کشف تراکنش‌های مشکوک به پول‌شویی بر اساس الگوی بافتاریِ حساب‌های بانکی

علی فرخیان[۱]
عبدالله چاله‌چاله[۲]

تاریخ دریافت:۹۶/۹/۷
تاریخ پذیرش:۹۷/۹/۱۱

چکیده:

مقابله با پول‌شـــویی مهم‌ترین ابزار مبارزه علیه جرم و جنایت است و بزرگ‌ترین چالش مبارزه با پول‌شـــویی در حوزه بانک، تشخیص «تراکنش‌های مشکوک به پول‌شویی» می‌باشد. عدم توجه به بافت صاحبان حســـاب‌های بانکی، باعث کارایی پایین روش‌های مقابله با پول‌شـــویی می‌شود. هدف این تحقیق ارائه یک روش شناسایی تراکنش‌های بانکی مشکوک به پول‌شویی با استفاده از تکنیک‌های آماری فرآیند داده‌کاوی در تحلیل «تراکنش‌های پَرت بافتاری» است که تراکنش‌های پول‌شـــویان را در مرحله یکپارچه‌سازیِ فرآیند پول‌شویی هدف قرار داده است. روش این پژوهش تحلیلِ محتوا بوده است و جامعه آماری تحقیق شامل ۱/۸ میلیون تراکنشِ ۱۰۰۸ نفر، طی مدت ۶ ســـال از ۴۸ بافت مختلف می‌باشد که به روش شبیه‌سازی رایانه‌ای ایجادشده است. توزیع‌های احتمالی استفاده‌شده در شبیه‌ساز مزبور بر اساس آزمون کولموگروف-اسمیرنوف از تراکنش‌های مقطعی ۵۰ نفر استخراج‌شده است، تراکنش‌های مذکور به روش میدانی جمع‌آوری‌شده‌اند. هرچند بـــه دلیل فراهم نبودن تعداد کافـــیِ تراکنش‌های بانکیِ واقعی از تراکنش‌های شبیه‌سازی‌شـــده رایانه‌ای اســـتفاده گردیده است، لیکن شبیه‌سازی توانایی ایجاد سناریوهایی را دارد که در دنیای واقعی، فراهم‌ســـازی آن‌ها ممکن نیست. آزمون ایده تحقیق، بیانگر نرخ ۱۰۰٪ پیش‌بینی درستِ تراکنش‌های مشکوک و نرخ ۱/۱۴٪ پیش‌بینی غلطِ تراکنش‌های عادی به‌عنوان مشکوک می‌باشد که نسبت به اغلب روش‌ها، پیشرفتی محسوس را نشان می‌دهد. نتایج تحقیق نشان داد که توجه به بافت صاحبان حساب بانکی در تشخیص پول‌شویی موجب ارتقاء دقت روش‌های مزبور می‌شود.

کلیدواژه‌ها: پول‌شویی، بافت، متغیر بافتاری، متغیر رفتاری، پنجره مجموعه کاری.

۱- کارشناس ارشد مهندسی نرم‌افزار، دانشگاه رازی، طاق‌بستان، کرمانشاه، ایران
۲ - استادیار و عضو هیئت‌علمی گروه کامپیوتر دانشکده فنی مهندسی، دانشگاه رازی، طاق‌بستان، کرمانشاه، ایران.
نویسنده مسئول، پست الکترونیک: Chalechale@razi.ac.ir

۱. مقدمه

پول‌شـــویی[۳]، فرآیند پنهان‌ســـازی و قانونی جلوه دادن منشـــأ پول کثیف[۴] و یا به عبارتی قانونی جلوه دادن عواید مالی حاصل از جرم و جنایت می‌باشـــد و بعد از بانکداری و صنعت خودرو، سومین کسب‌وکار[۵] بزرگ دنیاســـت (لوپزروجاس و اَکسلسون، ۲۰۱۲)، (رابینسون، ۱۹۹۶) و در چهار حوزه بانکـــداری، بیمـــه، مالیات و گمرک، رخ دادن آن متصور اســـت. به گزارش «دفتـــر مقابله با جرم و مواد مخدر ســـازمان ملـــل»[۶] (۲۰۱۷) دو تا پنج درصد درآمد ناخالص داخلـــی[۷] دنیا، معادلِ ۸۰۰ تا ۲۰۰۰ میلیارد دلار در چرخه فعالیت‌های پول‌شـــویی قرار دارد. قبل از حادثه ۱۱ ســـپتامبر اطلاق واژه پول‌شـــویی بیشـــتر متوجه تطهیر درآمدهای ناشـــی از تجارت مواد مخدر بود اما بعد از حادثه مزبور حوزه‌های جدیدی شـــامل اختلاس، رشوه، دزدی، تقلب در بانک‌ها، بیمه‌ها و بازارهای سرمایه، دایر کردن قمارخانه و خانه‌های فســـاد، قاچاق انســـان و کالا، تجارت اســـلحه، شرکت در شبکه‌های کلاه‌برداری و تأمین مالی تروریسم را نیز دربر گرفته است (ساکی، ۱۳۹۳).

بنا به گزارش کمیته ملی حملات تروریســـتی آمریکا در خصوص حادثه تروریستی ۱۱ سپتامبر، از خارجِ مرزهای آمریکا و به‌صورت الکترونیکی به حساب‌های تروریست‌ها پول واریزشده و توسط ایشان در داخل، برداشـــت گردیده است، ازآنجاکه تا قبل از این حوادث کنترل‌های حوزه پول‌شویی بیشتر متوجه عواید حاصل از مواد مخدر بوده اســـت، هشـــداری از طرف بانک‌ها در خصوص تراکنش‌های[۸] تروریســـت‌ها صادر نشـــد، لذا این ضعف در کنترل تراکنش‌های بانکی موجب شد که ایالات‌متحده آمریکا و اتحادیه اروپا در قوانین مربوط به پول‌شویی و نقل‌وانتقال پول تجدیدنظر کرده و اصلاحاتی انجام دهند و به‌تبع آن‌ها ســـایر کشـــورها نیز قوانین خود را در این حوزه به‌روزرسانی کردند. قانون مبارزه با پول‌شـــویی ایران، در تاریخ ۲ بهمن‌ماه سال ۱۳۸۶ به تصویب مجلس شورای اسلامی رسید و در ۱۷ بهمن همان سال مورد تائید شورای نگهبان قرار گرفت.

کارشناسان ســـازمان ملل از ایران به‌عنوان کشوری مســـتعد برای شست و شوی پول‌های کثیف نام‌برده‌اند (تذهیبـــی، ۱۳۹۴) و همچنین در گزارش[۹] «اداره بین‌المللی مبارزه با مواد مخدر و تنفیذ قانون»[۱۰] (۲۰۱۶) وزارت امور خارجه آمریکا -که هرساله گزارشی در خصوص وضعیت سایر کشورها از حیث مبارزه با پول‌شویی و جرائم مالی منتشر می‌کند- از ایران به‌عنوان کشوری نام‌برده شده است که در مورد آن بالاترین ســـطح نگرانی (نگرانی عمده[۱۱]) وجود دارد، در گزارش مزبور عبارت «حجم

---

3 . Money Laundering
4 . Dirty Money
5 . Business
6 . United Nations Office on Drugs and Crime
7 . GDP
8 . Transaction
9 .2016 INCSR
10 . INL
11 . Primary Concern

عظیم پول‌شـــویی در بازار ســـرمایه ایران رخ می‌دهد» آمده است و این در حالی است که با تصویب قانون مبارزه با پول‌شویی در ایران، مبانی قانونی مبارزه با پول‌شویی فراهم می‌باشد.

با توجه به حجم عظیم پول گردش یافته در فرآیند پول‌شویی، مخلوط شدن آن با کسب‌وکارهای قانونی، استراتژی‌ها و تصمیمات اقتصادی کشورها را بی‌اعتبار و یا به نفع پول‌شویان مصادره می‌کند و عـــلاوه بر آن، انگیزه کار و تلاش اقشـــار جامعه را از بین می‌بـــرد، لذا اختلال در جامعه را به دنبال دارد و درنتیجه عموم مردم را تحت تأثیر قرار می‌دهد (تذهیبی، ۱۳۹۴). عدم مبارزه موفق با فرآیند پول‌شویی در هر کشوری، گزارش‌های منفی مؤسسات بین‌المللی در این خصوص را به همراه دارد و موجب می‌شـــود، اعتماد سرمایه‌گذاران داخلی و خارجی جلب نشده و جذب سرمایه‌گذاری‌های بلند اتفاق نیفتد. علاوه بر آن پول‌های کثیف ممکن است به‌عنوان یک ابزار قدرتمند در راه‌یابی افراد فاسد به مناصب سیاســـی به‌کارگیری شده و ســـلامت فضای سیاسی کشورها را مختل کند. از طرفی عدم تحویل مرتکبان پول‌شـــویی به محاکم قضایی به‌نوعی مشوق رفتارهای خلاف قانون ایشان می‌باشد و بر هم خوردن آرامش فضای جامعه را در پی دارد. در اهمیت مبارزه علیه پول‌شویی لوئیس ج. فری[۱۲] رئیس اســـبق سازمان اف‌بی‌آی[۱۳] گفته اســـت: «مؤثرترین ابزار مبارزه با جنایت سازمان‌یافته، مبارزه علیه پول‌شویی است».

با توجه به موارد ذکرشـــده، مســـئله مطروحه، روش و چگونگی تشـــخیص تراکنش‌های مرتبط با فرآیند پول‌شـــویی اســـت. رفتارهای مالی افراد از پارامترهایی من‌جمله؛ درآمـــد ماهانه، تعداد افراد خانواده، وضعیت مالکیت مســـکن و یا مســـتأجر بودن، داشتن وسیله نقلیه شـــخصی و غیره متأثر می‌باشـــد، این پژوهش با دسته‌بندی افراد بر اساس پارامترهای مذکور و سپس مقایسه تراکنش‌های یک فرد با ســـایر افراد هم دســـته از حیث برخی متغیرهای خاص توسط ســـاختار داده‌های ابداعی و معیارهای آماری، تراکنش‌های غیرمعمول را به‌عنوان تراکنش‌های مشـــکوک به پول‌شـــویی اعلام می‌کند. اســـتفاده‌کنندگان بالقوه نتایج این تحقیق «بانک‌ها»، «مؤسســـات مالی»، «ســـازمان امور مالیاتی» و نهادهای نظارتی ازجمله «دیوان محاســـبات»، «ســـازمان بازرسی» و «وزارت اطلاعات» می‌باشند.

❐ ۲. بیان مسئله

فرآیند پول‌شـــویی شـــامل ســـه مرحله کلی جایگذاری[۱۴]، لایه‌گذاری[۱۵] و یکپارچه‌ســـازی[۱۶] است (رابینســـون، ۱۹۹۶)، (اســـکات، ۲۰۰۶) و در انجام هرکدام از این مراحل طیف وســـیعی از رویه‌ها و فعالیت‌ها وجود دارد و هر روزه نیز رویه جدیدی توســـط پول‌شـــویان ابداع می‌گردد اما درهرحال

12 . Louis Joseph Freeh, 5th Director of FBI
13 . FBI
14 . Placement
15 . Layering
16 . Integration

تراکنش‌هایِ بانکی، جزء اساسی و لاینفک فرآیند پول‌شویی هستند. پول‌شویان برای بهره‌برداری از عواید حاصل از جرائم، پول‌های کثیف را فارغ از منشأ آن‌ها به سیستم بانکی تزریق می‌کنند (جایگذاری) و با مجموعه‌ای از نقل و انتقالات سعی در گم کردن رد آن‌ها نموده (لایه‌گذاری) و نهایتاً مبالغ مذکور را به تعداد محدودی حساب منتقل می‌کنند (یکپارچه‌سازی)، علی‌رغم ظاهر ساده فرآیند پول‌شویی، باید توجه داشت که امکان شناسایی تراکنش‌های پول‌شویان با یک روش ساده و سرراست وجود ندارد زیرا سریع و ساده‌تر شدن انتقال پول از سویی و همه‌گیر شدن استفاده از انواع پول‌های جدید، مخصوصاً پول الکترونیکی[17]، موجب شده روزانه میلیون‌ها تراکنش بانکی انجام شود و جستجو در میان میلیون‌ها تراکنش انجام‌گرفته در شبکه بانکی و تطبیق آن با مصادیق پول‌شویی توسط نیروی انسانی غیرممکن است -خاصه آنکه متخلفان از مشاوره متخصصان حوزه‌های مختلف ازجمله آی‌تی، بانکداری، حسابداری، گمرک و حقوق به‌منظور پول‌شویی بهره می‌برند- بنابراین بزرگ‌ترین چالش مبارزه با پول‌شویی «کشف تراکنش‌های بانکی مشکوک به پول‌شویی» است.

روش‌های شناسایی پول‌شویی در حوزه بانک، به‌طورکلی به چهار دسته ۱- کنترل بر اساس قواعد[18]، ۲- شناسایی مشتریان و فعالیت‌های پرمخاطره[19] توسط کارکنان نهادهای مالی، ۳- تحلیلِ پیوند[20] و ۴- شناسایی داده‌های پَرت[21] تقسیم می‌شود (اداره برآورد فن‌آوری کنگره آمریکا، ۱۹۹۵). به‌منظور «تحلیل پیوند» و «شناسایی داده‌های پَرت» می‌توان از داده‌کاوی[22] استفاده کرد.

داده‌کاوی فرآیند به‌کارگیری تکنیک‌های حوزه‌هایی مانند آمار، یادگیری ماشین، ماشین‌های بردار پشتیبان[23]، نظریه گراف و مصورسازی برای تحلیل داده‌های خام و استنتاج اطلاعات مفید از آن‌ها می‌باشد و امروزه به‌عنوان یک حوزه عظیم کاربردی در اکثر شاخه‌های علم کاربرد دارد. مهم‌ترین روش‌های داده‌کاوی «دسته‌بندی»، «خوشه‌بندی» و «تحلیل داده‌هایِ پرت» است و ازآن‌جهت که داده‌کاوی قادر به کشف الگوهای نامشخص و استخراج دانش از میان انبوهی از داده‌هاست (هان و همکاران، ۲۰۱۱)، قابلیت به‌کارگیری برای کشف فرآیند پول‌شویی بر اساس روش‌های «تحلیلِ پیوند» و «شناسایی داده‌های پرت» را دارد. داده‌های پَرت به سه دسته داده‌های پَرت سراسری، بافتاری[24] و گروهی تقسیم و تحلیل می‌شوند (همان، ۲۰۱۱).

