یازدهمین همایش سالانه
بانکداری نوین و
نظام‌های پرداخت
پژوهشکده پولی و بانکی
بانک مرکزی جمهوری اسلامی ایران
بانک مرکزی جمهوری اسلامی ایران
شرکت ملی انفورماتیک
زیست بوم بانکداری هوشمند؛ حکمرانی دیجیتال و نظارت فنّاورانه
تهران، مرکز همایش‌های بین‌المللی برج میلاد ـ ۵ و ۶ آذر ۱۴۰۳ ISC

شناسایی تراکنش‌های مشکوک پول‌شویی با استفاده از یادگیری بازنمایی گراف[۱]

زهرا نوراله[۲] محسن رضوانی[۳]

چکیده

پول‌شویی به عنوان یکی از جرائم مالی بزرگ، چالشی جدی برای سیستم‌های مالی در دهه‌های اخیر بوده است. با توجه به پیچیدگی و پویایی این فعالیت‌ها، شناسایی و جلوگیری از آن نیازمند روش‌های پیشرفته‌ای است. در این پژوهش، با استفاده از یک چارچوب نوآورانه در سه مرحله، شناسایی زیرشبکه‌های مشکوک در تراکنش‌های مالی به منظور کشف پول‌شویی مورد بررسی قرار می‌گیرد. در مرحله اول، پیش‌پردازش و مدل‌سازی داده‌های تراکنش انجام شده تا گراف جهت‌دار تراکنش‌های مالی ایجاد گردد. در مرحله دوم، با استفاده از الگوریتم‌های تشخیص انجمن، ابرگره‌های حساب مشتریان مشخص می‌شود. بر اساس ماتریس عضویت و روابط عضویت حساب‌های مشتریان در انجمن‌ها، گراف «گروه تراکنش» تولید می‌شود. در ادامه ویژگی‌های ساختاری، جریان مالی و تعبیه زیرگراف متناظر با هریک از انجمن‌ها برای گراف گروه تراکنش به دست می‌آید، تا با استفاده از ویژگی‌های استخراج شده و الگوریتم‌های یادگیری عمیق، زیرشبکه‌های مشکوک در گراف گروه تراکنش بر اساس الگوهای آموخته شده در مرحله سوم شناسایی گردد. اثربخشی مدل پیشنهادی، با استفاده از داده‌های تراکنش مالی از بانک آنلاین رومانیایی لیبرا، با شناسایی فعالیت‌های پول‌شویی مشکوک در یک دوره سه ماهه، مورد ارزیابی قرار گرفت. نتایج به دست آمده نشان‌دهنده بهبود قابل توجهی نسبت به پژوهش‌های پیشین است، بهره‌گیری از الگوریتم‌های شناسایی انجمن و رویکردهای مبتنی بر قوانین برای تحلیل و ایجاد گراف گروه تراکنش، به منظور شناسایی رفتارهای گروهی و سازمان‌یافته مشکوک در تراکنش‌های بانکی، از نوآوری‌های این پژوهش می‌باشد. این دستاوردها می‌تواند به بهبود امنیت و کاهش ریسک‌های مرتبط با فعالیت‌های غیرقانونی در سیستم‌های مالی کمک کند.

واژه‌های کلیدی: یادگیری بازنمایی گراف، تشخیص ناهنجاری، پول‌شویی
طبقه‌بندی JEL: G10, G18, G19

---

[۱] مقاله از طرح پیشنهادی رساله دکتری تخصصی با عنوان «شناسایی زیرشبکه‌های ناهنجار در تراکنش‌های مالی با استفاده از یادگیری بازنمایی گراف» استخراج شده است.

[۲] بهسازان ملت و دانشجوی دکتری مهندسی کامپیوتر، هوش مصنوعی دانشگاه صنعتی شاهرود؛ enznoor@gmail.com
[۳] دانشیار دانشکده مهندسی کامپیوتر، دانشگاه صنعتی شاهرود؛ mrezvani@shahroodut.ac.ir

۱ مقدمه

پول‌شویی به فعالیت‌های مربوط به پنهان کردن درآمد واقعی، فرار مالیاتی، یا تبدیل پول به دست آمده غیرقانونی برای استفاده مشروع اشاره دارد [۲, ۱]. شناسایی فعالیت‌های غیرعادی و مشکوک در تراکنش‌های بانکی از جمله چالش‌های کلیدی در حوزه مبارزه با پول‌شویی محسوب می‌شود. دلیل اصلی این چالش، ماهیت پنهان و پیچیدگی فعالیت‌های پول‌شویی است، که با استفاده از روش‌هایی مانند تغییر الگوهای تراکنش‌های مالی، استفاده از شبکه‌های مخفی و شرکت‌های پوششی می‌باشد که منشأ وجوه غیرقانونی را پنهان می‌کنند. لذا با توجه به رشد روزافزون فعالیت‌های پول‌شویی و تغییرات مداوم در الگوهای آن، ابزارهای تشخیصی نیازمند بهبود مداوم هستند. در این راستا شناسایی زیرشبکه‌های مشکوک می‌تواند به مؤسسات مالی کمک کند تا فعالیت‌های غیرقانونی مانند تقلب، اختلاس، یا فعالیت‌های مرتبط با جرایم سازمان‌یافته را متوقف کنند. پیشرفت‌های اخیر در تشخیص ناهنجاری، به‌ویژه در حوزه تجزیه‌وتحلیل گراف مبتنی بر تکنیک‌های یادگیری عمیق، راه را برای استراتژی‌های مؤثرتر در کشف این شبکه‌های پیچیده مالی غیرقانونی هموار کرده است.

در بیشتر آثار موجود، اقدامات هر حساب بانکی به‌عنوان رفتار هویتی مستقل بررسی می‌شود، بدون در نظر گرفتن تعاملات مالی در سطح گروهی از مشتریان. این روش ممکن است به شناسایی ناهنجاری‌های فردی کمک کند، اما توانایی کافی برای کشف الگوهای پیچیده‌تر که ممکن است در شبکه‌های تراکنشی گسترده‌تر وجود داشته باشند، ندارد. در پژوهش‌های اخیر، نشان داده شده است که در شبکه تراکنش‌های مالی بین حساب‌های مشتریان، مجموعه‌ای از حساب‌ها می‌توانند به‌عنوان گروهی از مشتریان مشکوک شناسایی شوند [۳] این یافته‌ها بر اهمیت تحلیل تراکنش‌های مالی در سطح گروهی تأکید می‌کنند. با الهام از این مشاهدات، در این پژوهش به شناسایی زیرشبکه‌ای از شبکه معاملات حساب‌ها پرداخته می‌شود که به‌عنوان زیرشبکه معاملات مالی مشکوک شناخته می‌شود. این روش به‌جای تمرکز صرف بر رفتارهای فردی، به تعاملات مالی میان حساب‌ها توجه می‌کند و از این طریق قادر به شناسایی الگوهای ناهنجار و پیچیده‌تری است که در شبکه‌های مالی گسترده‌تر پدیدار می‌شوند. این رویکرد نه‌تنها دقت شناسایی ناهنجاری‌ها را افزایش می‌دهد، بلکه به کشف و مهار فعالیت‌های مشکوک و غیرقانونی مانند تقلب و پول‌شویی در سیستم‌های مالی کمک شایانی می‌کند.

در مدل پیشنهادی، از ویژگی ساختار انجمن که یکی از مهم‌ترین ویژگی‌های هر گراف می‌باشد، استفاده شده است. این بدان معنی است که رفتار معامله یک مشتری با سایر مشتریان با ویژگی‌های مشابه مقایسه می‌شود. این مقایسه با گروه همتا برای تشخیص پول‌شویی بسیار مهم است زیرا یک تراکنش ممکن است با توجه به سابقه مشتری غیرعادی باشد اما با توجه به حرکات رایج گروه با ویژگی‌های مشابه، عادی باشد [۴]. شناسایی انجمن‌ها به تشخیص ناهنجاری‌ها کمک می‌کند زیرا به ما اجازه می‌دهد تا الگوها و ارتباطات طبیعی را درک کرده و سپس مواردی را که از این الگوها منحرف می‌شوند، شناسایی کنیم. جنبه‌های اصلی در این پژوهش عبارتند از:

- استفاده از ویژگی‌های استخراج شده از تعبیه زیرگراف‌های شناسایی شده و جریان مالی در آنها برای شناسایی زیرشبکه معاملات مالی مشکوک

- شناسایی رفتارهای غیرمعمول با استفاده از ویژگی ساختار انجمن که با تحلیل داده‌ها و یافتن انجمن‌های معمول، می‌توان رفتارهایی را که از این الگوها منحرف می‌شوند، شناسایی کرد.
- شناسایی زیرشبکه‌ای از شبکه معاملات حساب‌ها که به عنوان زیرشبکه معاملات مالی مشکوک شناسایی می‌گردد.
- مقایسه مدل پیشنهادی با مدل‌های موجود در تحقیقات اخیر.

