزن‌دهی متوازن کلاسی است که زیان آموزش اندازه‌گیری‌شده را برجسته می‌کند، همراه با یک زیرگراف اعتبارسنجی که اندکی آسان‌تر است. صحت و F1 (شکل‌های ۳ و ٤) در ۲۵ دوره اول به‌سرعت افزایش می‌یابند، از یک مرحله اکتشافی کوتاه تا حدود دوره ۷۰ می‌گذرند و سپس در یک سطح پایدار نزدیک به ۰/۹۸–۰/۹۹ مستقر می‌شوند، به‌طوری‌که منحنی‌های آموزش و اعتبارسنجی اساساً روی هم قرار می‌گیرند. دقت (شکل ۵) به ۰/۹۳≈–۰/۹۵ همگرا می‌شود و فراخوانی (شکل ٦) در طول آموزش روی ۰/۹۹≈ ثابت می‌ماند. هیچ شواهدی از بیش‌برازش وجود ندارد [۳۷].



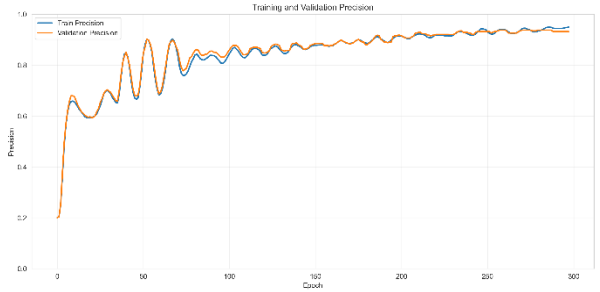
شکل (۲): زیان آموزش در برابر اعتبارسنجی در طول ۲۹۸ دوره



شکل (۳): معیار F1 آموزش در برابر اعتبارسنجی — منحنی‌ها به‌دقت یکدیگر را دنبال می‌کنند

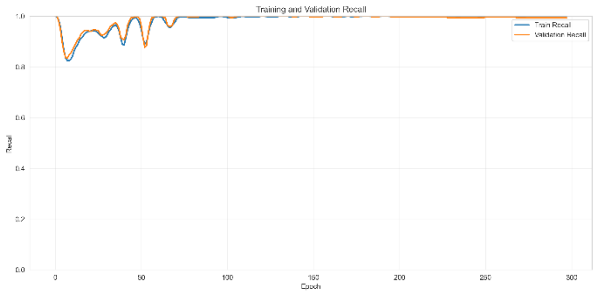


شکل ٤) صحت آموزش در برابر اعتبارسنجی، افزایش سریع سپس یک سطح پایدار نزدیک به ۰/۹۸ – ۰/۹۹



شکل (۵): منحنی‌های دقت رسیدن همگرایی به ≈ ۰/۹۳ – ۰/۹۵

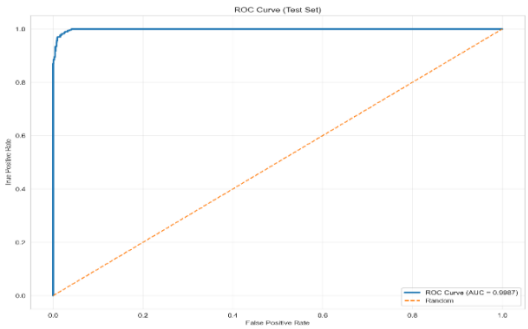
[1] Moderation



شکل (٦): منحنی‌های فراخوانی که در طول آموزش به ≈۰/۹۹ رسیده و ثابت می‌مانند

٤-٣ تفکیک‌پذیری: ROC و Precision–Recall

هر دو منحنی مستقل از آستانه، تفکیک‌پذیری تقریباً کامل را تأیید می‌کنند: AUC = ۰/۹۹۸۷ (شکل ۷) و میانگین دقت برابر با ۰/۹۹۵۱ (شکل ۸). مدل، دو کلاس را در فضای امتیاز مرتب می‌کند، بنابراین F1 برابر با ۰/۹۵٤ تنها منعکس‌کننده تبادل[1] در نقطه کار با آستانه ۰/۸۹ است، نه رتبه‌بندی ضعیف. پایین آوردن آستانه درصورتی‌که از دست دادن گروه‌های پرخطر پرهزینه‌تر از هشدارهای کاذب باشد، دقت را فدای فراخوانی می‌کند، تصمیمی که منحنی دقت–فراخوانی آن را به‌صراحت نشان می‌دهد [٤٣].

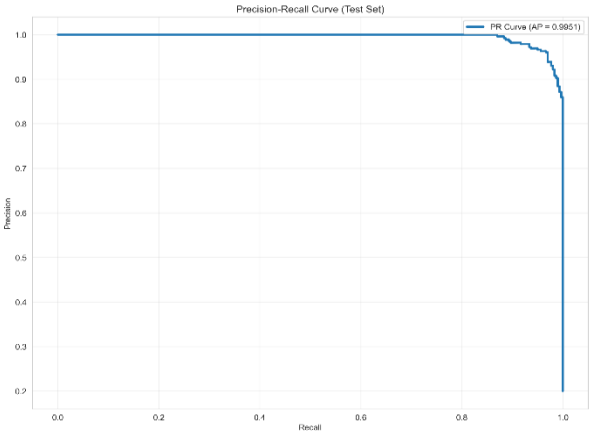


شکل (۷): منحنی ROC (آزمون)، AUC =۰/۹۹۸۷ ، مماس بر گوشه بالا–چپ

٤-٤ نتایج مقایسه‌ای

جدول ۸ و شکل ۹ تمام هفت مدل را بر روی مجموعه آزمون گزارش می‌دهند. سه سطح واضح پدیدار می‌شود.

