دانشگاه تربیت معلم

دانشکده فنی و مهندسی

پایان نامه

جهت اخذ درجه کارشناسی ارشد

رشته هوش مصنوعی

کشف و استخراج الگوهای ناهنجاری در کارتهای پرداخت الکترونیک

استاد راهنما:

دکتر رویا امجدی فرد

استاد مشاور:

دکتر جمشید شنبه‌زاده

دانشجو:

آرش زارع دستجردی

مهرماه ۱۳۹۰

بسم الله الرحمن الرحيم

قدردانی و تشکر

برخود لازم می‌دانم تا از خانواده‌ام به عنوان مشوقان اصلی من در تحصیل علم و دانش تشکر فراوان داشته‌باشم. همچنین از دوستان عزیزم که در طی نگارش این تحقیق کمک‌های شایانی به بنده کردند کمال تشکر را دارم.

برخود واجب می‌دانم از استاد راهنمای دلسوزم سرکار خانم دکتر رویا امجدی‌فرد به‌دلیل راهنمایی‌های کارگشایشان کمال تشکر را داشته‌باشم.

چکیده

افزایش روزافرون مبادلات پولی الکترونیکی و نقش پررنگ کارت‌های پرداخت الکترونیکی[1] در تسهیل این مبادلات و نوپابودن نحوه‌ی استفاده از این روش تبادل پولی، زمینه‌هایی را برای افراد متقلب و سودجو فراهم می-آورد تا با استفاده از آشنایی کمتر کاربران با راهکارهای صحیح تبادل الکترونیک، اقدام به کلاهبرداری از مشتریان بانک‌ها و موسسات مالی کنند.

در این تحقیق سعی‌شده‌است تا با استفاده از روش‌های استخراج الگو، رفتار مشتری در تعامل با بانک بررسی شده و نحوه خرج‌کردن مشتری و چگونگی عملیات برداشت از حساب وی با توجه به پارامترهای متفاوت استخراج شود و از الگوی[2] استخراج‌شده برای هر مشتری، جهت بررسی صحت و تطابق تراکنش[3]های جدید وی با الگوی رفتاری او استفاده شود. برای استخراج الگوی[4] رفتاری مشتری، کاشف رفتاری هر مشتری از ترکیب مدل-های مارکوف پنهان[5] که هریک وظیفه‌ی تصمیم‌گیری براساس شرایط خاصی را دارند تشکیل شده است. به‌ازای هریک از مدل‌های مارکوف پنهان موجود در کاشفِ مختصِ یک مشتری، میزان قابل اتکا بودن احتمال خروجی از آن به عنوان عاملی جهت بهبود دقت احتمال کل (احتمال حاصل از ادغام[6] نتایج) به‌کار می‌روند.

با استفاده از این تقسیم‌بندی اختیارات و وظایف، میزان خطای کاشف کاهش پیداکرده همچنین صحت تطابق تراکنش جدید با تمامی حالات ممکن رفتاری مشتری مقایسه شده و درنهایت احتمال خروجی از کاشف به ازای تراکنش دارای میزان بالایی از قابلیت اتکا نسبت به سایر روش‌ها است.

کلمات کلیدی: استخراج الگو، الگوی ناهنجار، کلاهبرداری، خوشه‌بندی، مدل مارکوف پنهان، ادغام دسته‌بند[7]

---

[1] Electronic Payment Card
[2] Pattern
[3] Transaction
[4] Pattern Recognition
[5] Hidden Markov Model
[6] Fusion
[7] Classifier Fusion