پس از مقدمه، ساختار این مقاله به شرح زیر در نظر گرفته شده است. مراحل پول‌شویی و روش‌های یادگیری بازنمایی گراف بیان می‌شود. کارهای مرتبط قبلی به همراه تکنیک‌ها و روش‌های استفاده شده برای تشخیص ناهنجاری مبتنی بر گراف در بخش سوم مقاله تشریح می‌شود. دادگان و مراحل آماده‌سازی داده‌ها برای ایجاد گراف گروه تراکنش، برای مطالعه موردی در بخش پنجم شرح داده می‌شود. همچنین نتایج اصلی مدل پیشنهادی و مقایسه آنها با نتایج مدل‌های موجود در بخش ششم نشان داده می‌شود. در نهایت، بخش هفتم نتیجه‌گیری و تحقیقات آتی را عنوان می‌کند.

۲ مفاهیم اولیه و پیشینه پژوهش

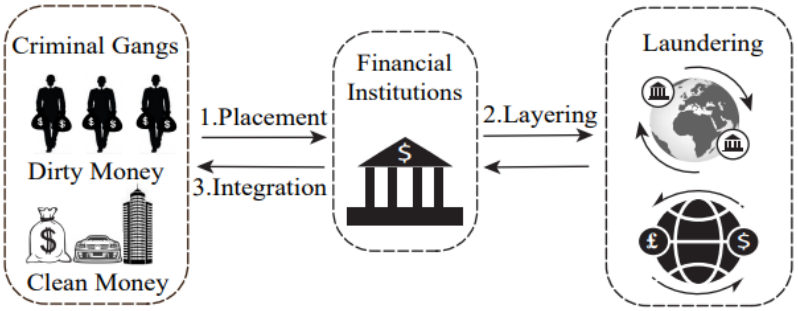
در طول چند دهه گذشته، تحقیقات در مورد شناسایی ناهنجاری به دلیل پیامدهای این رویدادها در طیف گسترده‌ای از رشته‌ها مانند امنیت، امور مالی و پزشکی، مورد توجه فزاینده‌ای قرار گرفته است. به همین دلیل برای جلوگیری از حوادث مخربی مانند کلاهبرداری مالی، نفوذ در شبکه و هرزنامه‌های اجتماعی، طیف گسترده‌ای از تکنیک‌های تشخیص ناهنجاری، از جمله روش‌های آماری، الگوریتم‌های یادگیری ماشین و مدل‌های یادگیری عمیق، توسعه یافته است. تشخیص ناهنجاری نیازمند توسعه تکنیک‌های مؤثری است که بتواند بر عوامل مختلفی که در وقوع ناهنجاری‌ها مؤثر است تأثیرگذار باشد. این تکنیک‌ها باید مقیاس‌پذیر، قابل تفسیر و سازگار با شرایط در حال تغییر باشند تا در کاربردهای دنیای واقعی مؤثر باشند [۵]. تشخیص ناهنجاری‌ها به دلیل عوامل مختلفی که می‌توانند بر وقوع آن‌ها تأثیرگذار باشند، مانند ماهیت داده‌ها، در دسترس بودن داده‌های برچسب‌گذاری‌شده، نوع ناهنجاری‌هایی که باید شناسایی شوند و حوزه برنامه، یک مشکل چالش‌برانگیز است. ناهنجاری‌ها می‌توانند اشکال مختلفی داشته باشند بنابراین تشخیص آن‌ها از رفتار عادی دشوار است. همچنین ناهنجاری‌ها می‌توانند در هر سطحی از انتزاع، از نقاط داده‌ای منفرد گرفته تا کل سیستم‌ها رخ دهند و می‌توانند ناشی از طیف وسیعی از عوامل از جمله خطای انسانی، اشکالات نرم‌افزاری، خرابی‌های سخت‌افزاری و فعالیت‌های مخرب باشند [۵,۶].

تحقیقات مرتبط با تشخیص ناهنجاری در شبکه تراکنش‌های مالی، به بررسی روش‌های متنوعی پرداخته است که این روش‌ها را می‌توان به سه دسته اصلی تقسیم کرد: روش‌های سنتی، روش‌های یادگیری، و روش‌های یادگیری ترکیبی. روش‌های سنتی با بررسی ویژگی‌های ساختاری داده‌ها از طریق تحلیل‌های آماری و ریاضیاتی، با استفاده از توزیع‌های آماری، ناهنجاری‌ها را شناسایی می‌کنند. روش‌های یادگیری از الگوریتم‌های یادگیری ماشین و هوش مصنوعی برای تشخیص الگوهای پیچیده‌تر در داده‌ها استفاده می‌کنند، و با بهره‌گیری از اتوانکدرها و شبکه‌های عصبی کانولوشن گرافی، می‌توانند ناهنجاری‌ها را تشخیص دهند. در نهایت، روش‌های ترکیبی و یادگیری عمیق نمایش گراف، از تلفیق اصول روش‌های سنتی و یادگیری عمیق بهره می‌برند، تا با استفاده از شبکه‌های عصبی

کانولوشن گرافی و شبکه‌های عصبی گرافی، به نحوی مؤثرتر، ناهنجاری‌ها را تشخیص دهند. در این پژوهش، بررسی روش‌های یادگیری نمایش گراف مورد توجه می‌باشد.

۱.۲ مراحل پول‌شویی

پول‌شویی به فعالیت‌های مربوط به پنهان کردن درآمد واقعی، فرار مالیاتی، یا تبدیل پول به دست آمده غیرقانونی برای استفاده مشروع اشاره دارد [۱]. فرآیند پول‌شویی از سه بخش اصلی تشکیل شده است. این مراحل شامل جانمایی و قرار دادن، لایه‌بندی و یکپارچه‌سازی می‌باشد [۲]. شکل ۱ طرح معمولی پول‌شویی را نشان می‌دهد [۷].



شکل ۱ . مراحل پول‌شویی

جانمایی، مرحله اولیه طرح پول‌شویی است که هدف اصلی آن قرار دادن پول غیرقانونی در سیستم‌های مالی است، بدون گرفتار شدن در محدودیت‌های قانونی. دو روش برای دستیابی به این هدف وجود دارد: سپرده اولیه و سپرده ثانویه. در سپرده اولیه، تلاش برای دور زدن بازرسان قانونی است، مثلاً با واریز مقدار کمی پول به صورت مکرر یا تأسیس بانک‌ها در کشورهایی با مقررات کمتر. در سپرده ثانویه، پول‌ها را از طریق خدمات و سرویس‌هایی که پول نقد دریافت می‌کنند به حساب‌های بانکی اشخاص حقوقی واریز می‌کنند. سپس این اشخاص حقوقی به عنوان صاحب حساب، پول‌های دریافتی را برای فرد مبدأ باز می‌گردانند. بدین صورت منشأ پول، غیرقابل‌شناسایی می‌گردد.

لایه‌بندی، مرحله دوم فعالیت‌های پول‌شویی است. در این مرحله که به انتقال مخفیانه نیز معروف است، افراد یا سازمان‌های پول‌شویی سعی می‌کنند پول‌های کثیف به دست آمده از فعالیت‌های غیرقانونی را به شکلی که ردپای فعالیت‌های غیرقانونی آن‌ها پنهان شود، به حساب‌های مختلف یا مکان‌های مختلف جهت استفاده آتی انتقال دهند. این مرحله اغلب شامل چندین فعالیت مختلف است که هدف آن‌ها پنهان کردن مبدأ واقعی پول‌ها و ایجاد مراحلی پیچیده برای پیگیری آن‌ها است. در این مرحله، پول از طریق یک شبکه پیچیده از تراکنش‌ها منتقل می‌شود که به طور مستقیم به پول اصلی متصل نیستند. همچنین، این تراکنش‌ها ممکن است به خریدوفروش برخی دارایی‌ها نیز مرتبط باشند. در این مرحله تقسیم پول به بخش‌های کوچک باعث سخت شدن تشخیص آن می‌شود. همچنین، استفاده از حساب‌های خارج از کشور در معاملات پول‌شویی به دلیل عدم وجود قوانین نظارتی مشترک بین ارگان‌های مرتبط نیز رایج است و اعمال نظارت بر آن‌ها را دشوار می‌کند.