سطح ۱ یادگیرنده‌های جدولی XGBoost (۱F = ۰/۹۹۲) و جنگل تصادفی (۰/۹۸۸) و رگرسیون لجستیک (۰/۹۸۸) قوی‌ترین هستند که همگی از گراف‌سیج پیشی می‌گیرند و در سطح ۲ مدل‌ها گراف‌سیج و پرسپترون چندلایه (۱F ≈ ۰/۹٥٤–۰/۹٥٦) قرار دارند. مدل گراف‌سیج در صحت با پرسپترون چندلایه برابری می‌کند اما دقت به‌طور چشمگیری بهتری دارد، آن هم با هزینه محاسباتی زیاد. در سطح ۳ GCN و GNN-CARE به ۱F = ۰/۳۳۳ سقوط می‌کنند (بخش ٤–۱۰). تسلط مدل‌های جدولی پیامد مستقیم این است که برچسب تابعی از ویژگی‌های ایستای گروه است (بخش ٤–۹)، که مدل‌های تقویت گرادیان و خطی قوی آن را تقریباً به‌طور کامل بازیابی می‌کنند [٤١]، [٤٠].



شکل (۸): منحنی دقت–فراخوانی (آزمون)، میانگین دقت (AP) = ۰/۹۹۵۱

٤-٥ هزینه محاسباتی

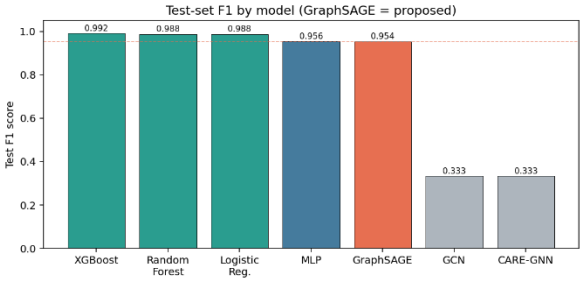
شکاف هزینه–فایده اصلی‌ترین یافته عملی است (شکل ۱۰). مدل گراف‌سیج برای آموزش به حدود ۹۸۲/۱ ثانیه (≈۳۳ دقیقه) زمان نیاز دارد، درحالی‌که این زمان برای XGBoost برابر ۱/٤۲ ثانیه و برای رگرسیون لجستیک ۰/۰۳ ثانیه است، اختلاف ۳٥۰ برابری تا ٦۰۰۰۰ برابری، درحالی‌که هیچ مزیت صحتی بر روی این معیار ارائه نمی‌دهد. برای وظیفه‌ای که سیگنال آن جدولی است، مدل گراف گزینه‌ای گران‌تر و اندکی ضعیف‌تر است، این تبادل کمی[2] به‌نوبه خود یک نتیجه قابل گزارش است.

[1] Trade-off

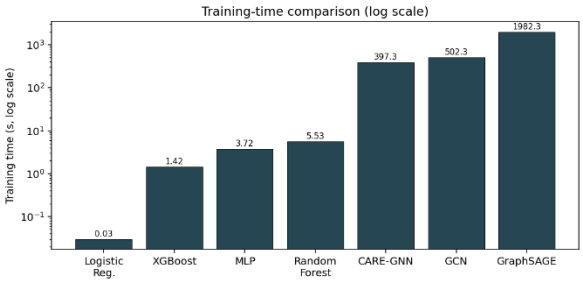
[2] Quantified Trade-off

جدول (۸): نتایج مقایسه‌ای مجموعه آزمون برای تمام مدل‌ها

مدل	صحت	دقت	فراخوانی	F1	AUC	آموزش (ثانیه)	استنتاج (ثانیه)
XGBoost (تنظیم شده)	۰/۹۹۶۷	۰/۹۸۳۶	۱/۰۰۰۰	۰/۹۹۱۷	۱/۰۰۰۰	۱/۴۲	۰/۰۰۷
Random Forest	۰/۹۹۵۳	۰/۹۹۰۰	۰/۹۸۶۷	۰/۹۸۸۳	۱/۰۰۰۰	۵/۵۳	۰/۰۶۳
Logistic Regression	۰/۹۹۵۳	۰/۹۹۳۳	۰/۹۸۳۳	۰/۹۸۸۳	۰/۹۹۹۹	۰/۰۳	۰/۰۰۰
MLP	۰/۹۸۲۰	۰/۹۳۳۳	۰/۹۸۰۰	۰/۹۵۶۱	۰/۹۹۸۵	۳/۷۲	۰/۰۰۰
گراف سیج (پیشنهادی)	۰/۹۸۲۰	۰/۹۷۲۳	۰/۹۳۶۷	۰/۹۵۴۲	۰/۹۹۸۷	۱۹۸۲/۳	۴/۵۷
GCN (تنظیم شده)	۰/۲۰۰۰	۰/۲۰۰۰	۱/۰۰۰۰	۰/۳۳۳۳	۰/۵۴۱۲	۵۰۲/۳	۶/۸۷
CARE-GNN (تنظیم شده)	۰/۲۰۰۰	۰/۲۰۰۰	۱/۰۰۰۰	۰/۳۳۳۳	۰/۴۹۵۴	۳۹۷/۳	۴/۴۵



شکل (۹): معیار F1 مجموعه آزمون بر اساس مدل. یادگیرنده‌های جدولی (سبز) پیشرو هستند؛ گراف‌سیج (نارنجی) مدل پیشنهادی است و مدل‌های GCN/CARE-GNN فرومی‌پاشند



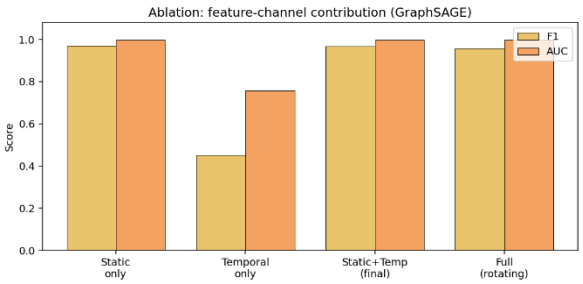
شکل (۱۰): مقایسه زمان آموزش (مقیاس لگاریتمی): گراف‌سیج بین ۳۵۰ تا ۶۰۰۰۰ برابر کندتر از مدل‌های جدولی است

۴-۶ مطالعه حذفی[1]

برای جداسازی مشارکت هر کانال ویژگی، گراف‌سیج تحت چهار پیکربندی ویژگی مجدد آموزش داده شد (جدول ۹ و شکل ۱۱).