به‌منظور بالا بردن دقت روش‌های مقابله با پول‌شویی گروه FATF[25] نقاطی را به‌عنوان نقاط آسیب‌پذیری پول‌شویان اعلام و توصیه کرده است که در هرکدام از روش‌های مقابله با پول‌شویی،

17 . Electronic Money
18 . Rule Based
19 . Risked Based
20 . Link Analysis
21 . Outlier Data Detection
22 . Data Mining
23 . Support Vector Machines
24 . Contextual
25 . Financial Action Task Force

تمرکز اصلی بر این نقاط باشـــد، نقاط مشخص‌شـــده عبارت‌اند از ۱- ورود پول نقد به سیستم مالی، ۲- انتقال در سیســـتم مالی (تراکنش‌های بین حساب‌ها و یا بین بانک‌ها)، ۳- جریان عبور پول نقد از مرزها.

بخش اعظـــم تراکنش‌های بانکی از نوع تراکنش‌های مبتنی بر بانکـــداری الکترونیکی و اینترنتی می‌باشد و مراجعه به شـــعب بانکی برای انجام تراکنش‌های بانکی روزبه‌روز کاهش می‌یابد لذا روش «شناسایی مشـــتریان و تراکنش‌های پرمخاطره توسط کارکنان شـــعب» کارایی لازم برای مقابله با پول‌شـــویی را ندارد، علاوه بر آن، تجربه و دانش عامل انســـانی نقش اساســـی در موفقیت این روش دارد، بنابراین با توجه به اشـــتغال افراد مختلف با توانایی متفاوت در سیســـتم بانکی، همیشه کارایی یکسانی را نمی‌توان انتظار داشت. ازآنجاکه پول‌شویان از مشاوره و همکاری نیروهای متخصص برای پنهان‌سازی فعالیت‌های خود بهره می‌برند، لذا در اغلب موارد تراکنش‌هایی که در راستای پول‌شویی انجام می‌گیرند، با روش «کنترل بر اســـاس قواعد» قابل‌کشـــف نخواهند بود. روش «تحلیل پیوند» نیازمند یک نقطه آغاز به‌عنوان ورودی اســـت که این نقطه آغاز باید توســـط سایر روش‌های مقابله با پول‌شویی تولید شود و در مرحله بعد توسط این روش، سایر عناصر دخیل و مرتبط در فرآیند مزبور کشـــف گردند، لذا تنها روش مســـتقل و کامل در بین چهار روش مقابله با پول‌شویی در حوزه بانک، روش «شناسایی داده‌های پَرت» است.

بر اســـاس مطالعه میدانی انجام‌گرفته در حال حاضر در برخی شـــعب بانک‌ها در راستای مبارزه با پول‌شویی در حین افتتاح حساب و یا در زمان انجام تراکنش بانکی در محل شعبه، بر اساس روش‌های «کنترل بر اساس قواعد» و «شناسایی مشتریان و فعالیت‌های پرمخاطره» فرم‌های «گزارش معاملات مشکوک»[۲۶] تنظیم می‌گردد، اما با توجه به گزارش‌های نهادهای بین‌المللی و رده‌بندی ایران در گروه کشورهای با حجم عظیم پول‌شویی، ناکافی بودن اقدامات مذکور روشن است، لذا می‌بایست از سایر روش‌ها نیز برای مبارزه با پول‌شویی بهره برد.

در این تحقیق یک روش مقابله با پول‌شـــویی در حوزه بانک و با تمرکز بر روش «شناســـایی داده پَرت» با استفاده از «تکنیک‌های آماریِ فرآیندِ داده‌کاوی» در «تحلیل داده‌های پَرت بافتاری»[۲۷] ارائه می‌شود به‌طوری‌که تراکنش‌هایِ بانکیِ پول‌شویان (یکی از نقاط آسیب‌پذیری پول‌شویان) در مرحله سوم از فرآیند پول‌شویی (یکپارچه‌سازی) را هدف قرار داده است.

❐ ۳. مبانی نظری

برخی داده‌ها تنها در صورتی شـــک‌برانگیز و تعجب‌آور هســـتند که در مقایسه با داده‌های حوزه یا بافت[۲۸] خود مقایسه شوند، این داده‌ها «داده پَرت بافتاری» نامیده می‌شوند. اشیایی که از حیث برخی صفات و یا خصوصیات خاص، همانند هستند و یا به‌عبارت‌دیگر در یک دسته قرار می‌گیرند، از حیث

26 . Suspicious Transactions Report
27 . Contextual Outlier Data Analysis
28 . Context

آن صفات، بافت یکسان دارند و به عبارتی یک بافت را تشکیل می‌دهند. به‌طور عادی[29] بچه‌های با سن‌های بین ۷ تا ۱۳ سال، دانش‌آموز مقطع ابتدایی و بین ۱۳ تا ۱۶ سال دانش‌آموز مقطع متوسطه اول و بین ۱۶ تا ۱۹ سال دانش‌آموز مقطع متوسطه دوم را تشکیل می‌دهند در این مثال از متغیر سن برای تعیین سه بافت ابتدایی و متوسطه اول و دوم استفاده شد، این نوع متغیرها را «متغیرهای بافتاری»[30] و دسته‌های ایجادشده بر اساس متغیرهای بافتاری، بافت نامیده می‌شوند. افرادی که در هر بافت قرار می‌گیرند لازم است برخی ضوابط مخصوص به آن بافت را رعایت کنند، مثلاً دانش‌آموز مقاطع مختلف تحصیلی می‌بایست یونیفرم بارنگ ویژه‌ای بپوشند و انتظار داریم توانایی حل برخی مسائل را داشته باشند، این ضوابط خاص همان «متغیرهای رفتاری»[31] می‌باشند که انتظار می‌رود در هر بافت خاص رعایت گردد و رفتاری خارج از آن بروز نکند. به‌عنوان‌مثال دیگر، عرض و طول جغرافیایی متغیرهای مورداستفاده جهت تعیین بافت یک منطقه جغرافیایی می‌باشند. دمای ۳۷ درجه سانتی‌گراد برای مناطق جنوبی ایران در زمستان ممکن است عادی باشد ولی همین دما برای یکی از شهرهای شمال غرب ایران در فصل زمستان غیرعادی[32] می‌باشد، یعنی عادی و یا غیرعادی بودن دمای ۳۷ درجه با توجه به بافت منطقه جغرافیایی مشخص می‌شود، در این مثال عرض و طول جغرافیایی متغیرهای بافتاری مورداستفاده جهت تعیین بافت یک منطقه جغرافیایی می‌باشند که بر اساس آن‌ها عادی و یا غیرعادی بودن دمای هر بافت قابل ارزیابی است، دما در این مثال متغیر رفتاری است.

ازآنجاکه تناسبی بین صورت‌حساب‌های بانکی مرتکبین جرم پول‌شویی (قاچاقچیان مواد مخدر، سارقین، گردانندگان باندهای فساد و فحشا، آدم‌رباها و غیره) و درآمدهای قانونی (شغل واقعی و قانونی این افراد) وجود ندارد، لذا وضعیت مالی ایشان در مقایسه با سایر افراد از بافت قانونی ایشان در دسته داده‌های پَرت بافتاری قرار گیرد، لذا روش‌های کشف داده‌های پَرت بافتاری برای تشخیص حساب‌ها و تراکنش‌های مالی ایشان قابل‌استفاده خواهد بود.

در شناسایی داده‌های پَرت بافتاری ابتدا باید بافت هر داده یا شیء مشخص شود، صفات خاصه‌ای[33] که بافت هر شیء را مشخص می‌کنند، «صفات خاصه بافتاری» و یا «متغیرهای بافتاری» گفته می‌شوند. هر داده‌ای که به‌صورت قابل‌ملاحظه‌ای از سایر داده‌های هم‌بافت خود انحراف داشته باشد، داده پَرت بافتاری خواهد بود، صفات خاصه‌ای که از آن‌ها برای سنجش و مقایسه داده‌ها استفاده می‌شود را «صفات خاصه رفتاری» و یا «متغیرهای رفتاری» گویند. تعیین صفات خاصه بافتاری و رفتاری خود بخشی از حل مسئله می‌باشد. برای تعیین بافت هر حساب بانکی «شغل» (و یا میزان درآمد ماهانه) صاحب حساب، «تعداد اعضای خانواده» (تعداد افراد تحت تکفل) صاحب حساب،

29 . Normal
30 . Contextual Variable
31 . Behavior Variable
32 . Abnormal or Outlier
33 . Attributes

«وضعیت مالکیت مسکن» (صاحب‌خانه و یا مستأجر بودن)، «وضعیت تملک اتومبیل» و «تعداد تومبیل‌های تحت تملک» پارامترهای مناسبی هستند و افرادی که از حیث صفات خاصه بافتاری وضعیت یکسانی دارند در یک بافت قرار می‌گیرند. توجه شود که مانده‌حساب روزانه افراد در طول یک ماه، تابعی از پارامترهای گفته‌شده می‌باشد و در غیاب اطلاعات مذکور، بهترین انتخاب «توزیع احتمال مترتب بر مانده‌حساب روزانه» خواهد بود.

برای صفات خاصه رفتاری «مقدار ماکزیمم مانده‌حساب»، «ماکزیمم مبلغ واریزی»، «ماکزیمم مبلغ برداشت»، «تعداد تراکنش در یک برهه زمانی»، «جمع مبالغ واریزی در یک برهه زمانی» و «جمع مبالغ برداشتی در یک برهه زمانی» را می‌توان در نظر گرفت. برهه زمانی را می‌توان «یک ماه» در نظر گرفت.

داده‌های پَرت سه دسته کلیِ «داده پَرت سراسری»، «داده پَرت بافتاری» و «داده پَرت گروهی» را شامل می‌شود، لذا سه نوع تراکنش مشکوک شاملِ «تراکنش‌های مشکوک سراسری»، «تراکنش‌های مشکوک بافتاری» و «تراکنش‌های مشکوک گروهی» وجود دارد. تراکنش‌های مشکوک گروهی خود به دودسته کلی تقسیم می‌شود:

۱- تعدادی از تراکنش‌های یک نفر که در یک برهه خاص انجام‌شده است.

۲- تراکنش‌های مشکوک بین افراد مختلف که نتیجه نقل و انتقالات بین حساب‌های افراد مختلف می‌باشد.

درصورتی‌که فضای کل شماره‌حساب‌ها، یک بافت در نظر گرفته شود، تراکنش مشکوک سراسری در دسته تراکنش‌های بافتاری قرار می‌گیرد، هرچند اگر تراکنشی در دسته تراکنش‌های مشکوک سراسری قرار گیرد، یقیناً در بافت خویش نیز جزو تراکنش‌های مشکوک بافتاری خواهد بود.

اثبات: (برهان خلف) فرض می‌کنیم یک تراکنش مشکوک سراسری، مجموعه غیر تهی و تک عضوی GST باشد، به‌طوری‌که جزو تراکنش‌های مشکوک بافتاری نیست. می‌دانیم بافت‌ها مجموعه‌های جدا از هم هستند، به‌عبارت‌دیگر دوبه‌دو اشتراکشان تهی است و اجتماع تراکنش‌های تمام بافت‌ها، مجموعه تمام تراکنش‌ها را تشکیل می‌دهند، یعنی داریم:

$$\bigcup_{i=1}^{n} Context_i = TransactionsSet \qquad (۱)$$

از طرفی طبق فرض (برهان خلف)، تراکنش مشکوک سراسری جزو تراکنش‌های مشکوک بافتاری نمی‌باشد، بنابراین داریم:

$$TransactionsSet = \left(\bigcup_{i=1}^{n} Context_i\right) \cup GST \qquad (۲)$$

بنابراین،

$$TransactionsSet - GST = \left(\bigcup_{i=1}^{n} Context_i \cup GST\right) - GST \quad (۳)$$

لذا داریم:

$$TransactionsSet - GST = \left(\bigcup_{i=1}^{n} Context_i \cup GST\right) \cap \overline{GST}$$

$$TransactionsSet - GST = \left(\bigcup_{i=1}^{n} Context_i \cap \overline{GST}\right) \cup (GST \cap \overline{GST}) \quad (۴)$$

می‌دانیم $A \cap \bar{A} = \Phi$ بنابراین داریم، $GST \cap \overline{GST} = \Phi$ و از ناسازگار بودن GST و $\bigcup_{i=1}^{n} Context_i \bigcup_{i=1}^{n} Context_i$ تحت مجموعه کل TransactionsSet از (۲) می‌توان نوشت $\overline{GST} = \bigcup_{i=1}^{n} Context_i$، بنابراین داریم:

$$\begin{aligned} TransactionsSet - GST &= \left(\bigcup_{i=1}^{n} Context_i \cap \overline{GST}\right) \cup (GST \cap \overline{GST}) \\ &= \left(\bigcup_{i=1}^{n} Context_i \cap \bigcup_{i=1}^{n} Context_i\right) \cup (\Phi) \\ &= \bigcup_{i=1}^{n} Context_i \end{aligned} \quad (۵)$$

$$TransactionsSet - GST = \bigcup_{i=1}^{n} Context_i$$

از تناقض به وجود آمده بین رابطه (۱) و (۵) نتیجه می‌شود که فرض برهان خلف غلط بوده و تراکنش مشکوک سراسری در حوزه پول‌شویی جزو تراکنش‌های مشکوک بافتاری نیز می‌باشد. بدیهی است دسته اول از تراکنش‌های مشکوک گروهی با توجه به اینکه مربوط به یک شماره‌حساب می‌باشند جزو تراکنش‌های مشکوک بافتاری قرار می‌گیرند و دسته دوم تراکنش‌های گروهی هم ازآنجاکه اجتماع چندین زیرمجموعه می‌باشد و هر زیرمجموعه خود متعلق به یک نفر است و در بافت خود آن فرد موردبررسی قرار می‌گیرد، می‌توان نتیجه گرفت تمام انواع تراکنش‌های پرت در دسته تراکنش‌های پرت بافتاری قابل‌کشف می‌باشند.