مرحله نهایی در فرآیند پول‌شویی، یعنی یکپارچه‌سازی، هدف آن پاکسازی پولی است که از فعالیت‌های غیرقانونی جمع‌آوری شده است. این پول‌ها ممکن است برای خرید سهام، خرید ملک یا سایر دارایی‌های مشابه استفاده شوند. اگر این مرحله با موفقیت انجام شود، پول شسته شده ممکن است بدون شناسایی توسط آژانس‌های مالی استفاده می‌شود [۷].

۲.۲ نگاشت و یادگیری نمایش گراف

برای گراف‌های بزرگ با میلیاردها گره، نمایش گراف با چالش‌هایی مانند پیچیدگی محاسباتی بالا، قابلیت موازی‌سازی کم و عدم امکان استفاده از روش‌های یادگیری ماشین در آنالیز و پردازش گراف مواجه است [۸]. برای بهینه‌سازی تحلیل اطلاعات گراف‌های بزرگ، نحوه نمایش و بازنمایی داده‌های گراف بسیار اهمیت دارد. در سال‌های اخیر، تلاش‌های زیادی در زمینه بازنمایی گراف انجام شده است. در فرایند بازنمایی گراف، هدف اصلی تبدیل رئوس و یال‌ها به یک فضای برداری است تا ویژگی‌ها و ارتباطات بین موجودیت‌ها به صورت عددی قابل استفاده باشند. ایده‌ی اصلی روش‌های تعبیه گراف این است که گره‌های نزدیک به هم در گراف در فضای برداری نیز نزدیک به هم هستند و ارتباط هندسی گره‌ها به هم در فضای برداری هم نزدیک به هم هستند. تعبیه گراف به چندین شکل انجام می‌شود، از جمله تعبیه گره، تعبیه یال، تعبیه زیر گراف و تعبیه کل گراف. همچنین، از شبکه‌های عصبی گراف[1] و تعاملات گرافی[2] برای مدل‌سازی تعاملات بین گره‌ها استفاده می‌شود، که اغلب برای پیش‌بینی ارتباطات و کشف الگوهای ارتباطی به کار می‌روند، و می‌توانند برای یادگیری بازنمایی گراف استفاده شوند [۹].

۳.۲ تحلیل گراف با رویکرد تشخیص انجمن و یادگیری زیرگراف

بیشتر بانک‌ها بر اساس تجربه و درک خود از فعالیت‌های مالی غیرقانونی، قوانینی را تنظیم نموده‌اند تا به کمک هشدارهای مبتنی بر قوانین تنظیم شده، بتوانند به نقل‌وانتقالات مالی مشکوک اشاره کنند [۱۰]. به همین منظور مؤسسات مالی سیستم‌هایی را برای شناسایی تراکنش‌های مشکوک بر اساس قوانین ثابت و مشروط به آستانه‌های خاص توسعه داده‌اند. این رویکرد دارای نقطه‌ضعف، تولید نسبت بالایی از مثبت کاذب است زیرا سیستم نمی‌تواند به صورت پویا قانون را بر اساس تغییرات رفتاری در مجرمان تنظیم کند [۱۱،۱۲]. راجپوت و همکاران [۱۳] در سال ۲۰۱۴ یک سیستم خبره مبتنی بر آنتولوژی پیشنهاد کردند. این سیستم با استفاده از اطلاعات آنتولوژیکی که درباره موجودیت‌ها، ارتباطات و ویژگی‌های مختلف در زمینه تراکنش‌های مالی مشکوک تعریف شده‌اند، به تحلیل دقیق و شناسایی الگوهای غیرمعمول و فعالیت‌هایی که ممکن است به فساد مالی یا کیفری منجر شود، می‌پردازد. ابزارهای AML مبتنی بر قوانین می‌توانند در شناسایی تراکنش‌های مشکوک مؤثر باشند، اما در مواجهه با طرح‌های پیچیده پول‌شویی ممکن است با شکست مواجه شوند. [۱۰]

در تشخیص ناهنجاری در تراکنش‌های بانکی، تمرکز بر روی تراکنش‌های گروهی، نه‌تنها امکان بهبود تشخیص ناهنجاری‌های پیچیده‌تر را فراهم می‌کند، بلکه به ارتقای امنیت و پیشگیری از کلاهبرداری‌های مالی کمک می‌نماید.

---

[1] Graph Neural Networks – GNNs

[2] Graph Interactions

سوج و همکاران در سال ۲۰۱۶، یک سیستم تشخیص ناهنجاری برای شناسایی فعالیت‌های پول‌شویی توصیف می‌کنند که به‌جای تمرکز بر افراد و تاریخچه تراکنش‌ها، بر روی تحلیل و تجزیه رفتار گروه‌ها تمرکز دارد. این سیستم برای شناسایی فعالیت‌های نامتعارف و مشکوک، از یک ترکیب استفاده می‌کند که شامل تجزیه‌وتحلیل شبکه و یادگیری نظارت شده است. به عبارت دیگر، به جای استفاده از روش‌های مبتنی بر افراد، این سیستم به بررسی رفتار و تعاملات گروه‌ها می‌پردازد تا میزان ناهنجاری و وجود فعالیت‌های غیرمعمول را شناسایی کند، که این امر می‌تواند در شناسایی فعالیت‌های پول‌شویی که ممکن است تنها به‌صورت گروهی مشاهده شوند، مؤثر باشد. عملکرد این سیستم بدین صورت می‌باشد که جوامع کوچک و معنی‌دار را در شبکه شناسایی و استخراج می‌کند و از دانش کسب‌وکار موجود برای بهبود فرآیند تشخیص استفاده می‌کند. متعاقباً، تکنیک‌های یادگیری تحت نظارت برای آموزش طبقه‌بندی‌کننده‌ها در این جوامع شناسایی‌شده اعمال می‌شوند. این رویکرد به سیستم اجازه می‌دهد تا به طور مؤثر بین فعالیت‌های عادی و مشکوک تمایز قائل شود و در نتیجه توانایی خود را برای شناسایی فعالیت‌های پول‌شویی افزایش دهد [۱۴].

در پژوهش دیگری که توسط چنج و همکاران در سال ۲۰۲۳ انجام شد، تشخیص زیرگراف گروه تراکنش کاربر در حوزه تشخیص فعالیت‌های پول‌شویی مورد بررسی قرار گرفته است. یک رویکرد مبتنی بر یادگیری گراف عمیق آگاهانه گروهی برای شناسایی پول‌شویی سازمان‌یافته پیشنهاد می‌کند. در این رویکرد یک رمزگذار انجمن محور برای نمایش گره‌ها و ویژگی‌ها در گراف تراکنش‌های کاربر و شناسایی رفتارهای گروهی به هم پیوسته معرفی می‌کند. همچنین، از یک طرح بهبود محلی برای تجمیع گره‌هایی با ویژگی‌های تراکنش مشابه در گروه‌ها برای بهبود دقت تشخیص استفاده می‌کند [۷].

در پژوهش چن و همکاران در سال ۲۰۲۱ یک مدل سلسله‌مراتبی GAGNN برای پیش‌بینی کیفیت هوای شهر در سراسر کشور، با استفاده از شبکه عصبی گراف، بر اساس گروه‌بندی شهرها پیشنهاد شده است. این مدل از گراف شهر و گراف گروه شهر برای مدل‌سازی وابستگی‌های فضایی و نهان بین شهرها استفاده می‌کند. با استفاده از GAGNN، وابستگی‌های پنهان در بین شهرها شناسایی شده و گروه‌های شهری تشکیل می‌شوند. سپس یک ماژول رمزگذاری همبستگی گروهی برای یادگیری همبستگی‌های بین گروه‌های شهر ارائه شده است. پس از ساخت گراف، GAGNN مکانیزم ارسال پیام را برای مدل‌سازی وابستگی‌های بین شهرها و گروه‌های شهر پیاده‌سازی می‌کند. این پژوهش از نقاط قوت شبکه‌های کانولوشن گراف و میدان‌های تصادفی مارکوف برای دستیابی به تشخیص انجمن نیمه نظارت شده استفاده می‌کند و بهبودهایی در شبکه‌های نسبت داده شده[1] ارائه می‌دهد. گروه‌هایی از گره‌ها را با هم پردازش می‌کند و باعث کاهش محاسبات و سربار حافظه می‌شود. با اعمال نفوذ در تعاملات گروهی، بازنمایی گره‌های غنی‌تر را به تصویر می‌کشد. پویایی گروه و رفتارهای جمعی را در برنامه‌هایی مانند شبکه‌های اجتماعی یا سیستم‌های توصیه بهتر به تصویر می‌کشد. با کاهش پیچیدگی محاسباتی از طریق پردازش گروهی، در نمودارهای بزرگ، مقیاس بهتری دارد. تعریف گروه‌های معنی‌دار از گره‌ها می‌تواند چالش‌برانگیز باشد و به دانش حوزه نیاز دارد. پیاده‌سازی GAGNN در مقایسه با شبکه‌های عصبی گراف سنتی پیچیدگی می‌افزاید، که می‌تواند بر طراحی و

[1] attributed networks

یکپارچگی تأثیر بگذارد. اثربخشی به شدت به تعریف دقیق و اعمال نفوذ گروه‌های مرتبط بستگی دارد. بهینه‌سازی مدل‌های GAGNN ممکن است نیاز به تنظیم پارامترهای پیچیده پردازش گروهی داشته باشد که به طور بالقوه زمان‌های آموزشی را افزایش می‌دهد [۱۵].