جدول (۹): مطالعه حذف بر روی کانال‌های ویژگی

پیکربندی	صحت	دقت	فراخوانی	F1	AUC
فقط ویژگی‌های ایستا	۰/۹۸۸۷	۰/۹۷۳۲	۰/۹۷۰۰	۰/۹۷۱۶	۰/۹۹۹۲
فقط ویژگی‌های زمانی	۰/۶۳۶۰	۰/۳۲۲۸	۰/۷۴۶۷	۰/۴۵۰۷	۰/۷۵۷۲
ایستا + زمانی (final snapshot)	۰/۹۸۷۳	۰/۹۷۳۱	۰/۹۶۳۳	۰/۹۶۸۲	۰/۹۹۸۸
کامل (ایستا + زمانی، چرخشی)	۰/۹۸۳۳	۰/۹۶۶۱	۰/۹۵۰۰	۰/۹۵۸۰	۰/۹۹۸۷



شکل (۱۱): ویژگی‌های ایستا تقریباً تمام سیگنال را حمل می‌کنند، ویژگی‌های زمانی به‌تنهایی ضعیف هستند و هنگام اضافه‌شدن اندکی به عملکرد آسیب می‌زنند

ویژگی‌های ایستا اساساً تمام سیگنال تفکیک‌پذیر را حمل می‌کنند (به‌تنهایی F۱ = ۰/۹۷۲)، که پیامد مستقیم تعریف برچسب است. ویژگی‌های زمانی به‌تنهایی ضعیف هستند (F۱ = ۰/۴۵۱، AUC =

[1] Ablation Study

۰/۷۵۷): پویایی‌های AR(۱)/فوران تنها همبستگی ضعیفی با برچسب دارند. جالب‌توجه است که افزودن اطلاعات زمانی اندکی عملکرد را کاهش می‌دهد، بنابراین کانال زمانی برای این برچسب بیش از سیگنال، نویز وارد می‌کند. درنتیجه، مدل‌سازی زمانی در اینجا باید به‌جای منبع دقت، به‌عنوان یک مطالعه عمومیت‌بخشی در نظر گرفته شود [۴۷]، [۴۸].

۴-۷ پایداری و معناداری آماری

در سراسر پنج بذر، گراف‌سیج پایدار است (جدول ۱۰)، اما مقایسه زوجی با بهترین خط پایه ادعای برتری را تأیید نمی‌کند.

جدول (۱۰): پایداری پنج‌بذری برای گراف‌سیج

معیار	گراف‌سیج (mean ± std)
صحت	۰۱۳/۰ ± ۹۷۶/۰
دقت	۰۶۰/۰ ± ۹۱۹/۰
فراخوانی	۰۱۱/۰ ± ۹۶۷/۰
F1	۰۲۸/۰ ± ۹۴۲/۰
AUC	۰۰۱/۰ ± ۹۹۸/۰

آزمون یک‌طرفه رتبه‌ای علامت‌دار ویلکاکسون [۴۵]، [۴۶] برای فرضیه «XGBoost < GraphSAGE» بر روی پنج امتیاز 1F زوجی، آماره‌ای برابر با ۰/۰ و مقدار p برابر با ۱/۰ را نشان می‌دهد که در سطح معناداری $\alpha = ۰/۰۵$، معنادار نیست. در هر بذر، مقدار 1F مربوط به XGBoost (حدود ۰/۹۸۸ تا ۰/۹۹۲) از گراف‌سیج (۰/۸۹۷ تا ۰/۹۷۰؛ میانگین ۰/۹۴۱۵ در برابر ۰/۹۹۰۱) بالاتر بوده است. بنابراین، در این معیار مصنوعی، مدل گرافی ازنظر آماری بهتر از خط‌پایه جدولی عمل نمی‌کند و درعین‌حال به‌مراتب گران‌تر از آن است.

۴-۸ تحلیل خطا

میانگین مقادیر ویژگی به ازای دسته پیش‌بینی (جدول ۱۱) نشان می‌دهد که خطاها به‌جای تکیه‌بر میان‌برهای تصادفی، در مرز واقعی کلاس متمرکزشده‌اند.

منفی‌های کاذب در ۱۹ گروه پرخطری که نادیده گرفته‌شده‌اند؛ به‌طور چشمگیری بزرگ‌تر هستند (میانگین ۱۵۲۱ شرکت‌کننده در برابر ≈۸۴۰ در کل) و بالاترین نرخ سکوت (۰/۷۱۵) را دارند: گروه‌های بزرگ و خاموش نزدیک به مرز که با یک آستانه آگاه از اندازه یا ویژگی قابل بازیابی هستند. مثبت‌های کاذب (۸) کوچک‌تر (۵۱۵ شرکت‌کننده) با فعالیت/مشارکت بالاتر هستند: گروه‌های کوچک با فعالیت مرزی که به‌اشتباه پرچم‌گذاری شده‌اند. مثبت‌های صحیح و منفی‌های صحیح میانگین ویژگی تقریباً یکسانی دارند، که تأیید می‌کند مدل از یک همبستگی جعلی بهره‌برداری نمی‌کند.

جدول (۱۱): میانگین ویژگی‌ها بر اساس نتیجه پیش‌بینی

دسته بندی	تعداد	activity	silent	engage	reply	partici pants
مثبت صحیح	۲۸۱	۳۳۹/۰	۶۶۱/۰	۳۴۳/۰	۹۷/۱	۸۳۷
منفی صحیح	۱۱۹۲	۳۲۸/۰	۶۷۲/۰	۳۳۳/۰	۹۷/۱	۸۴۵
مثبت کاذب	۸	۳۹۵/۰	۶۰۵/۰	۴۰۵/۰	۳۸/۲	۵۱۵
منفی کاذب	۱۹	۲۸۵/۰	۷۱۵/۰	۲۹۵/۰	۵۳/۲	۱۵۲۱

۴-۹ اعتبارسنجی توزیع و اهمیت ویژگی‌ها

اعتبارسنجی توزیع داخلی (جدول ۱۲) درواقع یک عیب‌یابی برچسب است. سه ویژگی که امتیاز ریسک را تعریف می‌کنند، کلاس‌ها را تقریباً به‌طور کامل از هم جدا می‌کنند (KS ≈ ۹۸/۰ p = ۰ و JS ≈ ۸۱/۰)، درحالی‌که ویژگی‌های غایب در امتیاز ازنظر آماری بین کلاس‌ها غیرقابل‌تشخیص هستند [۳۵]، [۳۶] (KS ≈ ۰۲/۰ و p ≫ ۰۵/۰).