❐ ۴. پیشینه پژوهش

آلسکرو و همکاران (۱۹۹۷) سامانه‌ای مبتنی بر شبکه‌های عصبی سه لایه پس‌انتشار جهت شناسایی تخلف و کلاه‌برداری مرتبط با کارت‌های اعتباری معرفی کرده‌اند و اعلام داشته‌اند به نتایج مطلوبی در تشخیص تخلف و کلاه‌برداری رسیده‌اند. در این مقاله ذکرشده است از هر توزیع احتمال (مانند توزیع گاوسی) و یا ترکیب چند توزیع احتمال با پارامترهای مناسب جهت شبیه‌سازی تراکنش‌های بانکی و زمان انجام تراکنش‌ها می‌توان استفاده کرد.

چـــن و همـــکاران (۲۰۰۴) برای هر فرد دارنده کارت اعتباری مدلی را با اســـتفاده از تکنیک‌های مبتنی بر بردار پشـــتیبان ماشـــین بر اســـاس یک پرسشنامه ۱۰۵ تا ۱۲۰ ســـؤالی تشکیل داده‌اند. تراکنش جدید فرد، توسط مدل ساخته‌شده مربوطه ارزیابی می‌شود و درصورتی‌که با پیش‌بینی مدل همخوانی نداشته باشد، به‌عنوان تراکنش مشکوک اعلام می‌گردد. تعداد ۱۰۰ دانشجو در این تحقیق مشارکت داشته و نگارندگان نتیجه گرفته‌اند رویه مورداستفاده در تشخیص تراکنش‌های پَرت موفق بوده است.

ژائو و همکاران (۲۰۰۶) در شناســـایی تراکنش‌های مشـــکوک به پول‌شویی از ۴۰ توصیه سازمان FATF استفاده کرده و یک سیستم «کنترل بر اساس قواعد» مبتنی بر «عاملِ هوشمند» ارائه داده‌اند. درراه حل مزبور رفتارهای غیرعادی و یا پرت، کشف و تراکنش‌های ریسکی گزارش می‌شود.

ونگ و دونگ (۲۰۰۹) یک الگوریتم کشف تراکنش‌های مشکوک به پول‌شویی بر اساس خوشه‌بندی مبتنی بر «درخت پوشـــای مینیمم اصلاح‌شـــده» ارائه کرده‌اند. با تعریف دو متریک شباهت و عدم شـــباهت خاص و اســـتفاده از آن‌ها به‌عنوان معیار فاصله در خوشـــه‌بندی داده‌ها، نگارندگان مدعی شـــده‌اند بعد از تزریق ۶۰ مورد پول‌شـــویی مصنوعی به مجموعه داده‌ها، در بهترین حالت موفق به کشف ۷۲ درصد از موارد پول‌شویی شده‌اند.

لی‌خـــاک و همکاران (۲۰۰۹) یک الگوریتم شناســـایی موارد مشـــکوک به پول‌شـــویی مبتنی بر داده‌کاوی و شـــبکه‌های عصبی ارائـــه کرده‌اند. بنا بر ادعای نگارندگان، الگوریتم ارائه‌شـــده در واحد مبارزه با پول‌شـــویی بانک سرمایه‌گذاری بین‌المللی ایرلند آزمون شده و نتایج به‌دست‌آمده قابل‌قبول بوده اســـت. الگوریتم ارائه‌شده جهت خوشـــه‌بندی تراکنش‌ها از روش K-Means استفاده کرده و خروجی به‌دست‌آمده جهت آموزش یک شبکه عصبی از نوع پس‌انتشار -به‌عنوان آموزش‌دهنده موارد نرمال و مشکوک- به‌کاررفته است.

فووا و همکاران (۲۰۱۰) یک مقاله مروری در مورد روش‌های داده‌کاوی بکار رفته برای کشف انواع کلاه‌برداری مالی ارائه کرده‌اند، نگارندگان نتیجه گرفته‌اند دو کمبود عمده در حوزه تشـــخیص کشف کلاه‌برداری بر اســـاس روش‌های داده‌کاوی وجود دارد. اولین کمبود، نبود داده واقعی عمومی[۳۴] برای انجام آزمایش اســـت و دومین کمبود، عدم انتشار متدهای و تکنیک‌های موفقِ کاربردی شده در این حوزه می‌باشد.

لوپزروجاس و اَکسلســـون (۲۰۱۲) در راهکار تشخیص پول‌شـــویی خود از الگوریتم‌های یادگیری ماشـــین شـــامل «درختِ تصمیم» و «قواعد تصمیم» اســـتفاده کرده‌اند و دلیل آن را فهم راحت‌تر خروجی این دسته الگوریتم‌ها برای انسان گفته‌اند. در آزمایشِ ایده این الگوریتم از داده‌های مصنوعی استفاده‌شده است. استفاده از داده‌های مصنوعی همیشه همراه این ریسک است که این داده‌ها بیان درســـتی از دنیای واقعی را ارائه نمی‌دهند و ممکن اســـت نتایج به‌دست‌آمده، منحرف‌کننده باشد اما از طرفی می‌توان توســـط آن‌ها ســـناریوهایی را آماده کرد که ممکن است داده واقعی آن‌ها هیچ‌گاه فراهم نشود.

34 . Public

کوثری‌لنگری و همکاران (۱۳۹۲) با به‌کارگیری درخت تصمیم سعی در کشف و دسته‌بندی الگوی رفتاری کاربران سامانه‌های بانکداری اینترنتی نموده‌اند و الگوهای به‌دست‌آمده را برای کشف تقلب و رفتارهای مشکوک در آن حوزه استفاده کرده‌اند. در این پژوهش درخت تصمیم با استفاده از ۱۰۰،۰۰۰ رکورد آموزشی و توسط چهار الگوریتم Chaid، ex_Chaid، C۴,۵ و C۵,۰ آموزش‌دیده است و برای ارزیابی درخت تصمیم از ۷۰ رکورد آزمایشی استفاده‌شده است، بهترین نتیجه با درصد موفقیت ۹۱ درصد متعلق به الگوریتم C۵,۰ بوده است. رکوردهای اطلاعاتی استفاده‌شده از یک بانک خصوصی اخذشده است.

محمدی و کاظمی‌فرد (۱۳۹۳) یک راه‌حل مبتنی بر خوشه‌بندی بر اساس الگوریتم K-Means برای مبارزه با پول‌شویی ارائه کرده‌اند. تشخیص موارد مشکوک به پول‌شویی در این مقاله بر اساس قانون مبارزه با پول‌شویی مصوب سال ۱۳۸۶ و دستورالعمل‌های ده‌گانه بانک مرکزی برای مبارزه با پول‌شویی است. داده‌های مورداستفاده در این مقاله به‌صورت تصادفی تولیدشده‌اند و تعدادی داده مشکوک به درون داده‌ها تزریق‌شده است، درصد کشف درست داده‌های مشکوک در این مقاله ۷۴ درصد اعلام‌شده است.

منجوناث (۲۰۱۵) ضمن معرفی بسیار مختصر روش‌های دسته‌بندی و خوشه‌بندی، بیان کرده است که هرکدام از تکنیک‌های زیرمجموعه دسته‌بندی و خوشه‌بندی قابلیت به‌کارگیری در سامانه‌های ضد پول‌شویی[۳۵] را دارند، همچنین نتیجه گرفته است که به دلیل تعداد زیاد تراکنش‌های روزانه در سیستم بانکی، امکان رسیدگی به آن‌ها بدون کمک نرم‌افزارهای کامپیوتری وجود ندارد.

سورِش و همکاران (۲۰۱۶) به ارائه یک راهبرد شناسایی عامل پول‌شویی در مرحله لایه‌گذاری با استفاده از نظریه گراف‌ها پرداخته‌اند. روش ارائه‌شده در این مقاله به‌منظور غلبه بر عدم وجود تکنیک‌های موفق ضد پول‌شویی در کشور هند و نیز زمان‌بر بودن راهبردهای اعلامی «ریزرو بانک»[۳۶] این کشور برای تشخیص تراکنش‌های مشکوک به پول‌شویی ارائه‌شده است و شامل دو مرحله است، ابتدا بر اساس تکنیک درهم سازی، حساب‌هایی که به‌طور مکرر و زیاد در تراکنش‌ها دخیل هستند تشخیص داده می‌شوند، در مرحله بعد حساب‌های پرتکرار و سایر حساب‌های مرتبط با آن‌ها در یک ساختار گرافی به هم متصل و به‌عنوان مشکوک اعلام می‌گردند.

❐ ۵. سؤال پژوهش

با توجه به اهمیت و ضرورت مبارزه با پول‌شویی، سؤالی که این تحقیق در راستای پاسخگویی به آن انجام‌شده، این است که با چه روشی و چگونه می‌توان تراکنش‌های مشکوک به پول‌شویی را از میان میلیون‌ها تراکنش تشخیص داد؟

35 . Anti-Money Laundering (AML)

36 . Reserve Bank of India

❐ ۶. روش‌شناسی پژوهش

ایـــن پژوهـــش ازنظر ماهیت جزو پژوهش‌هـــای توصیفی طبقه‌بندی می‌شـــود و روش آن تحلیلِ محتوا اســـت. جهت اجرا و ارزیابی ایده‌های این تحقیق، توســـط پژوهشگران آن، نرم‌افزار اختصاصی تحلیل تراکنش‌های بانکی، طراحی و برنامه‌نویســـی شده است، تحلیل‌های مزبور بر اساس روش‌های داده‌کاوی مبتنـــی بر آمار و احتمال و تجزیه‌وتحلیل ریاضی انجام می‌گیرد. جهت تولید تراکنش‌های بانکـــی موردنیاز بـــرای آزمایش ایده تحقیق نیز نرم‌افزار شبیه‌ســـاز رایانـــه‌ای اختصاصی، طراحی و برنامه‌نویســـی گردیده اســـت. از تراکنش‌های بانکی که به روش میدانی جمع‌آوری‌شده، توزیع‌های احتمال مورداستفاده در شبیه‌سازِ تراکنش‌هایِ بانکی استخراج گردیده است. شبیه‌سازی انجام‌گرفته ازنظر نوع، شبیه‌ســـازیِ تجربی[۳۷] می‌باشد. پژوهش به لحاظ هدف، پژوهشی توسعه‌ای-کاربردی است و نتایج این تحقیق به‌طور بالقوه برای نهادهای مالی، مالیاتی و نظارتی قابل‌استفاده است.

زبان برنامه‌نویس و نرم‌افزارهای استفاده‌شده در این پژوهش. نرم‌افزار Microsoft Excel Ver.2010 برای کارهای اولیه و نرم‌افزار Professional EasyFit Ver.5.5 برای کارهای آماری و احتمالاتی به‌کارگیری شده است. نرم‌افزارهای اختصاصی این تحقیق با زبان برنامه‌نویســـی C#.Net Microsoft نوشته و در محیط برنامه‌نویسی Microsoft Visual Studio Ultimate Ver.2012 تولیدشـــده‌اند. نرم‌افزارهای اختصاصی نوشته‌شـــده برای این تحقیق «نرم‌افزار یکپارچه‌سازی حســـاب‌های بانکی»، «نرم‌افزار شبیه‌ساز تولید حساب‌های بانکی»، «نرم‌افزار تزریق تراکنش‌های مصنوعی پول‌شویی»، «نرم‌افزار تشخیص تراکنش‌های مشکوک به پول‌شویی» و «نرم‌افزار ارزیابی» می‌باشـــند. برای مدیریت پایگاه داده این تحقیق، زبان اسکریپت نویسی SQL و نرم‌افزار Microsoft SQL Server Management Studio Ver.2005 مورد بهره‌برداری قرارگرفته است.

جامعه و نمونه آماری. تراکنش‌های موردنیاز برای پیاده‌سازی ایده این تحقیق به روش شبیه‌سازی رایانه‌ای فراهم‌شده است. رویه کار بدین‌صورت بوده است که ابتدا صورتحساب‌های بانکی (شامل تمام انواع تراکنش‌ها اعم از واریز، برداشت، انتقال و غیره و از نوع الکترونیکی و غیر الکترونیکی) یک‌ساله تعداد ۵۰ نفر مربوط به بانک‌های قرض‌الحسنه مهر ایران، دی، ملی و رسالت به روش میدانی فراهم گردیده است (اطلاعات ۲ نفر به دلیل تعداد کمِ تراکنشِ سالانه در این مرحله حذف شدند و ۴۸ نفر باقی ماندند). بعد از ادغام و یکپارچه‌سازی تراکنش‌های بانکی هر فرد از بانک‌های مختلف (برخی در بیش از یک بانک حســـاب داشـــته‌اند) و یکپارچه‌سازی نهایی تمام تراکنش‌ها «رویه حاکم بر زمان و تاریخ انجام تراکنش هر فرد» و «توزیع‌های احتمالی پیوســـته مترتب بر مانده‌حساب» و «توزیع‌های احتمالی گسســـته مترتب بر تعداد تراکنش ماهانه» بر اســـاس آزمون کولموگروف-اسمیرنوف[۳۸] به دســـت آمد. سپس با استفاده از توزیع‌های احتمالی به‌دست‌آمده و رویه حاکم بر زمان و تاریخ انجام تراکنش افراد، یک نرم‌افزار اختصاصی برای تولید تعداد تراکنش‌ها و شبیه‌ســـازیِ مبلغ (مانده‌حساب و واریز یا برداشت)، زمان و تاریخ انجام تراکنش بانکی ایجاد گردید[۳۹].

---

37 . Experimental Simulation

38 . Kolmogorov–Smirnov test

۳۹. لازم به ذکر است در شبیه‌ساز مزبور برای بخش‌های شماره‌حساب، شماره سند، کد شعبه و سایر اطلاعات مربوطه به ویژگی تراکنش‌ها، مدل‌هایی ایجاد و برنامه‌نویسی شده است.