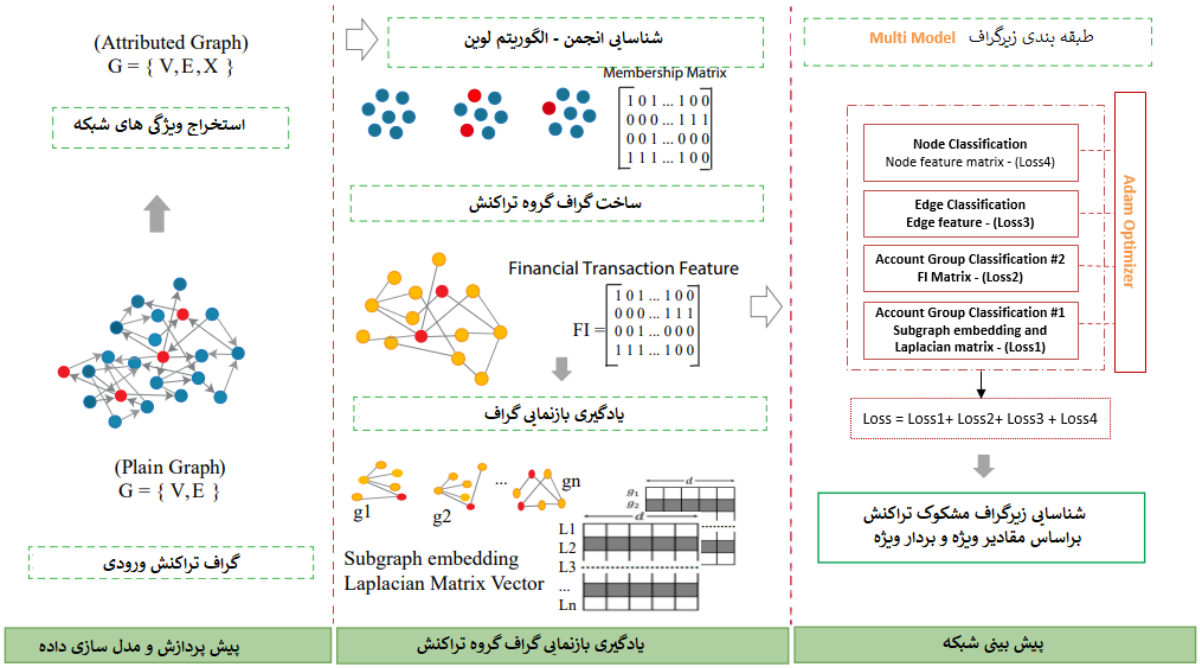
در پژوهشی که توسط سان و همکاران در سال ۲۰۲۱ انجام شد، شبکه عصبی زیرگراف SUGAR برای وظایف طبقه‌بندی گراف ارائه شده است. این مدل با استفاده از ترکیب مکانیزم تقویتی و اطلاعات متقابل خود نظارتی برای انتخاب و جاسازی زیرگراف‌های مهم از شیوه سلسله مراتبی استفاده می‌کند. مدل ارائه شده یاد می‌گیرد به جای اینکه کل گراف را به عنوان یک واحد در نظر بگیرد، زیرگراف‌های مهم را به شیوه‌ای سلسله مراتبی انتخاب و جاسازی کند. این مدل از یک هدف یادگیری خود نظارتی استفاده می‌کند که اطلاعات متقابل بین جاسازی‌های زیرگراف و نمایش کلی گراف را به حداکثر می‌رساند. این به مدل کمک می‌کند تا ویژگی‌های زیرگراف متمایز و قابل‌انتقال بیشتری را بیاموزد. ارزیابی‌های آن‌ها نشان می‌دهند که مدل SUGAR عملکرد بهتری نسبت به روش‌های دیگر شبکه عصبی گراف در طبقه‌بندی گراف دارد. یکی از مزیت‌های SUGAR توانایی آن در افزایش تفسیرپذیری مدل‌های طبقه‌بندی گراف با تمرکز بر الگوهای سطح زیرگراف است. با بازسازی گراف بر اساس زیرگراف‌های برجسته انتخاب شده، SUGAR بینش‌هایی را در مورد عناصر ساختاری خاصی که به تصمیم‌گیری‌های طبقه‌بندی کمک می‌کنند، ارائه می‌دهد، در نتیجه درک واضح‌تری از چگونگی تأثیر ویژگی‌های گراف بر نتایج ارائه می‌دهد [۱۶].

۳ مدل پیشنهادی و یافته‌های پژوهش

در این پژوهش، به بررسی و شناسایی زیرگراف‌های ناهنجار در تراکنش‌های مالی پرداخته می‌شود. تراکنش‌های مالی به دلیل حجم بالا و پیچیدگی‌های موجود در الگوهای تراکنشی، مستعد وقوع فعالیت‌های ناهنجار و غیرقانونی مانند تقلب و پول‌شویی هستند. شناسایی این ناهنجاری‌ها می‌تواند نقشی اساسی در بهبود امنیت و سلامت سیستم‌های مالی داشته باشد. روش‌های سنتی شناسایی ناهنجاری‌ها اغلب قادر به کشف الگوهای پیچیده و پنهان در داده‌های تراکنشی نیستند. در این پژوهش، با استفاده از شبکه‌های عصبی گراف (GNNs)، به عنوان یک ابزار قدرتمند در پردازش و تحلیل داده‌های ساختاریافته، رویکردی نوین برای شناسایی زیرگراف‌های ناهنجار ارائه می‌شود. شبکه‌های عصبی گراف به دلیل توانایی در یادگیری از ساختار گراف و ویژگی‌های گره‌ها و یال‌ها، قابلیت بالایی در شناسایی الگوهای ناهنجار پیچیده در داده‌های تراکنشی دارند.

معماری مدل پیشنهادی شامل سه لایه، پیش‌پردازش و مدل‌سازی داده‌های تراکنش‌های مالی، لایه یادگیری بازنمایی گروه تراکنش شناسایی شده و لایه پیش‌بینی شبکه می‌باشد که در شکل ۲ نشان داده شده است. لایه اول، پیش‌پردازش و مدل‌سازی داده‌های تراکنش می‌باشد. در این لایه، ویژگی‌های ساختاری از پیش تعریف‌شده مانند درجه گره، مسیرها و سایر ویژگی‌های ساختاری شبکه استخراج می‌شود. در لایه دوم، با استفاده از گراف جهت‌دار و الگوریتم‌های تشخیص انجمن، ابرگره‌های حساب مشتریان مشخص می‌شود. با استفاده از گراف جهت‌دار و الگوریتم‌های تشخیص انجمن، ابرگره‌های حساب مشتریان مشخص می‌شود. بر اساس ماتریس عضویت و روابط عضویت حساب‌های مشتریان در انجمن‌ها، گراف «گروه تراکنش» تولید می‌شود. در ادامه ویژگی‌های ساختاری،

جریان مالی و جاسازی زیرگراف متناظر با هریک از انجمن‌ها برای گراف گروه تراکنش به دست می‌آید، تا با استفاده از ویژگی‌های استخراج شده و الگوریتم‌های یادگیری عمیق، زیرشبکه‌های مشکوک در گراف گروه تراکنش بر اساس الگوهای آموخته شده شناسایی گردد.



شکل ۲ . معماری مدل پیشنهادی شناسایی ناهنجاری در سطح زیرگراف

در لایه سوم، پیش‌بینی شبکه، زیر شبکه مشکوک بر اساس معماری شبکه عصبی شناسایی می‌گردد بدین صورت که گراف گروه حساب مشتریان (گروه تراکنش مالی) براساس برچسب‌های عادی و غیرعادی آموزش داده شده تا گروه حساب‌هایی که معاملات مشکوک دارند شناسایی شود. پس از آن در گام آخر در این لایه، براساس مقادیر ویژه و بردار ویژه برای گره‌های هر زیرگراف (گروه حساب مشتریان)، تراکنش‌هایی که با امتیاز پایین هستند از زیرگراف مربوطه حذف می‌گردد تا زیرگراف مشکوک نهایی شناسایی شود.