تحلیل اهمیت ویژگی‌ها این موضوع را از سمت مدل نیز تأیید می‌کند (شکل ۱۳). هم دیدگاه مبتنی بر ناخالصی (جنگل تصادفی [۴۰]) و هم دیدگاه خطی (رگرسیون لجستیک) توافق دارند که silent_ratio ،activity_ratio و engagement_level غالب هستند و درمجموع حدود ۷۹٪ از اهمیت جنگل تصادفی و بزرگ‌ترین ضرایب لجستیک را با علائم موردانتظار تشکیل می‌دهند (جدول ۱۳). طبقه‌بندها به‌درستی تعریف ریسک مولد را بازیابی کرده‌اند.

جدول (۱۲): اعتبارسنجی توزیع داخلی برچسب ۰ در برابر برچسب ۱

ویژگی	KS	مقدار آماره p-value	JS	μ (۰ برچسب)	μ (۱ برچسب)
activity_ratio	۹۸۲/۰	۰/۰	۸۱۲/۰	۳۸۷/۰	۱۰۹/۰
silent_ratio	۹۸۲/۰	۰/۰	۸۱۲/۰	۶۱۳/۰	۸۹۱/۰
engagement_level	۹۷۷/۰	۰/۰	۸۱۰/۰	۳۹۳/۰	۱۱۴/۰
reply_count	۰۱۸/۰	۷۰۷/۰	۰۳۱/۰	۰۲/۲	۹۸/۱
participants	۰۲۴/۰	۳۳۱/۰	۰۳۷/۰	۸۲۸	۸۳۶
growth_proxy	۰۱۹/۰	۵۹۰/۰	۰۷۲/۰	۰۱۲/۰	۰۱۲/۰

جدول (۱۳): برترین ویژگی‌ها بر اساس اهمیت جنگل تصادفی و ضریب رگرسیون لجستیک

ویژگی	اهمیت جنگل تصادفی	ضریب رگرسیون لجستیک
silent_ratio	۳۱۴/۰	۲۵/۷+
activity_ratio	۲۵۸/۰	۲۵/۷−
engagement_level	۲۱۶/۰	۱۷/۸−
interaction_score	۰۸۶/۰	۱۳/۲−
active_density	۰۳۵/۰	۹۶/۱−

۴-۱۰ ساختار گراف و فروپاشی GCN/CARE-GNN

آمارهای گراف آموزشی (جدول ۱۴) هم ضعیف بودن رابطه فرافکنی و هم علت شکست دو مورد از مدل‌های گراف را توضیح می‌دهد.

جدول (۱۴): آمارهای گراف آموزشی

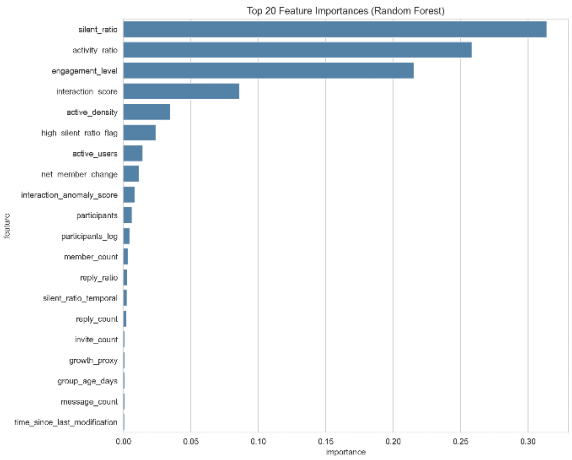
معیار	مقدار
گره‌های ترکیبی	۱۱۰٫۰۰۰
یال‌های ترکیبی	۷٫۶۴۱٫۳۲۹
چگالی ترکیبی	۰۰۱۲۶/۰
یال‌های دوبخشی	۷٫۶۴۱٫۲۶۰
یال‌های فرافکنی	۶۹
درجه میانگین	۹/۱۳۸
حداکثر درجه	۵٫۰۰۷
مؤلفه‌های همبند (ترکیبی)	۱
مؤلفه‌های همبند (فرافکنی)	۹٫۹۶۳
خوشه‌بندی میانگین (فرافکنی)	۰۰۲۴/۰

گراف ترکیبی از یک مؤلفه همبند واحد تشکیل‌شده است که بیش از ۶۴/۷ میلیون یال دوبخشی را در برمی‌گیرد؛ در مقابل، گراف فرافکنی تنها شامل ۶۹ یال است که میان ۹۹۶۳ مؤلفه مجزا توزیع‌شده‌اند. در چنین ساختاری، هر دو مدل GCN و CARE-GNN در عمل به یک طبقه‌بند ساده تبدیل شدند که همه نمونه‌ها را در کلاس «سخت‌قانون» قرار می‌داد. همچنین، آستانه‌های بهینه تصمیم‌گیری این دو مدل به کمترین مقدار ممکن رسید؛ موضوعی که نشان می‌دهد هیچ الگوی امتیازدهی معناداری برای تفکیک دو کلاس توسط این مدل‌ها شکل نگرفته است. دو مکانیزم مسئول این پدیده هستند:

الف) هموارسازی بیش‌ازحد در گرافی بزرگ و با گره‌های پردرجه: علت اصلی این رفتار را نمی‌توان صرفاً به عمق شبکه نسبت داد؛ بلکه نحوه ترکیب اطلاعات گره با اطلاعات همسایگان نقش تعیین‌کننده‌ای در آن دارد. در GCN، انتشار اطلاعات از طریق ماتریس نرمال‌سازی متقارن $\tilde{D}^{-1/2}\tilde{A}\tilde{D}^{-1/2}$ انجام می‌شود. درنتیجه، گره‌هایی با درجه بالا (که در این گراف تا ۵۰۰۷ نیز می‌رسد) هنگام به‌روزرسانی نمایش خود، اطلاعاتشان را با میانگین تعداد بسیار زیادی از همسایگان ترکیب می‌کنند [۲۲]،[۳۸]. این فرآیند باعث می‌شود ویژگی‌های متمایزکننده گره به‌تدریج کمرنگ شود و پس از تنها دو مرحله انتشار، نمایش برداری گروه‌ها تا حد زیادی به یکدیگر شبیه شود. در چنین شرایطی، توانایی مدل در تفکیک کلاس‌ها از بین می‌رود؛ پدیده‌ای که با عنوان هموارسازی بیش‌ازحد در GCN های به‌کاررفته روی گراف‌های متراکم شناخته می‌شود[۳۸]،[۳۹]. پیامد نهایی این وضعیت نیز گرایش مدل به پیش‌بینی کلاس غالب برای همه نمونه‌ها است. در مقابل، روش تجمیع قابل‌یادگیری مبتنی بر الحاق در گراف‌سیج، که راه‌حل توصیه‌شده در این پژوهش است [۸]، دچار این فروپاشی نمی‌شود.