جدول ۱. مشخصات پایگاه داده استفاده‌شده در تشخیص تراکنش‌های مشکوک به پول‌شویی

ردیف	تعداد بافت	تعداد در هر بافت (نفر)	تعداد کل	تاریخ آغاز	تاریخ پایان	زمان شبیه‌سازی	تعداد تراکنش تولیدی	حافظه مصرفی (کیلوبایت)
۱	۴۸	۲۰	۹۶۰	۱۳۹۰/۱/۱	۱۳۹۴/۱۲/۳۰	۰۵:۱۶:۲۱	۱،۷۸۷،۰۰۳	۴۰۲،۸۸۴
۲	۴۸	۱	۴۸	۱۳۹۵/۱/۱	۱۳۹۵/۱۲/۳۰	۰۰:۰۳:۳۷	۱۹،۳۰۶	۲۷،۲۴۸

به‌منظور آزمایش ایده این تحقیق تراکنش‌های ۵ ساله تعداد ۲۰ نفر از هر بافت (۹۶۰ نفر) در فاصله زمانی ابتدای سال ۱۳۹۰ تا انتهای سال ۱۳۹۴ شبیه‌ساز شد، تعداد تراکنش‌های ایجادشده در این مرحله ۱،۷۸۷،۰۰۳ تراکنش است، ردیف اول از جدول ۱ و به تعداد ۱ نفر از هر بافت (۴۸ نفر) نیز در فاصله زمانی اول تا انتهای سال ۱۳۹۵ ایجاد شد، تعداد تراکنش‌های ایجادشده این مقطع نیز ۱۹،۳۰۶ تراکنش می‌باشد، ردیف دوم جدول ۱.

روش پاسخ به بخش اول سؤال پژوهش. ازنقطه‌نظر تئوریک برای کشف داده‌های پَرت می‌توان از همه روش‌های داده‌کاوی (خوشه‌بندی، دسته‌بندی و روش‌های مبتنی بر آمار) بهره برد اما با توجه به جوانب مسئله کاربردیِ «کشف تراکنش‌های مشکوک به پول‌شویی» و تحلیل آن، برخی روش‌ها کارایی خود را از دست می‌دهند.

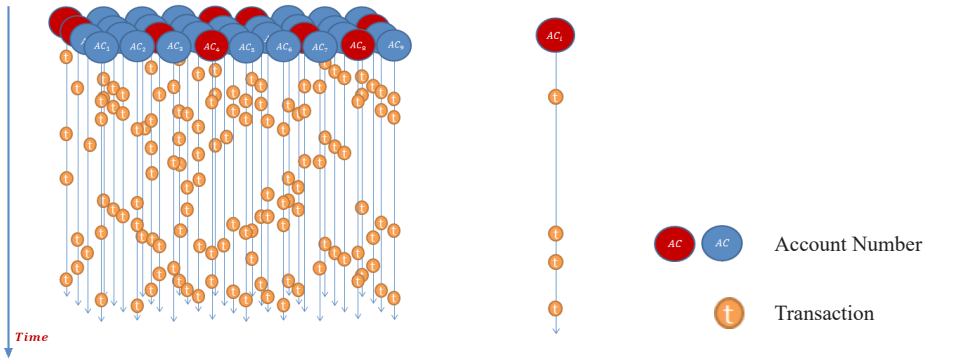
خوشه‌بندی روشی است که داده‌ها را بر اساس یک معیار خاص (فاصله) در خوشه‌هایی تقسیم‌بندی می‌کند و هر داده‌ای که در خوشه‌ها قرار نگیرد را به‌عنوان داده پَرت و یا غیر نرمال اعلام می‌کند. به سه دلیل خوشه‌بندی برای کشف تراکنش‌های مشکوک به پول‌شویی مناسب نیست، اول آنکه در تشخیص تراکنش‌های پَرت گروهی به دلیل ذات تراکنش‌های پَرت گروهی که خود تشکیل یک خوشه را می‌دهند دچار خطا شده و آن‌ها را خوشه نرمال در نظر می‌گیرد. دوم آنکه تراکنش‌های با مقادیر کوچک و خارج از محدوده سایر تراکنش‌ها را نیز مشکوک به پول‌شویی اعلام می‌کند و سوم آنکه وابسته به معیار فاصله است که مقادیر مختلف این معیار، نتایج متغیر و غیر یکسانی را تولید می‌کند. در صورت تعیین مقدار درستی برای معیار فاصله این روش تنها برای کشف تراکنش پَرت سراسری مناسب است.

دسته‌بندی روشی است که در آن بر اساس خواص و رفتار داده‌های موجود یک مدل رفتاری ساخته می‌شود و با استفاده از مدل مزبور، داده‌های جدید در دسته‌هایی تقسیم می‌شوند. به دو دلیل روش دسته‌بندی برای کشف تراکنش‌های مشکوک به پول‌شویی مناسب نیست. اول آنکه به دلیل ماهیت پویای روش‌های پول‌شویی و ناشناخته بودن خیلی از آن‌ها، مدل کردن آن غیرممکن است و در صورت ساخت مدل از روش‌های پول‌شویی کشف‌شده فعلی، همچنان در کشف روش‌های جدید پول‌شویی ناتوان است. دوم آنکه پول‌شویان رفتارهای مالی خود را با بهره‌گیری از کارشناسان خبره، رفتار نرمال جلوه می‌دهند لذا مدل کردن رفتارهای نرمال نیز در کشف پول‌شویی کارایی کاملی ندارد.

یک دسته از روش‌های تشخیص داده‌های پَرت در مقایسه با داده‌های نرمال روش‌های مبتنی بر آمار است. در روش‌های مبتنی بر آمار معیارهایی برای تعیین داده پَرت وجود دارد -روش «خلاصه

پنج عددی» از آن نوع اســـت- لذا روش‌های مبتنی بر آمار برای پاســـخگویی به بخش اول ســـؤال تحقیق بهترین راهکار اســـت و در این پژوهش با استفاده از مفهوم «داده پَرت بافتاری» و با استفاده از «روش‌های تشـــخیص مبتنی بر آمار» به کشف تراکنش‌ها و حســـاب‌های مشکوک به پول‌شویی پرداخته‌شده است.

شکل ۱. نمایش شماتیک حساب‌ها و تراکنش‌های آن‌ها

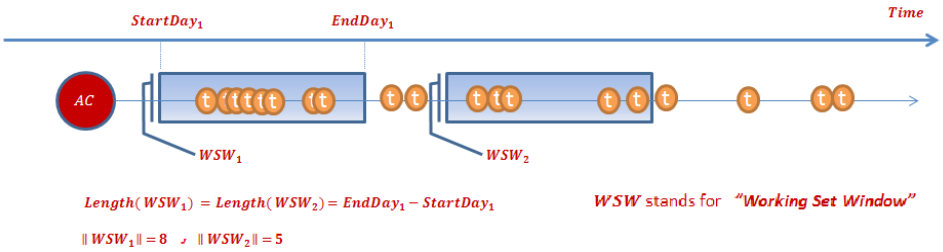


روش پاســـخ به بخش دوم سؤال پژوهش- الگوریتم پیشـــنهادی برای تشخیص تراکنش‌های مشـــکوک به پول‌شویی. قبل از ارائه الگوریتم پیشنهادیِ تشخیص و کشف تراکنش‌های مشکوک به پول‌شویی، چند داده‌ساختار[40] که در این زمینه ایجادشده است را معرفی می‌کنیم. اولین داده‌ساختار، مربوط به شماره‌حساب و تراکنش‌های انجام‌گرفته آن شماره‌حساب می‌باشد، شکل ۱.

شماره‌حساب AC_i و تراکنش‌های آن در بخش ب شکل ۱ بروی محور زمان نمایش داده‌شده است و در بخش ج شـــکل ۱ تعدادی حســـاب (AC_i) و برخی تراکنش‌های آن‌ها در مقطعی از زمان نشان داده‌شـــده‌اند، در این قسمت دو بافت که بر اســـاس صفات بافتاری تمایز پیداکرده‌اند قابل‌تشخیص اســـت. در راســـتای کشـــف فعالیت‌های مشـــکوکی که در برهه‌ای از زمان رخ می‌دهند و گروهی از تراکنش‌ها در آن درگیر هســـتند (داده پَرت بافتاری گروهی)، مفهوم داده‌ســـاختار «پنجره مجموعه کاری»[41] را ایجاد کرده‌ایم، شکل ۲.

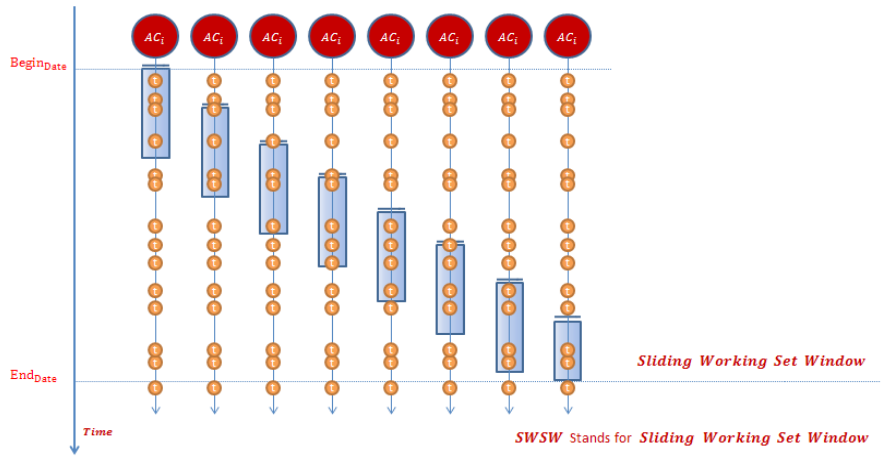
40 . Data Structure
41 . Working Set Window

شکل ۲. پنجره مجموعه کاری



مجموعه‌ای از تراکنش‌های رخداده از $Start_{Date}$ تا $Start_{Date} + WorkingSetWindowLength$ در «پنجره مجموعه کاری» قرار می‌گیرند، در شکل ۲ دو WSW به نمایش درآمده است، در WSW_1 تعداد ۸ تراکنش و در WSW_2 تعداد ۵ تراکنش قرار دارد، تعداد تراکنش‌های قرارگرفته در WSW، «اندازه پنجره مجموعه کاری» تعریف می‌شود و فاصل زمانیِ ابتدا و انتهای WSW را «طول پنجره مجموعه کاری» می‌نامیم.

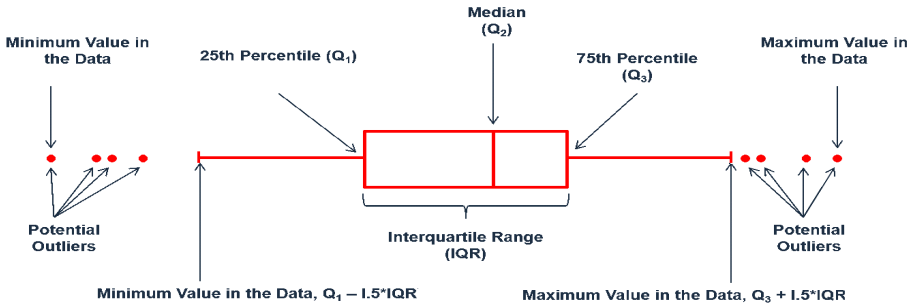
شکل ۳. پنجره مجموعه کاری لغزان



فرض کنید می‌خواهیم یک حساب خاص را از ابتدا تا انتهای سال برای کشف پول‌شویی موردبررسی قرار دهیم و طول پنجره مجموعه کاری را ۳۰ روز یعنی $\text{Length (WSW)} = 30$ در نظر گرفته‌ایم،

بدیهی است که برای بررسی کل یک سال می‌بایست هر بار پنجره مجموعه کاری را یک روز به جلو ببریم و متغیرهای رفتاری را مقایسه کنیم، در این حالت داده‌ساختار «پنجره مجموعه کاری» را که توانایی لغزیدن و جلو رفتن دارد، «پنجره مجموعه کاری لغزان»[۴۲] می‌نامیم شکل ۳. بدین ترتیب تمام فاصله بین $Begin_{Date}$ تا End_{Date} جاروب شده و متغیرهای رفتاری خواسته‌شده موردبررسی قرار می‌گیرند. مقدار پیش‌فرض طول مجموعه کاری یک ماه است، این مقدار توسط کارشناس خبره بانک قابل‌تغییر است.

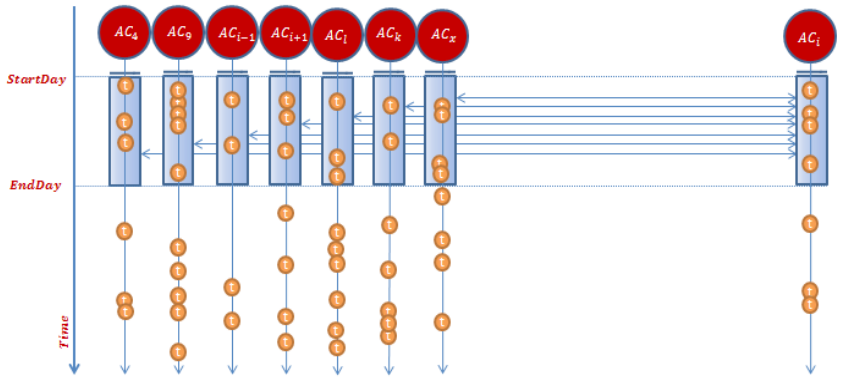
شکل ۴. نمودار جعبه‌ای (کلمن، ۲۰۱۶)



یکی از معیارهای تشخیص داده پرت در آمار، روش «خلاصه پنج عددی» می‌باشد (چمبرز، ۱۹۸۳) که نمایش شماتیک آن در شکل ۴ آورده شده است و بر اساس تعریف آزمون توکی[۴۳] (۱۹۷۷) هر داده‌ای که در بیرون فاصله $[Q_1 - d * IQR, Q_3 + d * IQR]$ برای d = ۱,۵ قرار گیرد، داده پرت ملایم[۴۴] و درصورتی‌که برای d = ۳ در بیرون فاصله مزبور قرار گیرد، داده پرت شدید[۴۵] خواهد بود، از این معیار برای مقایسه متغیرهای رفتاری بهره می‌بریم. در شکل ۵ پنجره‌های مجموعه کاری با یکدیگر مقایسه شده‌اند.