برای اثربخشی مدل پیشنهادی، داده‌های تراکنش مالی از بانک آنلاین رومانیایی لیبرا [۱۷]، برای شناسایی فعالیت‌های پول‌شویی مشکوک در یک دوره سه ماهه، استفاده گردید.

۱.۳ مجموعه دادگان تراکنش‌های مالی لیبرابانک

بانک اینترنتی لیبرا، یک بانک رومانیایی است که فهرست کاملی از تراکنش‌ها را در طول سه ماه ارائه کرده است که در این پژوهش از آن استفاده شده است [۱۷]. مشخصات این تراکنش‌ها در جدول ۱ نشان داده شده است. تراکنش‌های بانکی (نقل‌وانتقالات بین حساب‌ها) به صورت گراف ترسیم شده است. گره، حساب مشتری و لبه‌ها تراکنش‌های تجمیع شده هستند. چندین مرحله پیش‌پردازش نیز قبل از مدل‌سازی شبکه تراکنش انجام شده است. این مراحل شامل متعادل کردن تراکنش‌های برچسب‌گذاری شده با مجموعه داده تراکنش‌ها، ادغام موارد تکراری که

مبدأ و مقصد تراکنش یکسان بوده است. کنار گذاشتن سایر تراکنش‌ها مانند پرداخت‌هایی که از طریق کارت بانکی بوده است. مجموعه دادگان از توزیع نرمال پیروی می‌کنند. شبکه جهت‌دار ایجاد شده از این دادگان به شرح جدول ۲ می‌باشد.

جدول ۱

مشخصات مجموعه دادگان تراکنش‌های لیبرابانک

ویژگی مجموعه دادگان	
تعداد تراکنش	۴۵۵۸۸۰۵
میانگین ارزش معامله	۳۲۶۴
میانه ارزش معامله	۱۵۲
حداکثر ارزش معامله	۴۲ میلیون
حداقل ارزش معامله	۰/۰۱

گره‌های شبکه ایجاد شده، شامل حساب‌های بانکی می‌باشد که ۳۸۵۱۰۰ گره در این شبکه ترسیم شده است. ۵۹۷۱۶۵ لبه در این شبکه ایجاد شده است، که نشان‌دهنده نقل‌وانتقال بین دو حساب می‌باشد. میانگین درجه ورودی و خروجی ۱/۵۵ می‌باشد. تراکنش‌ها توسط خبرگان واحد AML برچسب‌گذاری شده است. در این بخش دو برچسب استفاده شده است. اولین برچسب «برچسب هشدار» است و براساس مجموعه قوانین ایجاد می‌گردد و به این معنی می‌باشد که تراکنش مربوطه باید توسط افراد متخصص بررسی گردد. دومین برچسب، «برچسب گزارش» می‌باشد و مشخص می‌کند که سطح سوءظن به اندازه کافی بالاست که باید برای بررسی به مقامات دولتی گزارش شود. (باید توجه داشت که یک گزارش لزوماً نشان‌دهنده جرم نیست.)

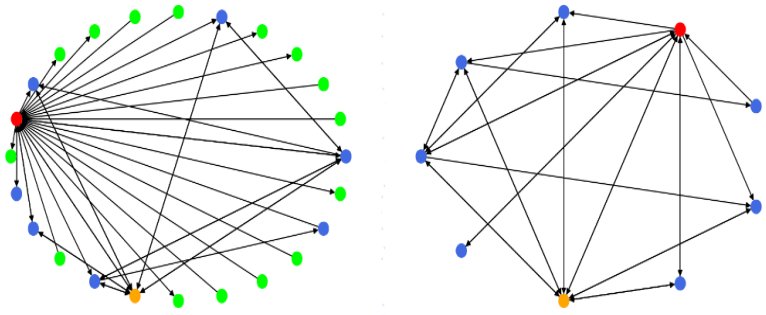
جدول ۲

ویژگی برچسب دادگان تراکنش‌های لیبرابانک

ویژگی برچسب دادگان	
تعداد تراکنش با برچسب هشدار	۵۱۷
تعداد تراکنش با برچسب گزارش	۱۱
تعداد گره‌های دارای برچسب هشدار	۶۰۰
تعداد گره‌های دارای برچسب گزارش	۱۵

پس از آماده‌سازی داده‌ها، یک گراف جهت‌دار $G = (V, E)$ ایجاد می‌شود. مجموعه گره‌ها (رأس) از لیست معاملات استخراج می‌شود که شماره حساب مشتریان می‌باشد. لبه‌های E نشان‌دهنده انتقال بین دو حساب می‌باشد. مجموع مبلغ منتقل شده و تعداد معاملات به عنوان وزن لبه‌ها در نظر گرفته شده است. وزن ناهنجاری برای هر گره برای هشدارها، به عنوان تعداد تراکنش‌هایی که به عنوان هشدار برچسب‌گذاری شده‌اند، تعریف می‌شود و برای

گزارش‌ها، نیز برابر با تعداد تراکنش‌هایی که برچسب «گزارش» دارند و گره مورد نظر در آن تراکنش شرکت داشته است، در نظر گرفته شده است. به عنوان مثال تعداد هشدار برای یک گره مانند a، برابر ۲۲ و گزارش برابر ۳ می‌باشد، یعنی این حداکثر هشدارهای این گره برابر ۲۲ و حداکثر تعداد گزارش برای این گره ۳ می‌باشد. در این پژوهش ناهنجاری یک هشدار و یا یک گزارش می‌باشد. ویژگی‌ها با استفاده از روش‌های ابتکاری عنوان شده در پژوهش انجام شده توسط دومیترسکو و همکاران [۱۷] از گراف استخراج می‌شوند.



شکل ۳. نمونه‌ای از اگونت و اگونت کاهش‌یافته

شکل ۳ مفاهیم اگونت و اگونت کاهش‌یافته را برای یک مورد واقعی نشان می‌دهد. گره قرمز مرکز اگونت است و تراکنش بین آن و گره نارنجی، مشکوک است. اگونت کاهش‌یافته با حذف گره‌های سبز به دست می‌آید. مجموعه‌ای از ویژگی‌های مشتق شده از اگونت و اگونت کاهش‌یافته، مانند چگالی لبه، مقادیر انتقال‌یافته، درجه گره مرکزی، امکان تشخیص ناهنجاری خوب را با استفاده از روش‌های یادگیری ماشین بدون نظارت فراهم می‌کند.

گراف نهایی شامل ۲۰۲۴۲۶ گره و ۳۷۰۹۱۸ یال می‌باشد. این گراف از اگونت‌های کاهش‌یافته در مرحله پیش‌پردازش حاصل شده است. برای ایجاد دو گراف مجزا ابتدا داده‌های گره به نسبت ۳۰ به ۷۰ برای ایجاد گراف آموزش و آزمون در نظر گرفته شد. تعداد ۱۶۱۵۹۲ گره با برچسب عادی و ۳۴۸ گره با برچسب مشکوک برای ایجاد گراف آموزش در نظر گرفته شد. همچنین تعداد ۴۰۳۸۶ گره با برچسب عادی و ۱۰۰ گره با برچسب مشکوک برای ایجاد گراف آزمون در نظر گرفته شد. ویژگی‌های دو گراف آموزش و آزمون در جدول ۳ نشان داده شده است.