ب) ناکافی بودن سیگنال رابطه‌ای در CARE-GNN: معماری CARE-GNN برای بهره‌گیری همزمان از دو نوع رابطه، یعنی رابطه دوبخشی و رابطه فرافکنی، طراحی‌شده است. بااین‌حال، ازآنجاکه گراف فرافکنی تنها ۶۹ یال دارد، یکی از مهم‌ترین منابع اطلاعاتی این مدل عملاً فاقد محتوای مؤثر است. از سوی دیگر، فرآیند تجمیع مبتنی بر فیلتر شباهت در رابطه دوبخشی نیز تحت تأثیر همان پدیده هموارسازی بیش‌ازحد قرار می‌گیرد و درنتیجه، توان مدل برای استخراج الگوهای متمایزکننده کاهش می‌یابد [٤]. بنابراین، حتی انتخاب زیرمجموعه‌ای از همسایگان نیز نمی‌تواند اطلاعات کافی برای جداسازی دو کلاس فراهم کند.

درمجموع، این نتایج را نباید به‌عنوان نشانه‌ای از ضعف ذاتی GCN یا CARE-GNN تفسیر کرد، بلکه بیشتر بیانگر آن است که در گراف‌های دوبخشی بسیار بزرگ و نامتوازن، گراف‌سیج به دلیل نحوه حفظ اطلاعات گره و شیوه تجمیع همسایگان، پایداری و توان تفکیک بالاتری از خود نشان می‌دهد [٤]،[٢٢].



شکل (۱۲): ۲۰ ویژگی برتر ازنظر اهمیت در جنگل تصادفی که سه ویژگی غالب هستند

٤-۱۱ خلاصه یافته‌ها

آشکارساز گراف‌سیج پیشنهادی تحت یک پروتکل استقرایی دقیق بدون هیچ‌گونه بیش‌برازش به عملکرد مطلق بسیارخوبی دست می‌یابد (صحت ۰/۹۸۲، F۱ = ۰/۹۵٤، AUC = ۰/۹۹۸۷) و در بین بذرها پایدار است. بااین‌حال، مهم‌ترین و صادقانه‌ترین یافته دارای ظرافت‌هایی است: بر روی این معیار مصنوعی، ساختار گراف هنوز مزیت قابل‌اندازه‌گیری نسبت به مدل‌های جدولی بسیار ارزان‌تر ارائه نمی‌دهد، زیرا برچسب، طبق طراحی، یک تابع قطعی از چند ویژگی گره ایستا است که به‌طور مستقل توسط مطالعه حذف، تحلیل‌های اعتبارسنجی توزیع و اهمیت ویژگی‌ها و آزمون غیرمعنادار ویلکاکسون تأیید شد. علاوه‌بر این، GCN و CARE-GNN تحت گراف دوبخشی بزرگ با درجه بالا و یک رابطه فرافکنی خالی فرومی‌پاشند (بخش ٤-۱۰). این نتایج مسیر روشنی را به‌سوی سهم قوی‌تر گراف ترسیم می‌کنند: بازتعریف برچسب به‌گونه‌ای که به ویژگی‌های همسایگی/زمانی بستگی داشته باشد (مانند نسبت ربات‌ها در میان اعضای مشترک، یا خطر ناشی از فوران) و غنی‌سازی گراف فرافکنی (یک آستانه ژاکارد پایین‌تر، یا هم‌عضویتیِ تجمیع‌شده در تمام تصاویر زمانی)، تا اطلاعات رابطه‌ای و زمانی به‌جای اینکه مازاد باشند، ضروری گردند [٤]، [٢٤]، [٤٧]، [٤٨]. زیرساخت برای چنین آزمایشی در حال حاضر در پایپ‌لاین وجود دارد.

۵- بحث

هدف این بخش، تفسیر نتایج در چارچوب یادگیری گراف و تبیین دامنه کاربرد رویکرد پیشنهادی است. مهم‌ترین یافته پژوهش، برتری مدل‌های جدولی با هزینه محاسباتی حدود ۳۵۰ تا ۶۰۰۰۰ برابر کمتر نسبت به گراف‌سیج ازنظر F1 است، نتیجه‌ای که بیش از آنکه ناشی از ضعف GNN ها باشد، به نحوه تعریف برچسب‌ها و ماهیت داده‌ها بازمی‌گردد.

۵-۱ پارادوکس مدل پیچیده در برابر مدل ساده

برچسب گروه‌های سخت‌قانون، همان‌گونه که در بخش ۳-۳-٤ بیان شد، از امتیاز ریسکی مبتنی بر PCA روی پنج ویژگی ایستای گروه (نرخ سکوت، ناهنجاری تعامل، رکود رشد، کم‌مشارکتی و فعالیت ربات) استخراج‌شده است. بنابراین، برچسب به توپولوژی گراف، ویژگی‌های همسایگان یا روابط بین‌گروهی وابسته نیست. درنتیجه، مدل‌های جدولی با دسترسی مستقیم به این ویژگی‌ها تمام سیگنال لازم برای طبقه‌بندی را در اختیاردارند و انتشار پیام در گراف اطلاعات جدیدی تولید نمی‌کند. این نتیجه با مطالعه حذفی (بخش ٤-٦)، اعتبارسنجی توزیع (بخش ٤-۹) و آزمون ویلکاکسون (بخش ٤-۷) نیز تأیید شد.