42 . Sliding Working Set Window
43 . Tukey's Test
44 . Mild Outlier
45 . Extreme Outlier

شکل ۵. مقایسه تراکنش‌های داخل پنجرهِ مجموعهِ کاری با سایر حساب‌های هم‌بافت



در فرآیند تشخیص تراکنش‌ها و حساب‌های مشکوک به پول‌شویی، ابتدا با پیمایش شماره‌حساب‌ها (AN_i) و اطلاعات صاحبان آن حساب‌ها، آن‌ها را بر اساس متغیرهای بافتاری دسته‌بندی کرده و برچسب اختصاصی C_i مربوط به بافت آن‌ها را معلوم می‌کنیم سپس در محدوده زمانی ($Begin_{Date}$, End_{Date}) برای هر شماره‌حساب AN_i از بافت C_iC_i با تابع توزیع احتمال fc_i، ۵ مرحله کلی مستقل ذیل با فرض $d = 3$ انجام می‌شود (یک یا بیشتر):

۱- مقدار ماکزیمم مانده‌حساب، واریزی و برداشت AN_i با مقدار ماکزیمم، مانده‌حساب، واریزی و برداشت تمام تراکنش‌ها $AN_j : j \neq i$ مقایسه می‌شود، درصورتی‌که مقدار ماکزیمم هرکدام از مقادیر مانده‌حساب، واریزی و برداشت AN_iAN_i، ماکزیمم تمام تراکنش‌ها باشند، پیغام مناسب برای مشکوک بودن آن تراکنش ذخیره می‌شود (تراکنش مشکوک سراسری است) شکل ۶.

۲- برای تمام تراکنش‌های T_{i_t} از AN_i، ۳ گام ذیل انجام می‌شود؛

۱-۲- برای مانده‌حساب، واریزی و برداشت تمام تراکنش‌های T_{k_t} از $AN_k : k \neq i$، بافت C_i، مقادیر $Q_{1_{c_i}}$ (چارک اول تراکنش‌های بافت C_i) و $Q_{3_{c_i}}$ (چارک سوم تراکنش‌های بافت C_iC_i) محاسبه می‌شود و از آنجا $IQR_{c_i} = Q_{3_{c_i}} - Q_{1_{c_i}}$ (دامنه بین چارکی) محاسبه می‌شود، درصورتی‌که هرکدام از مقادیر مانده‌حساب، واریزی و برداشت تراکنش T_{i_t}، از $Q_{3_{c_i}} + d * IQR_{c_i}$ بزرگ‌تر باشد، پیغام مناسب برای مشکوک بودن آن تراکنش در بافت ذخیره می‌شود.

۲-۲- مقدار ماکزیمم مانده‌حساب، واریزی و برداشت AN_i، با مقدار ماکزیمم، مانده‌حساب، واریزی و برداشت تمام مجموعه تراکنش‌های $AN_k : k \neq i$ بافت C_i مقایسه می‌شود، درصورتی‌که مقدار ماکزیمم هرکدام از مقادیر مانده‌حساب، واریزی و برداشت AN_i، ماکزیمم تمام تراکنش‌های بافت C_i باشند، پیغام مناسب برای مشکوک بودن آن تراکنش در بافت ذخیره می‌شود، شکل ۷.

۳-۲- احتمال رخ دادن مقدار مانده‌حساب تراکنش T_{i_t}، از f_{c_i} محاسبه می‌شود درصورتی‌که از مقدار Contextual_Outlier_Probabilty (از ورودی توسط فرد خبره وارد می‌شود) کمتر باشد، پیغام مناسب برای مشکوک بودن آن تراکنش در بافت ذخیره می‌شود.

۳- قرار می‌دهیم، $\text{Start}_{\text{Date}} = \text{Begin}_{\text{Date}}$ و در محدوده پنجره مجموعه کاری $\text{WSW} = (\text{Start}_{\text{Date}}, \text{Start}_{\text{Date}} + \text{WorkingSetWindowLength})$ برای تمام تراکنش‌های T_{i_t} از AN_i ۵ گام ذیل انجام می‌شود؛

۳-۱- درصورتی‌که محدوده $(\text{Begin}_{\text{Date}}, \text{End}_{\text{Date}})$ پوشش داده نشده است ادامه بده.

۳-۲- $\sum(T_{i_t}(\text{مبلغ واریزی}))$ (جمع مبالغ واریزی) و $\sum(T_{i_t}(\text{مبلغ برداشت}))$ (جمع مبالغ برداشت) محاسبه می‌شود.

۳-۴- برای تمام تراکنش‌های T_{k_t} از $AN_k : k \neq i$، بافت C_i، $\sum(T_{k_t}(\text{مبلغ واریزی}))$ و محاسبه می‌شود. مقادیر $Q_{1\_sum_{c_i}}$ (چارک اول حاصل جمع واریزی و یا برداشت بافت C_i) و $Q_{3\_sum_{c_i}}$ (چارک سوم حاصل جمع واریزی و یا برداشت بافت C_i) محاسبه می‌شود و ازآنجا $\text{IQR}_{\text{sum}_{C_i}} = Q_{3\_sum_{c_i}} - Q_{1\_sum_{c_i}}$ (دامنه بین چارکی حاصل جمع واریزی و یا برداشت بافت C_i) محاسبه می‌شود.

۳-۴- درصورتی‌که هرکدام از مقادیر $\sum(T_{i_t}(\text{مبلغ واریزی}))$ و یا $\sum(T_{i_t}(\text{مبلغ برداشت}))$ از $Q_{3\_sum_{c_i}} + d * \text{IQR}_{\text{sum}_{C_i}}$ بزرگ‌تر باشد، پیغام مناسب برای مشکوک بودن آن تراکنش در بافت ذخیره می‌شود.

۳-۵- $\text{WSW} = (\text{Start}_{\text{Date}}, \text{Start}_{\text{Date}} + \text{WorkingSetWindowLength})$، $\text{Start}_{\text{Date}} = \text{Start}_{\text{Date}} + 1$ (پنجره مجموعه کاری یک روز به جلو لغزانده می‌شود) و به گام ۳،۱ برو.

۴- درصورتی‌که کاربر تمایل به استفاده از پنجره مجموعه کاری با اندازه متغیر داشته است، ۴ گام ذیل انجام می‌گیرد:

۴-۱- $\text{WorkingSetWindowLength} = \text{WorkingSetWindowLength}_{\text{Min}}$

۴-۲- درصورتی‌که $\text{WorkingSetWindowLength} \leq \text{WorkingSetWindowLength}_{\text{Max}}$ الگوریتم مرحله ۳ را انجام بده.

۴-۳- $\text{WorkingSetWindowLength} = \text{WorkingSetWindowLength} + 1$

۴-۴- به گام ۴،۲ برو.

۵- قرار می‌دهیم، $\text{Start}_{\text{Date}} = \text{Begin}_{\text{Date}}$ و در محدوده پنجره مجموعه کاری $\text{WSW} = (\text{Start}_{\text{Date}}, \text{Start}_{\text{Date}} + \text{WorkingSetWindowLength})$ برای تمام تراکنش‌های T_{i_t} از AN_i ۵ گام ذیل انجام می‌شود:

۵-۱- درصورتی‌که محدوده $(\text{Begin}_{\text{Date}}, \text{End}_{\text{Date}})$ پوشش داده نشده است ادامه بده.

۵-۲- $|(T_{i_t})|$ (تعداد تراکنش‌ها) در محدوده WSW محاسبه می‌شود.

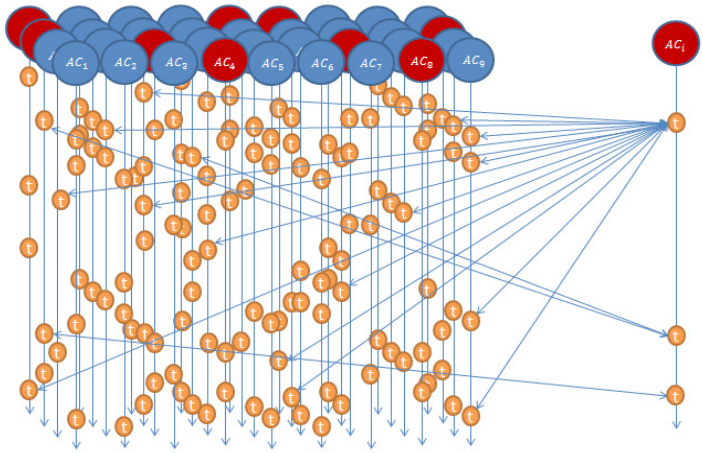
۵-۳- برای تمام تراکنش‌های T_{k_t} از $AN_k : k \neq i$، بافت C_i، $||T_{k_t}||$ (تعداد تراکنش‌ها) محاسبه می‌شود. مقادیر $Q_{1\_count_{c_i}}$ (چارک اول تعداد تراکنش‌های بافت C_i) و $Q_{3\_count_{c_i}}$ (چارک سوم تعداد تراکنش‌های بافت C_i) محاسبه می‌شود و ازآنجا $\text{IQR}_{\text{count}_{C_i}} = Q_{3\_count_{c_i}} - Q_{1\_count_{c_i}}$ (دامنه

بین چارکی تعداد تراکنش‌های بافت C_i) محاسبه می‌شود.

۴-۵- درصورتی‌که مقدار $||(T_{i_t})||$ خارج فاصله $[Q_{1_{count_{C_i}}} - d * IQR_{count_{C_i}}, Q_{3_{count_{C_i}}} + d * IQR_{count_{C_i}}]$ باشد، پیغام مناسب برای مشکوک بودن تعداد تراکنش‌های AN_i نسبت به بافت ذخیره می‌شود.

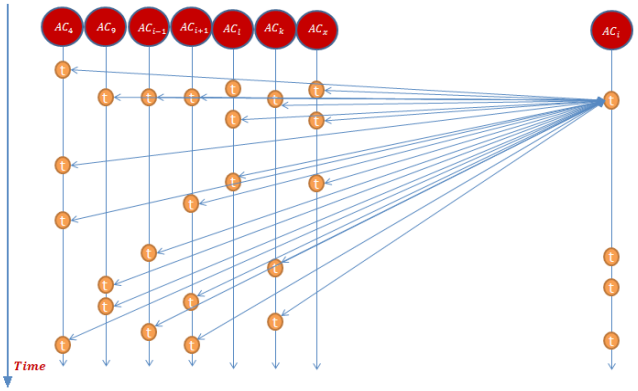
۵-۵- $Start_{Date} = Start_{Date} + 1$ و $WSW = (Start_{Date}, Start_{Date} + WorkingSetWindowLength)$ (پنجره مجموعه کاری یک روز به جلو لغزانده می‌شود) به گام ۵،۱ برو.

شکل ۶. مقایسه یک تراکنش با همه تراکنش‌ها



روش آزمایش الگوریتم پیشنهادی. رویه انجام آزمایش الگوریتم پیشنهادی به این صورت خواهد بود که توسط «نرم‌افزار تزریق تراکنش‌های مصنوعی پول‌شویی» به روش تصادفی ساده شماره‌حساب و یا شماره‌حساب‌هایی انتخاب می‌شود، سپس تراکنش یا تراکنش‌های نشان‌دار که جنبه‌های پول‌شویی دارند را به آن حساب تزریق می‌کنیم و در مرحله بعد توسط «نرم‌افزار تشخیص تراکنش‌های مشکوک به پول‌شویی» پایگاه داده تراکنش‌ها مورد کاوش قرار می‌گیرد و خروجی مناسب تولید می‌کند و علاوه بر آن تراکنش‌های مشکوک به پول‌شویی کشف‌شده، نشان‌دار می‌شوند. در مرحله بعد توسط «نرم‌افزار ارزیابی» بین تراکنش‌های پول‌شویی تزریق‌شده و کشف‌شده (هر دو گروه نشان‌دار بودند) مقایسه انجام می‌گیرد و درصد تراکنش‌های کشف‌شده مشخص می‌شود. ازآنجاکه در همه آزمایش‌ها می‌خواهیم همچنان تراکنش‌های مورد آزمایش ثابت بمانند، دقیقاً به یک‌رویه دو ردیف داده‌های جدول ۱ تهیه‌شده‌اند، تراکنش‌های مصنوعی را به‌ردیف دوم اضافه می‌کنیم و بعد از هر مرحله آزمایش داده‌های ردیف دوم نوسازی می‌شوند و به حالت قبل از آزمایش برگردانده می‌شوند، تراکنش‌های ردیف اول جدول ۱ به‌عنوان مرجع مقایسه بافت‌ها، مورداستفاده قرار می‌گیرند و محتوای آن در آزمایش‌ها تغییر نمی‌کند.

شکل ۷. مقایسه تراکنش یک حساب با تراکنش‌های سایر حساب‌های هم‌بافت



❐ ۷. تحلیل داده‌ها و یافته‌ها

آزمایش اول. در این آزمایش به پنج حساب، هرکدام یک تراکنش مشکوک به پول‌شویی، طبق جدول ۲ اضافه‌شده است، مبالغ ردیف‌های ۱، ۳ و ۵ از مبالغ تمام تراکنش‌های فضای آزمایش بزرگ‌تر است. نتیجه کاوش در جدول ۳ آمده است.

جدول ۲. آزمایش اول، تراکنش‌های تکی مشکوک در چند حساب

ردیف	شماره‌حساب‌انتخابی	مبلغ ماندەحساب	مبلغ واریزی	مبلغ برداشتی	تاریخ تراکنش	زمان تراکنش	بافت
۱	۰۱۲۲۲۹۲۵۳۵۰۰۸۷	۸۰۲۷۴۴۰۸۶	۶۷۱۰۰۳۶۶۷	۰	۱۳۹۵/۷/۵	۱۵:۲۰:۲۲	۳۹
۲	۰۱۹۶۶۳۷۳۹۸۰۰۱۲	۸۱۴۱۴۴	۰	۱۳۷۰۰۰۰۰۰	۱۳۹۵/۹/۱۹	۱۰:۰۵:۰۷	۴۴
۳	۰۲۷۳۳۱۱۰۳۹۰۰۵۰	۷۹۸۰۰۰۴۴۴	۷۹۷۹۶۰۰۶۳	۰	۱۳۹۵/۲/۲	۰۵:۳۵:۴۲	۱۸
۴	۰۳۷۲۹۴۵۸۹۳۰۰۱۳	۵۹۵۲۰۴۳۲۳	۵۹۰۲۸۹۰۰۳	۰	۱۳۹۵/۱۱/۱۳	۹:۴۱:۲۰	۲۳
۵	۰۱۸۰۴۶۸۶۳۷۰۰۷۹	۸۹۰۹۸۰۰۰۰	۸۴۰۵۳۶۰۵۷	۰	۱۳۹۵/۳/۲۹	۰۳:۵۲:۱۵	۱۹

درنتیجه تزریق تراکنش‌های مشکوک، برای جلوگیری از به هم خوردن رابطه بین مقادیر واریزی، برداشتی و مانده‌حساب تراکنش‌ها، ممکن است مبالغ اولین تراکنش ماقبل و برخی تراکنش‌های بعدی تغییر کنند، لذا ممکن است برخی از این تراکنش‌ها در دسته تراکنش‌های مشکوک قرار گیرند. در این آزمایش دو هدف دنبال می‌شود، اول بررسی این موضوع است که آیا الگوریتم پیشنهادی توانایی کشف چندین تراکنش غیر نرمال را در بین تعداد زیاد تراکنش‌های نرمال دارد و در ثانی آیا الگوریتم پیشنهادی دو تراکنش نرمال تزریق‌شده را نیز به‌درستی به‌عنوان تراکنش نرمال تشخیص

می‌دهد.