جدول ۳

ویژگی گراف‌های آموزش و آزمون

ویژگی گراف آموزش	تعداد گره با برچسب مشکوک	۳۳۵
	تعداد یال‌ها (تراکنش)	۲۳۰۹۰۷
	تعداد گره با برچسب عادی	۱۳۵۴۴۷
ویژگی گراف آزمون	تعداد گره با برچسب مشکوک	۷۴
	تعداد یال‌ها (تراکنش)	۱۶۲۲۲
	تعداد گره با برچسب عادی	۱۲۹۶۳

نتایج و یافته‌های پژوهش برای تشخیص ناهنجاری در دو سطح گره و زیرگراف با روش‌های سنتی و یادگیری عمیق در ادامه تشریح می‌گردد. رویکرد پیشنهادی توسط سه متریک ارزیابی می‌گردد. اولین متریک AUC[1] می‌باشد که به عنوان سطح زیر منحنی ROC[2] تعریف می‌شود. این معیار نشان‌دهنده توانایی مدل در تفکیک بین کلاس‌ها است و مقداری بین ۰ تا ۱ دارد، که عدد بزرگتر نشان‌دهنده عملکرد بهتر است. دومین متریک، دقت است. نسبت تعداد نمونه‌های درست تشخیص داده‌شده به کل نمونه‌های تشخیص داده‌شده است.

$$\frac{TP}{TP+FP} = Percision \quad (۱)$$

$$\frac{TP}{TP+FN} = Recall \quad (۲)$$

Recall یا TPR[3] نسبت تعداد نمونه‌های مثبت واقعی که به درستی تشخیص داده شده‌اند به کل تعداد نمونه‌های مثبت واقعی این معیارها کمک می‌کنند تا عملکرد مدل‌های یادگیری ماشین در تفکیک و تشخیص دسته‌ها ارزیابی شود. از آنجایی که نتایج کشف تراکنش‌های مشکوک پول‌شویی برای مؤسسات مالی حیاتی است، به طور کلی به نرخ دقت بالایی نیاز است. مقدار recall بالا نشان‌دهنده این است که مدل توانایی خوبی در شناسایی ناهنجاری‌ها دارد.

متریک سوم، معیار F-score است که یک معیار ترکیبی است که Precisionو recall را ترکیب می‌کند تا عملکرد کلی یک مدل در تشخیص ناهنجاری را ارزیابی کند. معیار F-score میانگین هارمونیکی از دقت و recall است و به صورت رابطه ۳ محاسبه می‌شود:

$$\frac{Percision*Recall}{Percision+Recall} * 2 = F - Score \quad (۳)$$

مقدار معیار F-score بین ۰ و ۱ است، که مقادیر بالاتر نشان‌دهنده عملکرد بهتر مدل در تشخیص ناهنجاری هستند. در واقع، مقدار بالاتر از ۰/۵ برای F-score نشان‌دهنده یک مدل خوب در تشخیص ناهنجاری است، در حالی که مقدار ۱ نشان‌دهنده دقت و recall کامل است.

۲.۳ تشخیص ناهنجاری در سطح گره

استفاده از تکنیک‌های یادگیری ماشین کمک می‌کند تا دقت تشخیص ناهنجاری افزایش پیدا کند. برخی از روش‌های یادگیری ماشین که در پژوهش‌های پیشین مورد استفاده قرار گرفته است، برای تشخیص ناهنجاری در سطح گره برای تراکنش‌های مالی لیبرابانک استفاده شد تا اثربخشی روش پیشنهادی را برجسته نماید.

نتایج ارزیابی روش‌های یادگیری ماشین و یادگیری عمیق در جدول ۵ و مقایسه نتایج آن‌ها در شکل ۴ نشان داده شده است. روش‌های یادگیری عمیق مقدار AUC و Recallبهتری نسبت به روش‌های یادگیری ماشین نشان می‌دهند. این امر به دلیل قدرت بالاتر این روش‌ها در مدل‌سازی ارتباطات پیچیده و ویژگی‌های غیرخطی داده‌ها است. فراپارامترهای تنظیم شده برای مدل‌های یادگیری استفاده شده در جدول ۴ نشان داده شده است.

---

[1] Area Under the Curve

[2] receiver operating characteristic curve

[3] True Positive Rate

جدول ۴

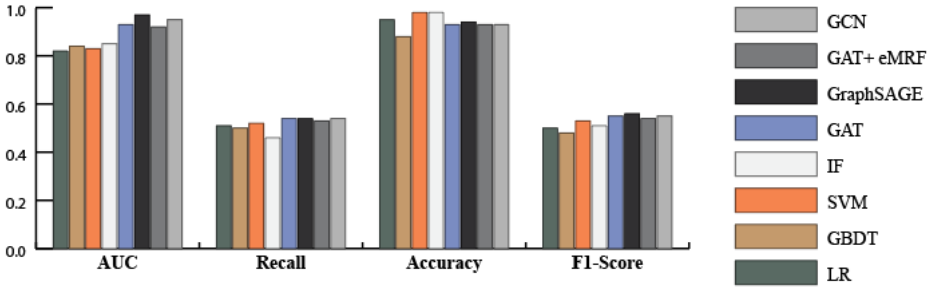
پارامترهای تنظیم شده برای مدل‌های یادگیری استفاده شده

مدل استفاده شده	پارامترهای تنظیم شده
Logestic Regresion (LR)	$LearningRate = 0.003$ momentum $= 0.002$, Batch size $= 128$
GBDT	$LearningRate = 0.002$, max depth $= 3$ estimators $= 100$, class weight $= 'balanced'$
SVM	class weight $= 'balanced'$, $kernel = 'linear'$
Isolation Forest	contamination $= 0.05$, class weight $= 'balanced'$
GCN	$LearningRate = 0.01$,epoch $= 100$ Hidden Feature $= 128$, weight $= 'balanced'$
GraphSAGE	$LearningRate = 0.01$,epoch $= 100$ Hidden Feature $= 100$,weight $= 'balanced'$
GAT + eMRF	$LearningRate = 0.05$,epoch $= 50$ Hidden Feature $= 16$, weight $= 'balanced'$

جدول ۵

نتایج ارزیابی تشخیص ناهنجاری با روش‌های یادگیری ماشین و یادگیری عمیق

مدل	AUC	Recall	Accuracy	Precision	F1-score
LR	۰/۸۲	۰/۵۱	۰/۹۵	۰/۸۲	۰/۵۰
GBDT	۰/۸۴	۰/۵۰	۰/۸۸	۰/۸۴	۰/۴۸
SVM	۰/۸۳	۰/۵۲	۰/۹۸	۰/۸۳	۰/۵۳
Isolation Forest	۰/۸۵	۰/۴۶	۰/۹۸	۰/۸۵	۰/۵۱
GCN	۰/۹۵	۰/۵۴	۰/۹۳	۰/۹۵	۰/۵۵
GraphSAGE	۰/۹۷	۰/۵۴	۰/۹۴	۰/۹۷	۰/۵۶
GAT	۰/۹۳	۰/۵۴	۰/۹۳	۰/۹۳	۰/۵۵
GAT + eMRF	۰/۹۲	۰/۵۳	۰/۹۳	۰/۹۲	۰/۵۴



شکل ۴ . مقایسه نتایج ارزیابی تشخیص ناهنجاری در سطح گره با استفاده از یادگیری سنتی و عمیق

۳.۳ تشخیص ناهنجاری در سطح زیرگراف

در این پژوهش، تلاش شده است تا با استفاده از یک مدل تحلیلی پیشنهادی، زیرگراف‌های مشکوک به پول‌شویی در داده‌های تراکنش مالی شناسایی و تحلیل شوند. تراکنش‌های پول‌شویی به صورت پراکنده در کل گراف تراکنش توزیع نشده است و بین برخی از حساب‌های فردی متمرکز است. بنابراین با این فرضیه که ماهیت معاملات پول‌شویی به صورت گروهی می‌باشد، در پژوهش چنچ و همکاران [۷] در سال ۲۰۲۳ گره‌های موجود در گراف به دو دسته حساب‌های عادی و گروهی تقسیم می‌گردد. با توجه به اینکه روش سنتی هنگام تجزیه‌وتحلیل قادر به تشخیص رفتار تراکنش در سطح گروه در ابعاد بالا نمی‌باشد. سیاستی برای دستیابی به تجمیع گره‌ها و پیش‌بینی ناهنجاری در سطح گروه معرفی شده است که با چالش‌هایی همراه می‌باشد. برخی از این چالش‌ها به شرح زیر می‌باشد:

- همپوشانی گره‌ها در گروه‌ها در نظر گرفته نمی‌شود.
- با توجه به اینکه فرآیند پول‌شویی در برخی از موارد شامل مجموعه‌ای از تراکنش‌های عادی نیز می‌باشد، حساب‌هایی وجود دارند که با تراکنش عادی به حسابی که تراکنش‌های مشکوک دارد متصل می‌شود. لذا برخی از تراکنش‌ها و حساب‌هایی که عادی می‌باشند به اشتباه در گروه تراکنش‌های مشکوک قرار می‌گیرند.
- مواردی از حساب‌هایی که ممکن است در ارتباط با پول‌شویی باشند و در گروه مشکوک قرار نگرفته باشد نیز وجود دارد.
- مبنای تجمیع گره‌ها (حساب‌های مشکوک) تراکنش‌های مشکوک می‌باشد، لذا در صورت وجود تراکنش‌های عادی امکان تجمیع بهینه نمی‌باشد و زمان اجرای این مدل‌ها نیز افزایش می‌یابد.