بااین‌حال، این نتیجه به همه مسائل تشخیص گروه‌های ناهنجار قابل‌تعمیم نیست. بر اساس این پژوهش و مطالعات پیشین، GNN ها در شرایط زیر برتری دارند [۸]، [۲۲]، [۲۳]، [۲٤]:

الف) روابط توپولوژیک غنی: زمانی که برچسب به الگوی ارتباط میان گروه‌ها، مانند اعضای مشترک یا شبکه‌های هماهنگ، وابسته باشد و این اطلاعات در ویژگی‌های محلی وجود نداشته باشد [٤]، [۲٤].

ب) ضعف ویژگی‌های ایستا: در گروه‌های تازه‌تأسیس یا گروه‌هایی که ویژگی‌های آماری آن‌ها به‌صورت عمدی عادی‌سازی شده است، ساختار همسایگی می‌تواند تنها منبع اطلاعات تفکیک‌کننده باشد.

ج) پویایی‌های زمانی: در صورت دسترسی به مسیرهای زمانی انتشار محتوا یا فعالیت‌های هماهنگ، GNN های فضایی–زمانی قادر به مدل‌سازی هم‌زمان وابستگی‌های ساختاری و زمانی هستند [۱۸]، [۲۰]، [۲۱]، [٤۷]، [٤۸].

د) گراف‌های ناهمگن و چندرابطه‌ای: این مدل‌ها می‌توانند اهمیت انواع مختلف روابط را به‌صورت خودکار بیاموزند، درحالی‌که مدل‌های جدولی به مهندسی دستی ویژگی‌ها وابسته‌اند [۱۷]، [۲۳].

ه) تعمیم استقرایی گراف‌سیج: به دلیل یادگیری تابع تجمیع، توانایی نمایش گره‌های دیده‌نشده را نیز دارد و برای سامانه‌های پویایی مانند تلگرام مناسب‌تر است [۸].

بنابراین، یافته این پژوهش به معنای برتری مطلق مدل‌های جدولی یا ناکارآمدی GNN ها نیست، بلکه نشان می‌دهد زمانی که برچسب مستقیماً از ویژگی‌های ایستای گره مشتق شود و گراف اطلاعات ساختاری محدودی داشته باشد، مدل‌های جدولی انتخاب مناسب‌تری هستند؛ در مقابل، در مسائل واقعی که برچسب از روابط ساختاری، هم‌عضویتی و پویایی‌های زمانی ناشی می‌شود، GNN ها همچنان مزیت روش‌شناختی خود را حفظ می‌کنند.

۵-۲ محدودیت‌های روش‌شناختی و دامنه اعتبارسنجی

تمامی آزمایش‌ها بر روی مجموعه‌داده‌ای شبیه‌سازی‌شده انجام‌شده‌اند. اگرچه توزیع‌های مصنوعی بر اساس ویژگی‌های گزارش‌شده برای شبکه‌های اجتماعی واقعی طراحی‌شده‌اند [۲۹]، [۳۲]، این داده‌ها همچنان فاقد برخی خصوصیات مهم شبکه‌های واقعی ازجمله نویزهای توپولوژیک، ساختارهای پیچیده اجتماع‌محور که به‌طور کامل با مدل بلوک تصادفی بازتولید نمی‌شوند و رفتارهای تطبیقی مهاجمان هستند [۲۹]، [٤]. ازاین‌رو، تعمیم نتایج به شبکه‌های واقعی تلگرام باید با احتیاط انجام شود و کمبود داده‌های واقعی برچسب‌خورده همچنان یکی از چالش‌های اصلی این حوزه است.

هم‌چنین، تعریف برچسب بر پایه PCA (بخش ۳-۳-٤) مستقیماً بر ویژگی‌های ایستای گروه استوار است. هرچند این انتخاب کنترل آزمایش و تکرارپذیری را افزایش می‌دهد، دامنه اعتبار نتایج را نیز به مسائلی محدود می‌کند که برچسب با ویژگی‌های محلی گره همسو باشد. در کاربردهای واقعی، ممکن است گروه‌های سخت‌قانون تنها از طریق الگوهای رفتاری هماهنگ، وابستگی‌های زمانی یا ارتباط با گروه‌های پرخطر قابل‌شناسایی باشند؛ در چنین شرایطی انتظار می‌رود مزیت مدل‌های جدولی کاهش‌یافته و GNN های فضایی–زمانی عملکرد بهتری ارائه دهند.

٦- نتیجه‌گیری و محدودیت‌ها

این پژوهش با موفقیت رویکردی مبتنی بر شبکه‌های عصبی گرافی، به‌ویژه گراف‌سیج را برای شناسایی گروه‌های سخت‌قانون که با رفتارهای اجباری و رشد غیرطبیعی مشخص می‌شوند، توسعه داد. به دلیل کمبود داده‌های واقعی و برچسب‌دار، یک پایپ‌لاین تولید داده‌های مصنوعی ایجاد شد که ساختارهای پیچیده شبکه‌های اجتماعی واقعی، الگوهای رفتاری متنوع، پویایی‌های زمانی و روابط دوبخشی کاربر–گروه را با دقت مدل‌سازی می‌کند. ارزیابی‌ها نشان داد که مدل پیشنهادی گراف‌سیج با دستیابی به‌دقت فراخوانی ۹۳/۷٪ و معیار AUC برابر با ۹۹/۸۷٪ در یک محیط کاملاً استقرایی، عملکرد قدرتمندی از خود نشان می‌دهد و نسبت به شبکه‌های گرافی دیگری نظیر GCN و CARE-GNN که در مواجهه با گراف‌های متراکم و دارای درجات بالا دچار مشکل هموارسازی بیش‌ازحد می‌شوند، بسیار مقاوم‌تر است.

بااین‌وجود، تحلیل‌های دقیق مقایسه‌ای و مطالعه حذف نشان داد که در چارچوب فعلی، ویژگی‌های ایستای گروه، غالب‌ترین سیگنال‌های تشخیص را فراهم می‌کنند؛ درنتیجه مدل‌های کلاسیک

جدولی مانند XGBoost با صرف هزینه محاسباتی بسیار پایین‌تر توانستند نتایجی رقابتی و حتی اندکی بهتر ارائه دهند. اگرچه استفاده از داده‌های مصنوعی امکان کنترل دقیق، تحلیل مکانیزم‌ها و تکرارپذیری کامل را فراهم کرد، اما از محدودیت‌های اصلی این رویکرد عدم امکان تأیید مستقیم با داده‌های پلتفرم‌های پیام‌رسان واقعی است. تحقیقات آینده باید بر توسعه برچسب‌ها و گراف‌های فرافکنی متمرکز شوند که وابستگی بیشتری به تعاملات ساختاری و پویایی‌های زمانی دارند، تا ارزش‌افزوده شبکه‌های عصبی گرافی در شناسایی رفتارهای هماهنگ پنهان و ناهنجار به‌وضوح برجسته‌تر گردد.