جدول ۳. نتیجه آزمایش اول، تراکنش‌های تکی مشکوک در چند حساب

ردیف	نتیجه جستجو برای کشف تراکنش مشکوک در بین تمام تراکنش‌ها	زمان مصرفی
۱	موفق	۰۰:۰۰:۱۰
۲	موفق*	
۳	موفق	
۴	موفق*	
۵	موفق	
خروجی برنامه	0122925350087: WITHDRAW Alert (Greater Than Any Other Withdraw in All Accounts) ----> Global Type BALANCE Alert (Greater Than Any Other Balance in All Accounts) ----> Global Type 01966373980012: 02733110390050: WITHDRAW Alert (Greater Than Any Other Withdraw in All Accounts) ----> Global Type DEPOSIT Alert (Greater Than Any Other Deposit in All Accounts) ----> Global Type BALANCE Alert (Greater Than Any Other Balance in All Accounts) ----> Global Type 03729458930013: 01804686370079: WITHDRAW Alert (Greater Than Any Other Withdraw in All Accounts) ----> Global Type DEPOSIT Alert (Greater Than Any Other Deposit in All Accounts) ----> Global Type BALANCE Alert (Greater Than Any Other Balance in All Accounts) ----> Global Type	

*تراکنش‌های ردیف ۲ و ۴ جزو تراکنش‌های مشکوک نیستند.

بر اســـاس جدول ۳، آزمایش با موفقیت انجام‌شـــده اســـت و هر دو هدف برآورده شد، با توجه به تغییر مبالغ واریزی و برداشت تراکنش‌ها به‌منظور حفظ تناسب بین این مقادیر، بیش از یک هشدار برای شماره‌حســـاب‌های ردیف‌های ۱، ۳ و ۵ جدول ۲ ثبت‌شـــده اســـت. تراکنش‌های اضافه‌شده به شماره‌حساب‌های ردیف ۲ و ۴ نرمال بودند و به‌درستی مورد مشکوکی برای آن‌ها اعلام نشد.

آزمایش دوم. در این آزمایش به پنج حســـاب، هرکدام یک تراکنش طبق جدول ۴ اضافه می‌شود. نتیجه کاوش در جدول ۵ آمده است.

جدول ۴. آزمایش دوم، تراکنش‌های تکی مشکوک و غیر مشکوک در چند حساب

ردیف	شماره‌حساب انتخابی	مبلغ مانده‌حساب	مبلغ واریزی	مبلغ برداشتی	تاریخ تراکنش	زمان تراکنش	بافت
۱	۰۲۹۰۴۳۸۹۲۲۰۰۵۸	۲۱۰۹۸۵۷۲۲	۰	۱۸۲۸۳۳۵۹۵	۱۳۹۵/۰۸/۵	۱۲:۵۹:۰۱	۱۲
۲	۰۲۴۵۸۴۷۷۲۴۰۰۸۷	۱۵۵۰۰۰۰۰۰	۰	۱۱۶۱۸۴۴۱۴	۱۳۹۵/۰۲/۰۵	۱۴:۳۰:۵۱	۳۸
۳	۰۱۹۶۶۳۷۳۹۸۰۰۱۲	۱۴۸۷۷۷۰۰۰	۰	۱۱۱۳۸۶۷۶	۱۳۹۵/۰۴/۲۰	۰۲:۱۵:۰۸	۴۴
۴	۰۱۳۹۴۲۰۴۱۹۰۰۲۲	۱۲۰۸۰۲۲۹۱	۴۰۰۰۰۰۰۰	۰	۱۳۹۵/۱۱/۲۰	۱۷:۴۱:۱۹	۲۷
۵	۰۳۸۲۳۲۲۲۰۲۰۰۱۹	۶۰۲	۰	۶۴۳۷۰۰۰	۱۳۹۵/۰۱/۱۶	۱۱:۳۹:۴۳	۲۶

جدول ۵. نتیجه آزمایش دوم، تراکنش‌های تکی مشکوک و غیر مشکوک در چند حساب

ردیف	نتیجه جستجو برای کشف تراکنش مشکوک در بافت	زمان مصرفی
۱	موفق	۰۰:۲۰:۳۲
خروجی برنامه	02904389220058: 5873 Bardasht Alert (Contexual heterogeneous) ----> Interquartile Type 5873 WITHDRAW Alert (Greater Than Any Other In it's OWN Context) ----> Context Type 19312 Varizi Alert (Contexual heterogeneous) ----> Interquartile Type 19312 Balance Alert (Contexual heterogeneous) ----> Interquartile Type 19312 DEPOSIT Alert (Greater Than Any Other In it's OWN Context) ----> Context Type 19312 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 02458477240087: 10423 WITHDRAW Alert (Greater Than Any Other In it's OWN Context) ----> Context Type 19313 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 01966373980012: 17512 WITHDRAW Alert (Greater Than Any Other In it's OWN Context) ----> Context Type 19314 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 01394204190022: 19315 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 03823222020019:	

در این آزمایش به شماره‌حساب‌های ردیف یک تا چهار جدول ۴، تراکنش‌ها به نحوی اضافه شد که در بافت هر حساب مشکوک باشند و در ردیف پنج داده مشکوکی اضافه نشده و صرفاً یک تراکنش معمولی اضافه گردیده است. هدف این آزمایش بررسی این مطلب است که آیا الگوریتم پیشنهادی توانایی کشف تراکنش‌های مشکوک در بافت را دارد یا نه؟ با توجه به جدول ۵ تمام تراکنش‌های مشکوک در بافت کشف گردید، برخی از آن‌ها بر اساس مقایسه در بافت و برخی هم توسط رویه خلاصه پنج عددی کشف شدند. برای حساب ردیف پنج هیچ مورد مشکوکی یافت نشد.

آزمایش سوم. در این آزمایش به یک حساب، تعداد ۲۰ تراکنش گروهی طبق جدول ۶ اضافه

گردید و نتیجه بررسی به‌عمل‌آمده در جدول ۷ آورده شده است.

جدول ۶. آزمایش سوم، تراکنش‌های گروهی یک‌ساله مشکوک در یک حساب

شماره‌حساب انتخابی	حداقل مبلغ مانده‌حساب	حداکثر مبلغ مانده‌حساب	تراکنش از تاریخ	تراکنش تا تاریخ	زمان تراکنش از	زمان تراکنش تا	بافت
۰۱۳۹۴۲۰۴۱۹۰۰۲۲	۹۰۰۰۰۰۰۰	۱۲۰۰۰۰۰۰۰	۱۳۹۵/۱/۰۵	۱۳۹۵/۱۲/۷	۰۰:۰۰:۰۰	۲۳:۵۹:۵۹	۲۷

جدول ۷. نتیجه آزمایش سوم، تراکنش‌های گروهی یک‌ساله مشکوک در یک حساب

ردیف	نتیجه جستجو برای کشف تراکنش‌های مشکوک در بافت	زمان مصرفی
۱	موفق*	۰۰:۰۱:۰۱
خروجی برنامه	7672 WITHDRAW Alert (Greater Than Any Other In it's OWN Context) ----> Context Type 7677 Bardasht Alert (Contexual heterogeneous) ----> Interquartile Type 7677 WITHDRAW Alert (Greater Than Any Other In it's OWN Context) ----> Context Type 19307 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 19308 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 19309 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 19310 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 19312 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 19314 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 19316 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 19317 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 19319 DEPOSIT Alert (Greater Than Any Other In it's OWN Context) ----> Context Type 19319 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 19321 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 19322 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 19323 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 19324 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 19311 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 19313 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 19315 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 19318 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 19320 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 19325 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type 19326 BALANCE Alert (GreaterThan Any Other In it's OWN Context) ----> Context Type	

*تعداد تراکنش‌های مشکوک یافت شده بیشتر از تعداد تراکنش‌های تزریق‌شده است

هدف این آزمایش بررسی رفتار الگوریتم پیشنهادی در مواجه با اضافه کرد تعدادی تراکنش خارج از بافت به یک شماره‌حساب می‌باشد؟ درنتیجه تزریق تراکنش‌های مشکوک، برای جلوگیری از به هم خوردن رابطه بین مقادیر واریزی برداشتی و مانده‌حساب تراکنش‌ها مبالغ برخی تراکنش‌ها تغییر کرده است و در دسته تراکنش‌های مشکوک قرارگرفته‌اند و تراکنش اضافه‌ای که به‌عنوان مشکوک شناسایی‌شده است بدین دلیل است. همه تراکنش‌هایی که با بافت همخوانی ندارند، با روش خلاصه پنج عددی و یا درنتیجه مقایسه با بافت مربوطه کشف‌شده‌اند، جدول ۷.

آزمایش چهارم. در این آزمایش به یک حساب تعداد ۹۰ تراکنش گروهی در یک ماه طبق جدول

۸ اضافه گردید و اندازه پنجره مجموعه کاری ۳۰ روز قرار داده شد.

جدول ۸. آزمایش چهارم، تراکنش‌های گروهی یک‌ماهه مشکوک در یک حساب

شماره‌حساب انتخابی	حداقل مبلغ مانده‌حساب	حداکثر مبلغ مانده‌حساب	تراکنش از تاریخ	تراکنش تا تاریخ	زمان تراکنش از	زمان تراکنش تا	بافت
۰۱۹۵۰۱۲۶۶۲۰۰۳۸	۱۰۰۰۰۰۰۰۰	۳۰۰۰۰۰۰۰۰	۱۳۹۵/۵/۱	۱۳۹۵/۵/۳۰	۰۰:۰۰:۰۰	۲۳:۵۹:۵۹	۱۴

جدول ۹. نتیجه آزمایش چهارم، تراکنش‌های گروهی یک‌ماهه مشکوک در یک حساب

ردیف	نتیجه جستجو برای کشف تراکنش‌های مشکوک بر اساس پنجره مجموعه کاری	زمان مصرفی
۱	موفق*	۰۰:۰۲:۰۹
خروجی برنامه	01950126620038 During (1394/05/01,1395/05/30), Deposite OR Withdraw Suspicious (Contexual WorkingSet) ----> WorkingSet Type 01950126620038 : During (1395/05/01, 1395/05/30) Number of Transactions Suspicious.	
*فقط بر اساس پنجره مجموعه کاری جستجو انجام‌شده است		

هـــدف این آزمایش برآورد کارایی پنجره مجموعه کاری و پنجره مجموعه کاری لغزان در مواجه با تراکنش‌های اضافه‌شده به یک حساب است. خروجی الگوریتم پیشنهادی در جدول ۹ نشان می‌دهد هم در جمع مبالغ واریزی و برداشت و هم در تعداد تراکنش‌های ۱۳۹۵/۵/۱ تا ۱۳۹۵/۵/۳۰ فعالیت مشکوک تشخیص داده‌شده است.

آزمایش پنجم. در این آزمایش به ترتیب ۲۵۰، ۷۵، ۴۰، ۳۹۸، ۱۰۰ تراکنش گروهی مشکوک در زمان‌های مختلف به ۵ حساب اضافه گردید، جدول ۱۰.

جدول ۱۰. آزمایش پنجم، تراکنش‌های گروهی مشکوک در حساب‌های مختلف

شماره‌حساب انتخابی	حداقل مبلغ مانده‌حساب	حداکثر مبلغ مانده‌حساب	تراکنش از تاریخ	تراکنش تا تاریخ	زمان تراکنش از	زمان تراکنش تا	بافت
۰۲۳۳۰۰۰۰۷۲۰۰۷۷	۴۰۰۰۰۰۰۰۰	۶۰۰۰۰۰۰۰۰	۱۳۹۵/۳/۱	۱۳۹۵/۵/۳۰	۰۰:۰۰:۰۰	۰۷:۵۹:۵۹	۳۴
۰۳۴۳۴۸۵۷۳۷۰۰۷۹	۱۰۰۰۰۰۰۰۰	۱۵۰۰۰۰۰۰۰	۱۳۹۵/۸/۱	۱۳۹۵/۱۱/۱	۰۰:۰۰:۰۰	۲۳:۵۹:۵۹	۱۱
۰۳۳۴۱۰۹۴۲۸۰۰۱۷	۵۰۰۰۰۰۰۰	۱۲۵۰۰۰۰۰۰	۱۳۹۵/۲/۱	۱۳۹۵/۵/۱	۰۸:۰۰:۰۰	۱۷:۵۹:۵۹	۸
۰۱۲۷۰۱۶۵۰۹۰۰۷۸	۱۰۰۰۰۰۰	۲۰۰۰۰۰۰	۱۳۹۵/۵/۱	۱۳۹۵/۷/۱	۰۰:۰۰:۰۰	۲۱:۵۹:۵۹	۲۲
۰۲۳۸۱۶۷۷۸۸۰۰۷۹	۱۰۰۰	۱۰۰۰۰۰۰۰۰	۱۳۹۵/۲/۱۰	۱۳۹۵/۶/۷	۰۰:۰۰:۰۰	۲۴:۵۹:۵۹	۲۷

جدول ۱۱. نتیجه آزمایش پنجم، تراکنش‌های گروهی مشکوک در حساب‌های مختلف

ردیف	نتیجه جستجو برای کشف تراکنش‌های مشکوک بر اساس پنجره مجموعه کاری	زمان مصرفی
۱	موفق*	۰۲:۱۹:۲۱
۲	موفق*	
۳	موفق*	
۴	موفق*	
۵	موفق*	
خروجی برنامه	02330000720077: WorkingSet Withdraw OR Deposite Alert (Contexual WorkingSet) ----> WorkingSet Type WorkingSet Transaction Numbers Alert (Contexual WorkingSet) ----> WorkingSet Type 03434857370079: WorkingSet Transaction Numbers Alert (Contexual WorkingSet) ----> WorkingSet Type 03341094280017: WorkingSet Withdraw OR Deposite Alert (Contexual WorkingSet) ----> WorkingSet Type WorkingSet Transaction Numbers Alert (Contexual WorkingSet) ----> WorkingSet Type 01270165090078: WorkingSet Withdraw OR Deposite Alert (Contexual WorkingSet) ----> WorkingSet Type WorkingSet Transaction Numbers Alert (Contexual WorkingSet) ----> WorkingSet Type 02381677880079: WorkingSet Withdraw OR Deposite Alert (Contexual WorkingSet) ----> WorkingSet Type WorkingSet Transaction Numbers Alert (Contexual WorkingSet) ----> WorkingSet Type	

*فقط بر اساس پنجره مجموعه کاری جستجو انجام‌شده است

هدف این آزمایش برآورد کارایی الگوریتم پیشنهادی در مواجه با تراکنش‌های مشکوک گروهی در چندین حساب مختلف در زمان‌های متفاوت است. بخش نتیجه جدول ۱۱، نشان‌دهنده کشف تمام تراکنش‌های گروهی هم از حیث جمع مبالغ واریزی و برداشـــت و هم از حیث تعداد تراکنش برای

چند حساب مختلف است.