در مرحله دوم طرح پیشنهادی، انجمن‌های هر یک از گراف‌های آموزش و آزمون شناسایی شدند. تعداد کل انجمن‌های شناسایی شده ۱۳۷۶۷ انجمن می‌باشد. برچسب هریک از انجمن‌های شناسایی شده براساس برچسب اعضای هریک از انجمن‌ها در نظر گرفته می‌شود. برای ایجاد گراف گروه‌تراکنش، پس از شناسایی انجمن‌ها برای هریک از گراف‌های آموزش و آزمون، مجدد گراف گروه‌تراکنش آموزش و گراف گروه‌تراکنش آزمون ایجاد گردید. مشخصات این دو گراف با تعداد انجمن‌های شناسایی شده در جدول ۶ نشان داده شده است.

جدول ۶

ویژگی گراف گروه تراکنش آموزش و آزمون

ویژگی گراف گروه تراکنش آموزش	تعداد گره (انجمن) با برچسب مشکوک	۲۲۷
	تعداد یال‌ها (تراکنش)	۸۳۱۹۴
	تعداد گره (انجمن) با برچسب عادی	۱۳۵۴۰
ویژگی گراف گروه تراکنش آزمون	تعداد گره (انجمن) با برچسب مشکوک	۵۷
	تعداد یال‌ها (تراکنش)	۴۶۶۴
	تعداد گره (انجمن) با برچسب عادی	۲۰۴۱

انجمن‌های شناسایی شده با سایز کمتر از ۱۰ تعداد گره مشکوک در آن‌ها بیشتر از یک مورد بوده است. ۳۰ درصد انجمن‌های شناسایی شده در آموزش و ۲۰ درصد انجمن‌های شناسایی شده در دادگان آزمون، سایز انجمن‌ها ۲ بوده و هر دو عضو آن مشکوک بوده است.

در دادگان آموزش ۹۰ درصد انجمن‌های شناسایی شده مشکوک، بین ۲ تا ۵۰ نود در آن عضو می‌باشد. همچنین ۶۰ درصد انجمن‌های مشکوک در حدود ۳۰ درصد اعضای آن‌ها مشکوک هستند و ۲۵ درصد انجمن‌های مشکوک بیشتر از ۵۰ درصد اعضای آن‌ها مشکوک می‌باشند.

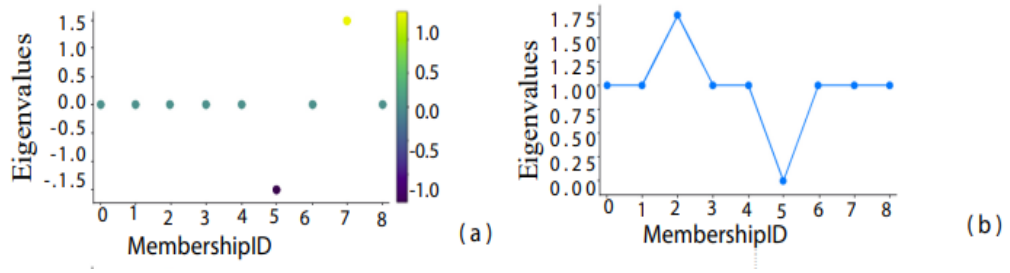
در این پژوهش سعی شد که مدل پیشنهادی با دو رویکرد با مدل پژوهش چنج و همکاران مقایسه شود. در رویکرد اول استفاده از استراتژی بهینه‌سازی برای آموزش و یادگیری مورد مقایسه قرار گرفت. در این رویکرد مدل M_1 مبتنی بر الگوریتم شبکه‌های$GAT + eMRF$، مدل M_2 مبتنی بر GAT و مدل M_3 بر اساس استراتژی بهینه‌سازی ترکیب شبکه‌های$GAT + eMRF$ و GAT می‌باشد، که باهم مقایسه شده است. نتایج در جدول ۷ نشان داده شده است. مدل M_3 در تشخیص گروه مشکوک، عملکرد بهتری داشته است. در رویکرد دوم، برای حل مسائل مربوط به چالش‌های عنوان شده به جای تجمیع گره‌ها براساس تراکنش مشکوک از الگوریتم دو مرحله‌ای لوین برای شناسایی انجمن‌ها استفاده گردید و براساس انجمن‌های شناسایی شده گراف گروه‌تراکنش ایجاد گردید. مدل M_4 مبتنی بر شبکه‌های$GAT + eMRF$ با این رویکرد با مدل M_1 و M_2 مقایسه شده است. استفاده از این روش علاوه بر اینکه چالش‌های فوق‌الذکر را برطرف می‌نماید، در دقت و امتیاز فراخوانی مشاهدات مشکوک بهتر عمل می‌کند.

جدول ۷

متریک‌های ارزیابی نتایج تشخیص ناهنجاری در سطح گروه تراکنش

مدل	AUC	Recall	Accuracy	Percision	F-score
$M_1(GAT + eMRF)$	۰/۸۳	۰/۵۰	۰/۸۱	۰/۸۲	۰/۵۰
$M_2\ (GAT)$	۰/۸۵	۰/۵۱	۰/۸۵	۰/۸۵	۰/۵۱
M_3(combined)	۰/۹۰	۰/۵۱	۰/۹۰	۰/۹۰	۰/۵۲
$M_4(Proposed\ Model)$	۰/۹۰	۰/۵۴	۰/۹۰	۰/۹۰	۰/۵۵

در زیرگراف‌هایی که با برچسب مشکوک طبقه‌بندی می‌شوند ممکن است حسابی که برچسب مشکوک دارد تراکنشی نیز داشته باشد که ماهیت آن عادی باشد. برای دقت در عملکرد مدل و حذف تراکنش‌های عادی از زیرگراف مشکوک، از مقادیر ویژه و بردار ویژه استفاده می‌شود. به کمک این ویژگی‌ها می‌توان الگوهای عادی را در داده‌ها شناسایی کرده و سپس انحرافات نسبت به این الگوها را در قالب امتیاز ناهنجاری تحلیل کرد. پس از مشخص کردن امتیاز ناهنجاری با استفاده از یک مقدار حد آستانه تراکنش‌هایی که امتیاز ناهنجاری کمتری نسبت به مقدار حد آستانه دارند از زیرگراف حذف می‌گردد تا در نهایت زیرگراف مشکوک شناسایی شود. به عنوان مثال در انجمن شماره ۳۷۶۴، که شامل ۹ عضو می‌باشد مقادیر ویژه و بردار ویژه در شکل ۵ نشان داده شده است مقدار حد آستانه ۰/۵ در نظر گرفته شده و ۸ عضو آن با امتیاز بالا به عنوان زیرگراف مشکوک نهایی در نظر گرفته می‌شود.



شکل ۵. مقادیر ویژه و بردار ویژه برای انجمن ۳۷۶۴

۴ جمع‌بندی و نتیجه‌گیری

با افزایش روزافزون فعالیت‌های پول‌شویی و تغییرات مستمر در الگوهای آن، ابزارهای تشخیصی باید بهبود یابند تا مؤسسات مالی بتوانند بهترین پاسخ را در برابر فعالیت‌های نامناسب مانند تقلب، اختلاس، و فعالیت‌های مرتبط با جرایم سازمان‌یافته ارائه دهند. شناسایی زیرشبکه‌های مشکوک می‌تواند به این مؤسسات کمک کند تا الگوهای نامتعارف در معاملات مالی را تشخیص داده و به‌طور مؤثری اقدامات پیشگیرانه انجام دهند. این اقدامات می‌توانند به حفظ اعتبار و امنیت مالی مشتریان و جامعه کمک کرده و از اثرات منفی فعالیت‌های غیرقانونی به‌طور قابل توجهی جلوگیری نمایند.

براساس این مشاهدات، رویکردی جدید برای شناسایی زیرشبکه‌های مشکوک در شبکه معاملات حساب‌های مشتریان بانک با استفاده از تکنیک‌های یادگیری عمیق و یادگیری بازنمایی گراف، پیشنهاد گردید. معماری مدل پیشنهادی برای تجزیه‌وتحلیل داده‌های تراکنش مالی از سه لایه یکپارچه تشکیل شده است. ابتدا، لایه پیش‌پردازش، ویژگی‌های ساختاری مانند درجه گره و مسیرهای شبکه را از داده‌های تراکنش استخراج می‌کند. در لایه دوم، با استفاده از رویکرد یادگیری ترکیبی یادگیری، ابتدا براساس معیار ماژولاریتی، ابرگره‌های حساب مشتریان مشخص می‌شود، سپس بر اساس ماتریس عضویت و روابط عضویت حساب‌های مشتریان در انجمن‌ها، گراف «گروه تراکنش» تولید می‌شود. این لایه همچنین با استفاده از تکنیک‌هایی یادگیری عمیق، یادگیری بازنمایی را بر روی انجمن‌های شناسایی‌شده برای جاسازی زیرگراف‌ها در فضای ویژگی انجام می‌دهد. در لایه سوم نیز از یک شبکه عصبی برای پیش‌بینی زیرشبکه‌های مشکوک در گراف گروه تراکنش استفاده می‌شود، تا فعالیت‌های مشکوک و بالقوه غیرقانونی

بر اساس الگوهای آموخته شده شناسایی گردد. در نهایت این پیش‌بینی‌ها، با ارزیابی مقادیر ویژه و بردارهای ویژه گره‌ها در زیرگراف‌ها با فیلتر کردن تراکنش‌هایی که دارای امتیاز ناهنجاری پایین هستند اصلاح می‌شوند و رفتار مشکوک برای بررسی بازرسان برجسته می‌گردد.