References

[1] D. Javed, N. Z. Jhanjhi, N. A. Khan, S. K. Ray, A. Al Mazroa, F. Ashfaq, and S. R. Das, "Towards the future of bot detection: A comprehensive taxonomical review and challenges on Twitter/X," Computer Networks, vol. 254, Art. no. 110808, 2024.

[2] A. Beutel, W. Xu, V. Guruswami, C. Palow, and C. Faloutsos, "CopyCatch: Stopping group attacks by spotting lockstep behavior in social networks," in Proc. 22nd Int. World Wide Web Conf. (WWW), Rio de Janeiro, Brazil, 2013, pp. 119–130.

[3] A. Buitrago López et al., "Agent-based simulation of online social networks and disinformation," arXiv preprint arXiv:2512.22082, 2025.

[4] Y. Dou, Z. Liu, L. Sun, Y. Deng, H. Peng, and P. S. Yu, "Enhancing graph neural network-based fraud detectors against camouflaged fraudsters," in Proc. 29th ACM Int. Conf. on Information and Knowledge Management (CIKM), 2020, pp. 315–324.

[5] A. Davies and N. Ajmeri, "Realistic synthetic social networks with graph neural networks," arXiv preprint arXiv:2212.07843, 2022.

[6] N. Jiang, F. Yin, B. Wang, and A. T. Crooks, "A large-scale geographically explicit synthetic population with social networks for the United States," Scientific Data, vol. 11, Art. no. 1204, 2024, doi: 10.1038/s41597-024-03970-1.

[7] N. Jiang, A. T. Crooks, H. Kavak, A. Burger, and W. G. Kennedy, "A method to create a synthetic population with social networks for geographically-explicit agent-based models," Computational Urban Science, vol. 2, Art. no. 7, 2022, doi: 10.1007/s43762-022-00034-1.

[8] W. L. Hamilton, R. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in Advances in Neural Information Processing Systems (NeurIPS), 2017, pp. 1024–1034.

[9] A. Bojchevski, O. Shchur, D. Zügner, and S. Günnemann, "NetGAN: Generating graphs via random walks," in Proc. 35th Int. Conf. on Machine Learning (ICML), PMLR, vol. 80, 2018, pp. 610–619.

[10] J. You, R. Ying, X. Ren, W. L. Hamilton, and J. Leskovec, "GraphRNN: Generating realistic graphs with deep auto-regressive models," in Proc. 35th Int. Conf. on Machine Learning (ICML), PMLR, 2018, pp. 5708–5717.

[11] A. Zareie et al., "Identifying coordination in online social networks through anomalous sharing behaviour," Online Social Networks and Media, vol. 50, Art. no. 100341, 2025.

[12] R. Rogers and N. Righetti, "Coordinated inauthentic behaviour on Facebook? A typology of manufactured attention," Platforms & Society, vol. 2, Art. no. 29768624251369784, 2025.

[13] M. Cinelli et al., "Coordinated inauthentic behavior and information spreading on Twitter," Decision Support Systems, vol. 160, Art. no. 113819, 2022.

[14] L. Luceri et al., "Coordinated inauthentic behavior on TikTok: Challenges and opportunities for detection in a video-first ecosystem," in Proc. Int. AAAI Conf. on Web and Social Media, vol. 20, no. 1, 2026.

[15] F. Cinus et al., "Exposing cross-platform coordinated inauthentic activity in the run-up to the 2024 US election," in Proc. ACM Web Conf. 2025, 2025.

[16] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," arXiv preprint arXiv:1710.10903, 2017.

[17] Y. Liu et al., "Heterogeneous graph convolutional network for rumor detection with multi-level interactive fusion and graph reconstruction," Scientific Reports, vol. 15, no. 1, Art. no. 31639, 2025.

[18] S. Shehnepoor et al., "Spatio-temporal graph representation learning for fraudster group detection," IEEE Trans. Neural Netw. Learn. Syst., vol. 35, no. 5, pp. 6628–6642, 2022.