آزمایش ششـــم. هدف این آزمایش معین کردن این نکته است که نرخ تشخیص غلط الگوریتم در اعلام تراکنش‌هایی نرمال و غیر مشـــکوک به‌عنوان تراکنش مشکوک به پول‌شویی چقدر است و چه تعداد از تراکنش‌های غیر مشکوک را به‌غلط تراکنش مشکوک به پول‌شویی اعلام می‌کند. به همین منظور توســـط «نرم‌افزار شبیه‌ساز تولید حساب‌های بانکی» برای ۴۸ نفر از ۴۸ بافت مختلف و برای مدت‌زمان یک ســـال طبق جدول ۱۲، تعداد ۱۰،۰۸۲ تراکنش تولید گردید و بدون تزریق هیچ نوع تراکنش مشکوکی، تحت کاوش بر اساس «متغیرهای رفتاری» قرار گرفت، جدول ۱۳.

جدول ۱۲. مشخصات داده‌های استفاده‌شده در برآورد نرخ پیش‌بینی غلط الگوریتم پیشنهادی

ردیف	تعداد بافت	تعداد در هر بافت (نفر)	تعداد کل	تاریخ آغاز	تاریخ پایان	زمان شبیه‌سازی	تعداد تراکنش تولیدی
۱	۴۸	۱	۴۸	۱۳۹۵/۱/۱	۱۳۹۵/۱۲/۳۰	۰۰:۲:۴۹	۱۰،۰۸۲

درمجموع تعداد ۱۱۵ تراکنش نرمال، به‌عنوان تراکنش مشـــکوک شناسایی‌شـــده است، لذا طبق رابطه (۶) نرخ پیش‌بینی غلط الگوریتم پیشنهادی معادل ۱/۱۴٪ می‌باشد.

$$False\ Positive\ Rate = \frac{115}{10082} * 100 = 1.14\ \% \qquad (۶)$$

جدول ۱۳. نتیجه آزمایش ششم، برآورد نرخ پیش‌بینی غلط الگوریتم پیشنهادی

عنوان	نتیجه
زمان مصرفی	۱۳:۲۱:۰۸
بخشی از خروجی برنامه	01781782880009: 27 Varizi Alert (Contexual heterogeneous) ----> Interquartile Type 27 Balance Alert (Contexual heterogeneous) ----> Interquartile Type 51 Varizi Alert (Contexual heterogeneous) ----> Interquartile Type 51 Balance Alert (Contexual heterogeneous) ----> Interquartile Type 52 Bardasht Alert (Contexual heterogeneous) ----> Interquartile Type 59 Balance Alert (Contexual heterogeneous) ----> Interquartile Type ..... 02276821220087: 2553 Varizi Alert (Contexual heterogeneous) ----> Interquartile Type ..... 01132097340018: 2632 Varizi Alert (Contexual heterogeneous) ----> Interquartile Type 2632 Balance Alert (Contexual heterogeneous) ----> Interquartile Type 2633 Bardasht Alert (Contexual heterogeneous) ----> Interquartile Type 2648 Balance Alert (Contexual heterogeneous) ----> Interquartile Type 2656 Varizi Alert (Contexual heterogeneous) ----> Interquartile Type 2656 Balance Alert (Contexual heterogeneous) ----> Interquartile Type 2657 Bardasht Alert (Contexual heterogeneous) ----> Interquartile Type 2666 Balance Alert (Contexual heterogeneous) ----> Interquartile Type 2668 Varizi Alert (Contexual heterogeneous) ----> Interquartile Type 2668 Balance Alert (Contexual heterogeneous) ----> Interquartile Type

مقایسه روش ارائه‌شده با روش‌های دیگر. در روش‌های مقابله با پول‌شویی منتشره در مقالات مرتبط با پول‌شویی حوزه مالی از این نکته غفلت شده که اطلاعات مالی هر فرد می‌بایست نسبت به افراد هم‌سنخ و یا به‌عبارت‌دیگر هم‌بافت وی موردسنجش قرار گیرد، لذا این پژوهش این نکته را مدنظر قرارداد. در اغلب مقالات مرتبط، اطلاعات کاملی از دیتاست[۴۶] استفاده‌شده، ارائه نشده است و در مواردی هم که اطلاعاتی از پایگاه داده ارائه‌شده، داده‌های مورداســـتفاده ارائه نشده است و لذا امکان مقایسه کاملی وجود ندارد، بااین‌وجود جدول ۱۴ خلاصه تعدادی از مقالات ارائه‌شـــده در این زمینه را از حیث «تعداد داده‌های پایگاه داده»، True False Positive Rate[۴۷]، False Negative Rate[۴۸]، Positive Rate[۴۹]، زمان مصرفی و روش مورداستفاده نمایش می‌دهد. با توجه به ارائه کامل‌تر اطلاعات در مقاله کوثری‌لنگری و همکاران (۱۳۹۲) نسبت به سایر مقالات جدول ۱۴ در ادامه مقایسه مختصر مقاله یادشده و تحقیق حاضر ارائه خواهد شد.

جدول ۱۴. مقایسه روش‌های مختلف کشف موارد مشکوک به پول‌شویی

ردیف	نام مقاله	تعداد داده‌های پایگاه داده	FPR	TPR	FNR	زمان مصرفی	روش
۱	آلسکرو و همکاران (۱۹۹۷)	۳۷۶	٪۰	٪۸۵	-	-	شبکه عصبی
۲	چن و همکاران (۲۰۰۴)	۱۲۰۰۰	-	٪۹۰	-	-	دسته‌بندی
۳	ونگ و دونگ (۲۰۰۹)	۶۴۹۴۱	-	٪۷۲	-		خوشه‌بندی
۴	لی‌خاک و همکاران (۲۰۰۹)	۲۰۰۰۰۰۰	-	٪۱۰۰	-	۵ دقیقه	خوشه‌بندی و هوش مصنوعی
۵	لوپزروجاس و اَکسلسون (۲۰۱۲)	۴۸۶۹۷۷	٪۱/۵	٪۹۹/۹	-	۸ ساعت	دسته‌بندی
۶	کوثری‌لنگری و همکاران (۱۳۹۲)	۱۰۰۰۰۰	-	٪۹۱	-	-	درخت تصمیم- الگوریتم C۵.۰
۷	محمدی و کاظمی‌فرد (۱۳۹۳)	۸۵۴	٪۷/۳	٪۷۴	٪۲۵	-	خوشه‌بندی
	این پژوهش	۱۸۰۶۳۰۹	٪۱/۱۴	٪۱۰۰	٪۰	در جداول نتیجه ارائه‌شده است	مقایسه‌های آماری بر اساس بافت حساب بانکی

46 . Dataset

۴۷ . FPR (تشخیص غلط داده‌های درست، به‌عبارت‌دیگر تشخیص تراکنش‌ها به‌عنوان تراکنش مشکوک به پول‌شویی درحالی‌که تراکنش‌ها نرمال می‌باشند)

۴۸ . TPR (تشخیص درست داده‌های درست، به‌عبارت‌دیگر تشخیص تراکنش‌های غیر مشکوک به پول‌شویی به‌عنوان تراکنش نرمال)

۴۹ . FNR (تشخیص غلط داده‌های غلط، به‌عبارت‌دیگر عدم‌تشخیص تراکنش مرتبط با پول‌شویی به‌عنوان تراکنش مشکوک به پول‌شویی)

در کوثری‌لنگری و همکاران (۱۳۹۲) برای کشف و اعلام به‌موقع سرقت اطلاعات حساب از درخت تصمیم و تعداد ۷ متغیر برای تعیین الگوی رفتاری کاربران سامانه‌های بانکداری اینترنتی استفاده‌شده است، لذا تعدادی قانون «اگر- آنگاه»[۵۰] برای الگوی رفتاری کاربران به‌دست‌آمده است از آن به بعد درصورتی‌که اطلاعات حساب اینترنتی کاربری به سرقت رفته و رفتاری خلاف عرف صاحب حساب مشاهده شود قابل ره‌گیری خواهد بود. تحقیق حاضر تشخیص پول‌شویی در حوزه بانک می‌باشد و تمام انواع تراکنش‌های داخل شعبه، عابر بانک، اینترنت بانک، موبایل بانک و غیره را شامل می‌شود اما در تحقیق کوثری‌لنگری و همکاران (۱۳۹۲) تراکنش‌های اینترنتی هدف‌گیری شده است. تفاوت بعدی آنکه در تحقیق حاضر «واریز» و «برداشت» و حتی «مجموعه‌ای از برداشت‌ها» و «مجموعه‌ای از واریزها» هدف‌گیری شده و اصولاً در فرآیند پول‌شویی اطلاعات حساب الزاماً سرقت نمی‌شود و خود صاحبان حساب‌ها و چندین نفر یک سلسله واریزها و برداشت‌ها را در راستای گم کردن منشأ درآمدها و پول‌شویی انجام می‌دهند که حتی ممکن است یک تراکنش به‌خودی‌خود اصلاً مشکوک نباشد، حال‌آنکه در تحقیق کوثری‌لنگری و همکاران (۱۳۹۲) هدف ممانعت و یا کشف فعالیت افرادی است که اطلاعات حساب را سرقت کرده‌اند و تراکنش مشکوکی را انجام داده‌اند. در تحقیق کوثری‌لنگری و همکاران (۱۳۹۲) از روش درخت تصمیم استفاده‌شده است، روش درخت تصمیم در ذیل روش‌های دسته‌بندی قرار دارد حال‌آنکه روش این تحقیق، تشخیص با استفاده از ترکیب مفهوم داده پرت و مفهوم داده بافتاری و روش‌های مبتنی بر آمار می‌باشد. در تحقیق حاضر تأکید بر تعیین و تشخیص بافت افراد و سپس اقدام به سایر مراحل تشخیص بر اساس بافت فرد می‌باشد حال‌آنکه در تحقیق کوثری‌لنگری و همکاران (۱۳۹۲) ایجاد درخت تصمیم و استخراج قواعد «اگر- آنگاه» بدون توجه به بافت افراد انجام می‌گیرد، تفاوت بعدی این است که در تحقیق حاضر تمام حساب‌های یک فرد به‌عنوان یک حساب کل در نظر گرفته‌شده است ولی در تحقیق مزبور در این خصوص اشاره‌ای نشده است. به‌کارگیری متغیرهایی خاص در فرآیند تشخیص، شباهت دو تحقیق می‌باشد و تفاوت آن‌ها در نوع متغیرهای به‌کارگیری شده در فرآیند مزبور می‌باشد، متغیرهای استفاده‌شده در کوثری‌لنگری و همکاران (۱۳۹۲) شامل «تعداد خطا» در ورود موفق به سامانه، تعداد دفعات «ورود موفق» به سامانه، تعداد مراکز «آی‌اس‌پی» استفاده‌شده جهت ورود به سامانه، تعداد «مرورگر» های مختلف جهت ورود به سامانه، تعداد «آی‌پی» های مختلف در زمان ورود کاربر به سامانه، بیشترین «تعداد حواله» با سقف پایین و میانگین «مبلغ حواله» که کاربر با سقف پایین انجام داده است می‌باشد و در تحقیق حاضر رویه به این صورت است که ابتدا بافت افراد بر اساس متغیرهای بافتاری شامل «شغل» (و یا میزان درآمد ماهانه) صاحب حساب، «تعداد اعضای خانواده» (تعداد افراد تحت تکفل) صاحب حساب، «وضعیت مالکیت مسکن» (صاحب‌خانه و یا مستأجر بودن)، «وضعیت تملک اتومبیل» و «تعداد اتومبیل‌های تحت تملک» تعیین و سپس بر اساس متغیرهای رفتاری شامل «مقدار ماکزیمم مانده‌حساب»، «ماکزیمم مبلغ واریزی»، «ماکزیمم مبلغ برداشت»، «تعداد تراکنش در یک برهه زمانی»، «جمع مبالغ واریزی در یک برهه زمانی» و «جمع مبالغ برداشتی در یک برهه

50 . IF –Then–Else

زمانی»، رفتار آن‌ها با سایر افراد هم‌بافت ایشان مقایسه می‌گردد، دو متغیر «تعداد حواله» و «مبلغ حواله» تا حدی به متغیرهای این تحقیق شبیه می‌باشند اما اصولاً در این تحقیق از میانگین متغیرها پرهیز شده است. تفاوت متغیرهای دو تحقیق به ایده‌های متفاوت آن‌ها برمی‌گردد. ازلحاظ تعداد نمونه استفاده‌شده در تحقیق حاضر حدود یک‌میلیون و هشت‌صد هزار تراکنش را به کار گرفته است و در تحقیق کوثری‌لنگری و همکاران (۱۳۹۲) حدود صد هزار تراکنش بکار رفته است.