فهرست

فصل ۱. مقدمه ۱

۱.۱ مقدمه ۲

۲.۱ چالش های پیشرو ۴

۳.۱ تعاریف ۷

۱.۳.۱ انواع ناهنجاری ۷

۲.۳.۱ ناهنجاری های تجمعی ۸

۴.۱ کاربردهای کشف ناهنجاری ۱۰

۱.۴.۱ کشف کلاهبرداری ۱۱

۱.۱.۴.۱ کشف کلاهبرداری در کارت‌های اعتباری ۱۱

۱.۱.۱.۴.۱ انواع کلاهبرداری کارت اعتباری ۱۱

۲.۴.۱ کشف نفوذ ۱۲

۵.۱ کارهای انجام‌شده در این تحقیق ۱۲

۶.۱ ساختار پایان‌نامه ۱۳

فصل۲. مروری بر پیشینه پژوهش ۱۴

۱.۲ مرور روش های باسرپرست ۱۵

۱.۱.۲ شبکه های عصبی ۱۶

۲.۱.۲ روش های مبتنی بر قاعده ۱۸

۱.۲.۱.۲ درخت الگوی مستمر ۱۸

۲.۲ مرور روش‌های بدون سرپرست ۲۰

۱.۲۲ خروجی کشف ناهنجاری ۲۰

۲.۲.۲ روش‌های مبتنی بر خوشه‌بندی ۲۱

۳.۲ مروری بر روش‌های ترکیبی ۲۳

۱.۳.۲ ترکیب مدل مارکوف پنهان و خوشه‌بندی ۲۳

فصل ۳. ابزارپژوهش ۲۵

۱.۳ مقدمه ۲۶

۲.۳ دلایل استفاده از خوشه‌بندی ۲۸
۳.۳ روش های خوشه‌بندی مورد استفاده ۲۸
۱.۳.۳ مروری بر الگوریتم K-means ۲۹
۲.۳.۳ مروری بر الگوریتم X-means ۳۰
۳.۳.۳ کاربرد خوشه‌بندی ۳۴
۴.۳ کاشف ۳۶
۵.۳ مدل مارکوف پنهان ۳۹
۱.۵.۳ تعاریف و کارکردهای مدل مارکوف پنهان ۳۹
۲.۵.۳ ساختار و اجزای مدل مارکوف پنهان ۴۰
۳.۵.۳ آموزش مدل مارکوف پنهان ۴۲
فصل۴. اجرای روش پیشنهادی ۴۶
۱.۴ آموزش زیرمجموعه‌های تصمیم‌گیری ۴۷
۱.۱.۴ کلیات روش آموزش ۴۷
۲.۱.۴ پیاده سازی آموزش کاشف مختص هر مشتری ۴۹
۲.۴ مجموعه داده ۵۰
۳.۴ آموزش کاشف ۵۰
۱.۱.۱.۳.۴ آموزش زیرمجموعه‌های تصمیم‌گیری کاشف ۵۱
۱.۱.۱.۱.۳.۴ خوشه‌بندی تراکنش‌ها ۵۱
۲.۱.۱.۱.۳.۴ آموزش اجزای زیرمجموعه‌های تصمیم‌گیری ۵۲
۱.۲.۱.۱.۱.۳.۴ پیاده‌سازی آموزش مدل مارکوف پنهان ۵۲
۲.۱.۱.۳.۴ تجمیع نتایج اجزاء تصمیم‌گیری در زیرمجموعه‌های تصمیم‌گیری ۵۳
۳.۱.۱.۳.۴ تجمیع نتایج زیرمجموعه های تصمیم‌گیری ۵۶
۴.۴ آزمون ۵۸
۱.۴.۴ رویه ارزیابی ۵۸
۲.۴.۴ آزمون‌ها ۵۹
۱.۲.۴.۴ آزمون مدل مارکوف پنهان و خوشه‌بندی خاص‌مشتری ۶۰

۲.۲.۴.۴ آزمون زیرمجموعه‌های تصمیم‌گیری ۶۵
۳.۲.۴.۴ آزمون کاشف ۶۶
فصل ۵. جمع بندی و پیشنهادها ۷۲
۱.۵ جمع بندی ۷۳
۲.۵ پیشنهادها ۷۵
مراجع ۷۶
واژگان ۸۰

فهرست اشکال

شکل ۱.۱ نمونه‌هایی از انواع ناهنجاری [2] ۳

شکل ۲.۱ نمونه‌های از بروز ناهنجاری در رفتار مشتریان در ماه فروردین ۸

شکل ۱.۲ مدل مارکوف پنهان مخصوص هر مشتری بر اساس خوشه های ثابت مبلغ تراکنش کم، متوسط و زیاد [16] ۲۳

شکل ۱.۳ نحوه قرارگیری نمونه‌ها و خوشه‌ها پس از خوشه‌بندی درون خوشه‌ای ۳۲

شکل ۲.۳ اتمام مرحله ۱ از الگوریتم X-Means [20] ۳۳

شکل ۳.۳ محاسبه امتیاز خوشه‌ها در الگوریتم X-Means [20] ۳۳

شکل ۴.۳ اتمام مرحله ۲ از الگوریتم X-Means [20] ۳۴

شکل ۵.۳ ساختار کاشف مختص هر مشتری ۳۷

شکل ۶.۳ زنجیره مارکوف با ۵ حالت [22] ۴۰

شکل ۷.۳ نحوه عملکرد الگوریتم forward-backward [23] ۴۲

شکل ۱.۴ نحوه ی تصمیم‌گیری در هر یک از زیر مجموعه‌های تصمیم گیر بر اساس تصمیمات اجزاء ۵۴