[19] M. Boyapati and R. Aygun, "BalancerGNN: Balancer graph neural networks for imbalanced datasets: A case study on fraud detection," Neural Networks, vol. 182, Art. no. 106926, 2025.
[20] D. Saldaña-Ulloa, G. De Ita Luna, and J. R. Marcial-Romero, "A temporal graph network algorithm for detecting fraudulent transactions on online payment platforms," Algorithms, vol. 17, no. 12, Art. no. 552, 2024.
[21] Z. Qu et al., "Temporal enhanced multimodal graph neural networks for fake news detection," IEEE Trans. Comput. Soc. Syst., vol. 11, no. 6, pp. 7286–7298, 2024.
[22] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in Proc. Int. Conf. on Learning Representations (ICLR), 2017.
[23] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 1, pp. 4–24, 2021.
[24] X. Ma, J. Wu, S. Xue, J. Yang, C. Zhou, Q. Z. Sheng, H. Xiong, and L. Akoglu, "A comprehensive survey on graph anomaly detection with deep learning," IEEE Transactions on Knowledge and Data Engineering, vol. 35, no. 12, pp. 12012–12038, 2023.
[25] H. Qiao, H. Tong, B. An, I. King, C. Aggarwal, and G. Pang, "Deep graph anomaly detection: A survey and new perspectives," arXiv preprint arXiv:2409.09957, 2024.
[26] B. Khemani, S. Patil, K. Kotecha, and S. Tanwar, "A review of graph neural networks: concepts, architectures, techniques, challenges, datasets, applications, and future directions," Journal of Big Data, vol. 11, Art. no. 18, 2024.
[27] M. Fey and J. E. Lenssen, "Fast graph representation learning with PyTorch Geometric," in ICLR Workshop on Representation Learning on Graphs and Manifolds, 2019.
[28] N. Patki, R. Wedge, and K. Veeramachaneni, "The synthetic data vault," in Proc. IEEE Int. Conf. on Data Science and Advanced Analytics (DSAA), 2016, pp. 399–410.
[29] P. W. Holland, K. B. Laskey, and S. Leinhardt, "Stochastic blockmodels: First steps," Social Networks, vol. 5, no. 2, pp. 109–137, 1983.
[30] I. T. Jolliffe, Principal Component Analysis, 2nd ed. New York, NY: Springer, 2002.
[31] W. Chen, K. Yang, Z. Yu, Y. Shi, and C. L. P. Chen, "A survey on imbalanced learning: latest research, applications and future directions," Artificial Intelligence Review, vol. 57, no. 6, Art. no. 137, 2024.
[32] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks," Annual Review of Sociology, vol. 27, pp. 415–444, 2001.
[33] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, Time Series Analysis: Forecasting and Control, 5th ed. Hoboken, NJ: Wiley, 2015.
[34] S. Sahu, "A Fast Parallel Approach for Neighborhood-based Link Prediction by Disregarding Large Hubs," arXiv preprint arXiv:2401.11415, 2024.
[35] H. H. Rashidi, S. Albahra, B. Hu, and B. P. Rubin, "A novel and fully automated platform for synthetic tabular data generation and validation," Scientific Reports, vol. 14, 2024.
[36] P. A. Apellániz, A. Jiménez, B. Arroyo Galende, J. Parras, and S. Zazo, "Synthetic Tabular Data Validation: A Divergence-Based Approach," IEEE Access, vol. 12, 2024.
[37] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," Journal of Machine Learning Research, vol. 15, pp. 1929–1958, 2014.
[38] Q. Li, Z. Han, and X.-M. Wu, "Deeper insights into graph convolutional networks for semi-supervised learning," in Proc. 32nd AAAI Conf. on Artificial Intelligence, 2018, pp. 3538–3545.
[39] K. Oono and T. Suzuki, "Graph neural networks exponentially lose expressive power for node classification," in Proc. Int. Conf. on Learning Representations (ICLR), 2020.
[40] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
[41] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proc. 22nd ACM SIGKDD Int. Conf. on Knowledge Discovery and Data Mining (KDD), 2016, pp. 785–794.
[42] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in Proc. Int. Conf. on Learning Representations (ICLR), 2015.
[43] T. Saito and M. Rehmsmeier, "The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets," PLoS ONE, vol. 10, no. 3, Art. no. e0118432, 2015.
[44] [44] J. Davis and M. Goadrich, "The relationship between precision-recall and ROC curves," in

Proc. 23rd Int. Conf. on Machine Learning (ICML), 2006, pp. 233–240.

[45] A. Foucart, A. Elskens, and C. Decaestecker, "Ranking the scores of algorithms with confidence," in Proc. ESANN 2025.

[46] J. Demšar, "Statistical comparisons of classifiers over multiple data sets," Journal of Machine Learning Research, vol. 7, pp. 1–30, 2006.

[47] Y. Zheng, L. Yi, and Z. Wei, "A survey of dynamic graph neural networks," Frontiers of Computer Science, vol. 19, no. 6, Art. no. 196323, 2025.

[48] E. Rossi, B. Chamberlain, F. Frasca, D. Eynard, F. Monti, and M. Bronstein, "Temporal graph networks for deep learning on dynamic graphs," in ICML Workshop on Graph Representation Learning, 2020.

Journal of Applied And Basic Machine Intelligence Research (ABMIR)

Applied and Basic Machine Intelligence Research

Vol. 4, No. 2, Autumn & Winter 2026, pp. 89 - 112

Detection of Hard-Law Groups in Telegram Messenger Using GraphSAGE Graph Neural Networks and Simulated Data with Temporal Features

Maedeh Arab Bafrani[1], Mohammad Ali Zare Chahooki[2*]

[1]Ph.D. Candidate, Department of Software Engineering, Faculty of Computer Engineering, Yazd University, Yazd, Iran

[2]Associate Professor, Department of Software Engineering, Faculty of Computer Engineering, Yazd University, Yazd, Iran

Article Information

Original Research Paper

Received:
2026 April 01

Accepted:
2026 August 18

Keywords:
Hard-law groups, Telegram messenger, Graph neural networks, GraphSAGE, Simulated dataset, Temporal features, Coercive behavior detection, Graph-based deep learning

Corresponding Author*:
chahooki@yazd.ac.ir

Abstract

This study addresses the detection of 'Hard-law' groups in the Telegram messaging platform, defined as groups that manipulate their organic growth by requiring users to perform mandatory actions, such as inviting additional users or joining specific channels, as a prerequisite for joining or continuing participation. Detecting such groups is essential for preserving platform integrity and mitigating abusive behaviors. The scarcity of labeled real-world data, the dynamic temporal nature of user behavior, and the need to jointly exploit structural and behavioral signals constitute the primary challenges in this domain. To address these challenges, a large-scale simulated dataset comprising 100,000 users, 10,000 groups, and 12 temporal snapshots was generated, capturing both static and temporal features of users and groups. Group labels were assigned using a principal component analysis (PCA)-based risk score, and a user–group bipartite graph was constructed. A GraphSAGE-based graph neural network was then trained using a fully inductive evaluation protocol, a weighted binary cross-entropy loss function, and decision threshold optimization, and its performance was compared with six baseline models. On an independent test set, the proposed model achieved an accuracy of 98.2٪, a positive-class precision of 97.2٪, a recall of 93.7٪, an F1-score of 95.4٪, an area under the ROC curve (AUC) of 99.87٪, and an average precision (AP) of 99.51٪, while demonstrating stable performance across all random seeds. The results indicate that, because the labels are structurally dependent on static group attributes, tabular machine learning models achieve slightly better performance with lower computational cost, whereas GraphSAGE remains robust on large-scale, high-degree graphs. These findings have been validated only on simulated data, and their generalization to real-world Telegram data remains an open challenge due to the lack of labeled datasets.

: 10.22034/ABMIR.2026.24432.1231

E-ISSN: 2821-2037