❐ ۸. نتیجه‌گیری و پیشنهادها

رفتارهای مالی افراد بنا بر مقدار درآمد، وضعیت مالکیت مسکن و چندین پارامتر دیگر از یکدیگر متمایز می‌شود و با توجه به بافت آن‌ها در یک محدوده قابل‌تصور است، به‌طوری‌که برخی تراکنش‌ها در یک بافت طبیعی و در بافتی دیگر غیرطبیعی است لذا در این پژوهش برای شناسایی تراکنش‌های مشکوک به پول‌شویی، یک‌راه حل مبتنی بر بافت هر حساب بانکی و تحلیل‌های آماری ارائه گردید، به‌طوری‌که ابتدا افراد بر اساس متغیرهای بافتاری مشابه در یک گروه قرار می‌گیرند و الگوی رفتارهای مالی ایشان همان الگوی رفتاری افراد هم‌بافت آن‌ها خواهد بود، از این مرحله به بعد متغیرهای رفتاری ایشان با سایر افراد هم‌بافت آن‌ها مقایسه می‌شود. مبلغ تراکنش‌ها و تعداد تراکنش‌ها دو وجه مهم در کشف پول‌شویی هستند. درروش ارائه‌شده این دو وجه به‌طور کامل توسط داده‌ساختارهای «پنجره مجموعه کاری»، «پنجره مجموعه کاری لغزان» برای مبالغ مانده‌حساب، واریزی و برداشت و نیز برای جمع مبالغ واریزی به‌حساب بانکی، جمع مبالغ برداشت از حساب بانکی و نیز برای تعداد تراکنش‌ها در «طول پنجره کاری» پوشش داده شد، برای تشخیص تراکنش‌های مشکوک به پول‌شویی علاوه بر مقایسه متغیرهای رفتاری بر اساس روابط کوچک‌تری و بزرگ‌تری از آزمون توکی (۱۹۷۷) نیز بهره برده شد. در عموم تحقیقات دیگر که در خصوص کشف تخلفات مالی ازجمله پول‌شویی صورت گرفته است از روش‌های مبتنی بر خوشه‌بندی و یا دسته‌بندی و تکنیک‌های ذیل این روش‌ها استفاده کرده‌اند، در این تحقیق برای اولین بار از ترکیب روش‌های مبتنی بر آمار با مفهوم داده پرت و متغیرهای بافتاری برای نیل به کشف تراکنش‌های مشکوک به پول‌شویی استفاده‌شده است. دیگر آنکه در این تحقیق علاوه بر در نظر گرفتن تراکنش‌ها به‌صورت مجزا، تراکنش‌ها به‌صورت مجموعه‌ای نیز مورد مقایسه قرار گرفتند، در این حالت مجموع مبالغ واریزی و برداشت و نیز تعداد تراکنش‌ها در برهه‌های زمانی خاص برای مقایسه استفاده‌شده است و این در حالی است که اغلب مقالات دیگر، تراکنش‌ها را مجزا بررسی کرده‌اند.

باهدف آزمایش ایده مطرح‌شده در این پژوهش، یعنی تأثیر توجه به بافت صاحبان حساب‌های بانکی در تشخیص فعالیت‌های مرتبط با پول‌شویی شش آزمایش با اهداف خاص ترتیب داده شد، در آزمایش‌های اول تا پنجم عملاً اعتبار روش پیشنهادی برای کشف تراکنش‌های مشکوک بعد از تزریق تعدادی تراکنش مشکوک مصنوعی به بوته آزمایش قرار گرفت که نتایج به‌دست‌آمده حاکی از موفقیت مدل مطروحه بود. در آزمایش ششم باهدف کشف میزان تشخیص غلط الگوریتم پیشنهادی ترتیب داده شد که درصد قابل‌قبول ۱/۱۴٪ (یک ممیز چهارده صدم درصد) تشخیص غلط را نتیجه

داد. به دلیل یکسان نبودن داده‌های آزمایش‌های این تحقیق و داده‌های سایر تحقیقات، مقایسه همه‌جانبه روش پیشنهادی این پژوهش و سایر روش‌ها ممکن نیست، بااین‌وجود بر اساس جدول ۱۴ و آزمایش‌های انجام‌شده، برتری این روش نسبت به سایر روش‌ها روشن است و پیشنهاد می‌گردد، نهادهای نظارتی و نهادهای مالی ازجمله بانک‌ها از نتایج این پژوهش مخصوصاً توجه به مفهوم بافت، استفاده نمایند. با توجه به اشتغال افراد متخصص در فرآیند شستشوی پول‌های کثیف و توانایی ایشان در درست نمایی و قانونی نمایی رفتارهای مالی پول‌شویان، صرف توجه و کنترل قوانین مصوب قانون‌گذاران، منجر به کشف تمام موارد تخلف نخواهد شد و این تأییدِ توجه به بافت افراد در تشخیص رفتارهای خلاف عرف می‌باشد زیرا ممکن است تراکنشی وفق قوانین و مقررات باشد اما با توجه به بافت عامل آن، مشخص شود از استطاعت اعضاء آن بافت خارج است. در روش‌های پیشنهادی سایر مقالات مرتبط با پول‌شویی در حوزه بانک، بافت افراد مدنظر قرار نگرفته است، لذا پیشنهاد می‌گردد پژوهشگران روش‌های مطروحه را با توجه به بافت افراد به‌روزرسانی کرده و نتایج را مجدداً مورد ارزیابی و مقایسه قرار دهند.

عدد میانگین با استفاده از واردکردن اعداد کوچک در مجموعه، تحت تأثیر قرار می‌گیرد، مثلاً یک تراکنش یک میلیارد تومانی در کنار ۱۰ تراکنش خرید شارژ هزار تومانی، میانگین حدود صد میلیون را نتیجه می‌دهد که حدود ۹۰۰ درصد نسبت به بزرگ‌ترین تراکنش مجموعه، کاهش نشان می‌دهد، لذا پیشنهاد می‌گردد استفاده از میانگین و ارائه راهکارهای مبتنی بر آن در روش‌های تشخیص تخلفات مالی من‌جمله پول‌شویی به‌طور هوشمندانه و در کنار سایر پارامترها به‌کارگیری شود.

پیشنهاد می‌شود تمام حساب‌های یک فرد به‌عنوان یک حساب کل در نظر گرفته شود، در مواردی که حساب‌های یک فرد مشکوک تشخیص داده شد و یا فردی سابقه تخلفات مالی داشت، پیشنهاد می‌گردد مجموعه حساب‌های افراد غیرمستقل خانواده وی باهم به‌عنوان یک حساب کل در نظر گرفته و بررسی شوند. همچنین استفاده از کد ملی به‌جای شماره‌حساب و افتتاح حداکثر یک حساب از هر نوع حساب در کل شعبات یک بانک -و دیدن آن به‌عنوان یک حساب کل- به بانک‌ها و نهادهای مالی پیشنهاد می‌گردد، در این صورت گام مهمی در شناسایی پول‌شویی در مرحله یکپارچه‌سازی انجام می‌گیرد.

با توجه به حجم عظیم محاسبات و مقایسه‌های انجام‌گرفته در حین انجام فرآیند تشخیص تراکنش‌های مشکوک و با توجه به ماهیت روش ارائه‌شده در این پژوهش، تکنیک‌های الگوریتم‌های موازی[51] و برنامه‌نویسی پویا[52] جهت نیل به کاهش انجام محاسبات مکرر و کاهش زمان اجرای الگوریتم در مرحله پیاده‌سازی نهایی قابل‌استفاده خواهد بود.

مشکل عمده پیش روی محققان حوزه مبارزه با پول‌شویی در حیطه تراکنش‌های بانکی، فراهم نبودن دیتاست واقعی از تراکنش‌های بانکی است، درصورتی‌که «ایده این پژوهش» و نتایج آزمایش آن بر روی تراکنش‌های شبیه‌سازی‌شده موردتوجه نهادهای سیاست‌گذار حوزه بانک‌ها قرار گیرد،

51 . Parallel Algorithms
52 . Dynamic Programming

پیشـــنهاد می‌گردد پروژه‌های تحقیقاتی با مشارکت بانک‌ها و مراکز دانشگاهی و محققان علاقه‌مند، متضمن به امکان دسترسی پژوهشگران به تراکنش‌های بانکی تعریف گردد.

مالیات از دو منظر در رابطه با پول‌شـــویی دارای اهمیت اســـت، اول آنکه فرار مالیاتی خود یکی از وجه‌های مهم پول‌شـــویی اســـت زیرا افراد مالیات درآمدهای قانونی خود را پرداخت نکرده‌اند و لذا مرتکب پول‌شـــویی شده‌اند، دوم آنکه گردش مالی صاحبان درآمدهای غیرقانونی و مالیات پرداختی آن‌ها همخوانی ندارد لذا میزان مالیات پرداختی افراد و شرکت‌ها می‌تواند به‌عنوان عاملی برای کشف پول‌شـــویی و حتی جرم و جنایت استفاده شود، لذا پیشنهاد می‌گردد مالیات پرداختی افراد به‌عنوان متغیر بافتاری در نظر گرفته‌شده و با استفاده از آن گردش مالی و دارایی‌های افراد در بانک‌ها جهت کشـــف پول‌شویی وارسی شود. همچنین پیشنهاد می‌شـــود سازمان امور مالیاتی کشور مقدار مالیات پرداختی افراد را جهت بهره‌برداری و محاســـبه درآمدهای قانونی و مقایسه با موجودی حساب‌ها در اختیار نهادهای نظارتی و بانک‌ها قرار دهند.

عمده‌ترین محدودیت‌های پیش روی این تحقیق، عدم وجود دیتاســـت تراکنش‌های بانکی افراد و اطلاعات مرتبط با متغیرهای بافتاری (وضعیت مسکن، تملک ماشین، تعداد افراد تحت تکفل و غیره) ایشـــان، کامل نبودن و ناکافی بودن اطلاعات ارائه‌شده از سوی ســـازمان‌های ذی‌ربط با مستمسک محرمانه بودن و عدم انتشار اطلاعات روش‌های موفق به‌کارگیری شده در سایر کشورها بوده است.

منابع

۱- تذهیبی، فریده. (۱۳۹۴). پول‌شویی و روش‌های مبارزه با آن. تهران: انتشارات جنگل.

۲- ساکی، محمدرضا. (۱۳۹۳). آشنایی با جرم پول‌شویی. تهران: انتشارات جاودانه.

۳- کوثری‌لنگری، روح‌الله. مقدم‌چرکری، نصرالله. وحدت، داوود. (۱۳۹۲). «به‌کارگیری الگوریتم‌های درخت تصمیم‌گیری جهت کشف رفتارهای مشکوک در بانکداری اینترنتی». پژوهشگاه علوم و فناوری اطلاعات ایران، ۲۸(۳)، تهران.

۴- محمدی، آزاده. کاظمی‌فرد، محمد. (۱۳۹۳). «یک‌راه حل مبتنی بر داده‌کاوی برای مبارزه با پول‌شویی در پایگاه داده بانک». اولین کنفرانس داده‌کاوی در صنعت بانکداری.

5. Aleskerov, E., Freisleben, B., & Rao, B. (2002). Cardwatch: A neural network based database mining system for credit card fraud detection. Proceedings of the IEEE/IAFE 1997 Computational Intelligence for Financial Engineering (CIFEr) (pp. 220-226), 1997. New York, NY: IEEE.

6. Bureau of International Narcotics and Law Enforcement Affairs. Countries/ Jurisdictions of Primary Concern – Iran. (2016). (Online). Available at: https://www.state.gov/j/inl/rls/nrcrpt/2016/vol2/253407.htm (Accessed: 31 November 2016).

7. Chambers, J.M., Cleveland, W.S., Kleiner, B., & Tukey, P. A. (1983). "Graphical Methods for Data Analysis", Boston: Wadsworth International Group.

8. Chen, R. C., Chiu, M. L., Huang, Y. L., & Chen, L.T. (2004). "Detecting credit card fraud by using questionnaire-responded transaction model based on support vector machines". Springer Berlin Heidelberg, 3177, 800-806.

9. Coleman, D. Box Plot with Minitab. (2015). (Online). Available at: https://www.leansigmacorporation.com/box-plot-with-minitab (Accessed: 31 November 2016).

10. Gao, S., Xu, D., Wang, H., & Wang, Y. (2007). Intelligent anti-money laundering system. IEEE International Conference on Service Operations and Logistics, and Informatics (pp. 851-856), 2006. Shanghai, China: IEEE.

11. Han, J., Pei, J., & Kamber, M. (2011). Data mining: concepts and techniques, San Francisco. CA, U.S.: Morgan Kaufmann.

12. Le-Khac, N. A., Markos, S., & Kechadi, M. T. (2009). "Towards a new data mining-based approach for anti-money laundering in an international investment bank". International Conference on Digital Forensics and Cyber Crime. Springer, 31, 77-84.

13. Le-Khac, N. A., Markos, S., & Kechadi, M. T. (2010). "A data mining-based solution for detecting suspicious money laundering cases in an investment bank". Second International Conference on Advances in Databases Knowledge and Data Applications.

14. Lopez-Rojas, E. A., & Axelsson, S. (2012). "Multi Agent Based Simulation (MABS) of Financial Transactions for Anti Money Laundering (AML)". 17th Nordic Conference on Secure IT Systems. Karlskrona: Blekinge Institute of Technology.

15. Manjunath, K. V., (2015). "Data Mining Techniques for Anti Money Laundering". International Journal of Advanced Research in Science, Engineering and Technology. 2(8), 819-823.

16. Phua, C., Lee, V., Smith, K., & Gayler, R. (2010). A comprehensive survey of data mining-based fraud detection research. arXiv preprint arXiv:1009.6119.

17. Robinson, J. (1996). The Laundrymen: inside money laundering, the world's third-largest business. New York, U.S.: Arcade.

18. Schott, P. A., World Bank., & International Monetary Fund. (2006). Reference guide to anti-money laundering and combating the financing of terrorism. Washington, D.C., U.S.: World Bank.

19. Suresh. Ch., Thammi-Reddy, K., & Sweta, N. (2016). "A Hybrid Approach for Detecting Suspicious Accounts in Money Laundering Using Data Mining Techniques". International Journal of Information Technology and Computer Science (IJITCS), 8(5), 37-43.

20. Tukey, J.W., (1977). Exploratory data analysis. Mass, U.S.: Addison-Wesley Pub. Co.

21. U.S. Congress, Office of Technology Assessment, Information. (1995). Technologies for Control of Money Laundering, Washington, DC: U.S.. Government Printing Office.

22. United Nations Office on Drugs and Crime. Money-laundering and globalization. (2007). (Online). Available at: https://www.unodc.org/unodc/en/money-laundering/globalization.html (Accessed: 31 November 2016).

23. Wang, X., & Dong, G. (2009). "Research on money laundering detection based on improved minimum spanning tree clustering and its application". 2009 Second International Symposium on Knowledge Acquisition and Modeling, 2, 62-64.