اثربخشی مدل پیشنهادی، با استفاده از داده‌های تراکنش مالی از بانک آنلاین رومانیایی لیبرا، با شناسایی فعالیت‌های پول‌شویی مشکوک در یک دوره سه ماهه، ارزیابی شد. نتایج تحقیقات انجام شده تاکنون، نشان می‌دهد که استفاده از الگوریتم‌های شبکه‌های عصبی و رویکردهای یادگیری عمیق نمایش گراف، توانایی تشخیص مشتریان بانکی درگیر در فعالیت‌های مشکوک پول‌شویی را نسبت به نمونه‌های مشابه بهبود می‌دهد. این دستاوردها می‌تواند به بهبود امنیت و کاهش ریسک‌های مرتبط با فعالیت‌های غیرقانونی در سیستم‌های مالی کمک کند.

۵ پیشنهادهای کاربردی

یکی از چالش‌های کلیدی، شناسایی فعالیت‌های غیرعادی و مشکوک در تراکنش‌های بانکی است. پیچیده‌تر شدن روش‌های تقلب، اهمیت شناسایی الگوها و زیرشبکه‌های مشکوک را بیش از پیش نشان می‌دهد. شناسایی زیرشبکه‌های مشکوک می‌تواند به مؤسسات مالی کمک کند تا فعالیت‌های غیرقانونی مانند تقلب، اختلاس، یا فعالیت‌های مرتبط با جرایم سازمان‌یافته را متوقف کنند. پیشرفت‌های اخیر در تشخیص ناهنجاری، به‌ویژه در حوزه تجزیه‌وتحلیل گراف مبتنی بر تکنیک‌های یادگیری عمیق، راه را برای استراتژی‌های مؤثرتر در کشف این شبکه‌های پیچیده مالی غیرقانونی هموار کرده است. لازم به ذکر است که رویکرد مبتنی بر تحلیل شبکه به عنوان جایگزینی برای سیستم‌های مبتنی بر قانون، نمی‌باشند بلکه به عنوان یک ابزار مکمل در کشف طرح‌های کلاهبرداری پیچیده‌تر یا جدید عمل می‌کنند.

فهرست منابع

[1] DUMITRESCU, Bogdan; BĂLTOIU, Andra; BUDULAN, Ştefania. Anomaly detection in graphs of bank transactions for anti money laundering applications. IEEE Access, 2022, 10. Jg., S. 47699-47714.

[2] Chopra S., Hadsell R., Le Cun Y. Learning a Similarity Metric Discriminatively, with Application to Face Verification; Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05); San Diego, CA, USA. 20–25 June 2005; pp. 539–546.

[3] Leman Akoglu, Hanghang Tong, and Danai Koutra. Graph based anomaly detection and description: a survey. Data mining and knowledge discovery, 29:626–688, 2015.

[4] JUN, Tang. A cross datasets referring outlier detection model applied to suspicious financial transaction discrimination. In: International Workshop on Intelligence and Security Informatics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006. S. 58-65.

[5] GUJRAL, Ekta. Survey: Anomaly Detection Methods. 2023.

[6] N. R. Prasad, S. Almanza-Garcia, and T. T. Lu, "Anomaly detection," Comput. Mater. Contin., vol. 14, no. 1, pp. 1–22, Jul. 2009, doi: 10.1145/1541880.1541882.

[7] CHENG, Dawei, et al. Anti-money laundering by group-aware deep graph learning. *IEEE Transactions on Knowledge and Data Engineering*, 2023.

[8] November. Anomaly detection with graph convolutional networks for insider threat and fraud detection. In *MILCOM 2019-2019 IEEE Military Communications Conference* (*MILCOM*) (pp. 109-114). IEEE.

[9] A. Goodge, B. Hooi, S.-K. Ng, and W. S. Ng, "Lunar: Unifying local outlier detection methods via graph neural networks," in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 36, no. 6, 2022, pp. 6737–6745.

[10] LUNA, Devendra Kumar, et al. Finding shell company accounts using anomaly detection. In: Proceedings of the ACM India Joint International Conference on Data Science and Management of Data. 2018. pp. 167-174.

[11] CHANDRADEVA, Lakshika Sammani, et al. Monetary transaction fraud detection system based on machine learning strategies. In: *Fourth International Congress on Information and Communication Technology: ICICT 2019, London, Volume 1*. Springer Singapore, 2020. S. 385-396.

[12] MOUSTAFA, Tamer Hossam, et al. Anti money laundering using a two-phase system. *Journal of Money Laundering Control*, 2015, 18. Jg., Nr. 3, S. 304-329.

[13] RAJPUT, Quratulain, et al. Ontology based expert-system for suspicious transactions detection. *Computer and Information Science*, 2014, 7. Jg., Nr. 1, S. 103.

[14] Savage, D., Wang, Q., Chou, P., Zhang, X. and Yu, X., 2016. Detection of money laundering groups using supervised learning in networks. *arXiv preprint arXiv:1608.00708.*

[15] CHEN, Ling, et al. Group-aware graph neural network for nationwide city air quality forecasting. ACM Transactions on Knowledge Discovery from Data, 2023, 18. Jg., Nr. 3, S. 1-20.

[16] Sun, Q., Li, J., Peng, H., Wu, J., Ning, Y., Yu, P.S. and He, L., 2021, April. Sugar: Subgraph neural network with reinforcement pooling and self-supervised mutual information mechanism. In *Proceedings of the Web Conference 2021* (pp. 2081-2091).

[17] DUMITRESCU, Bogdan; BĂLTOIU, Andra; BUDULAN, Ştefania. Anomaly detection in graphs of bank transactions for anti money laundering applications. IEEE Access, 2022, 10. Jg., S. 47699-47714.

یازدهمین همایش سالانه
بانکداری نوین و
نظام‌های پرداخت
پژوهشکده پولی و بانکی
بانک مرکزی جمهوری اسلامی ایران
بانک مرکزی جمهوری اسلامی ایران
شرکت ملی انفورماتیک
زیست بوم بانکداری هوشمند؛ حکمرانی دیجیتال و نظارت فنّاورانه
تهران، مرکز همایش‌های بین‌المللی برج میلاد ـ ۵ و ۶ آذر ۱۴۰۳

Detection of Suspicious Money Laundering Transactions Using Graph Representation Learning

Zahra Nourollah **Mohsen Rezvni**[1]

Abstract

Money laundering, one of the most significant financial crimes, has posed a persistent challenge to financial systems in recent decades. The complexity and dynamic nature of such activities demand advanced and innovative detection methods. This study proposes a three-stage framework to detect suspicious sub-networks in financial transactions, aiming to uncover money laundering activities effectively. In the first stage, transaction data is preprocessed and modeled as a directed graph of financial transactions. In the second stage, community detection algorithms are applied to identify supernodes representing customer account clusters. Using the membership matrix and relationships among customer accounts within these communities, a "transaction group" graph is constructed. Subsequently, structural features, financial flows, and embeddings of subgraphs corresponding to each community are extracted from the transaction group graph. In the third stage, these extracted features are analyzed using deep learning algorithms to detect suspicious sub-networks within the transaction group graph based on learned patterns. The proposed model's effectiveness was evaluated using financial transaction data from the Romanian online bank, Libra, successfully identifying suspicious money laundering activities over a three-month period.

The results demonstrate a significant improvement over existing methods. The use of community detection algorithms combined with rule-based approaches to analyze and construct transaction group graphs for identifying suspicious group behaviors in banking transactions represents a key innovation of this study. These findings contribute to enhancing financial system security and mitigating risks associated with illegal activities.

Keywords: Graph Representation Learning, Anomaly Detection, Money Laundering
JEL Classification: G10, G18, G19

[1] enznoor@gmail.com