شکل ۲.۴ وزن اعتباری بر اساس دوره زمانی ۵۵

شکل ۳.۴ تعیین وضعیت نهایی تراکنش بر حسب نتایج زیرمجموعه‌ها ۵۷

شکل ۴.۴ نتایج مدل مارکوف پنهان و خوشه بندی خاص مشتری براساس مجموعه صفات مختلف براساس معیار میزان ضرر تحمیلی ۶۳

شکل ۵.۴ نتایج مدل مارکوف پنهان و خوشه بندی خاص مشتری بر اساس مجموعه صفات متخلف براساس معیار میزان دقت ۶۴

شکل ۶.۴ مدل مارکوف مشتری شماره ۲ در مجموعه صفات ۶۵

شکل ۷.۴ مقایسه نتایج کاشف، زیرمجموعه های تصمیم گیری و الگوریتم پایه [16] بر اساس میزان ضرر تحمیلی ۶۹

شکل ۸.۴ مقایسه نتایج کاشف، زیرمجموعه های تصمیم گیری و الگوریتم پایه [16] براساس میزان دقت ۷۰

شکل ۹.۴ نمودار نحوه تعیین وضعیت ۱۰۰ تراکنش اخیر مشتری شماره ۲ ..... ۷۰

فهرست جداول

جدول ۱ مراحل اجرای یک دور الگوریتم X-means ۳۴
جدول ۲ صفات موجود در تراکنش مشتری ۵۰
جدول ۳ تراکنش‌های مورد پردازش هر یک از مشتریان مورد آزمون ۶۰
جدول ۴ نتایج مدل مارکوف پنهان و خوشه‌بندی خاص مشتری بر اساس مبلغ تراکنش به ازای تعدادی از مشتریان مورد آزمون ۶۱
جدول ۵ نتایج مدل مارکوف پنهان و خوشه‌بندی خاص مشتری بر اساس صفات مبلغ، موجودی و تاریخ تراکنش به ازای تعدادی مشتریان مورد آزمون ۶۲
جدول ۶ احتمالات شروع مدل مارکوف پنهان شکل ۶.۴ ..... ۶۴
جدول ۷ نتایج آزمون به ازای زیرمجموعه‌های تصمیم گیری متفاوت ۶۶
جدول ۸ نتایج آزمون کاشف مشتریان ۶۷
جدول ۹ نتایج کلی روش خوشه‌بندی ثابت و در نظر گرفتن تمامی تراکنش های قبلی برای آزمون تراکنش جدید [16] ۶۸

فصل ۱ :

مقدمه

۱.۱ مقدمه

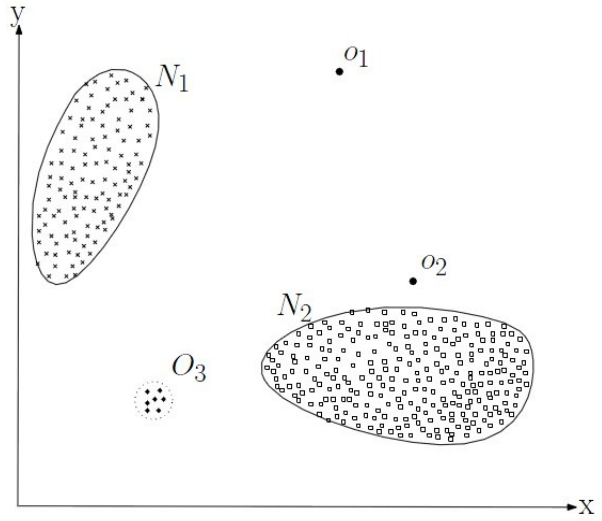
گسترش مبادلات مالی برخط[1] و استفاده‌ی روزافزون بانک‌ها و موسسات مالی و مشتریان آن‌ها از این نوع خدمات موجب تسهیل عملیات بانکی و مالی برای هر دو طرف و کاهش هزینه‌های جاری طرفین در طی تبادلات مالی می‌شود. از سوی دیگر نوپا بودن این‌گونه مبادلات و آشنایی کمتر کاربران موسسات با جزئیات این رویه‌ها، زمینه‌هایی را برای اشخاص ثالث فراهم می‌کند تا با استفاده از دانش کم کاربران و با استفاده از نقاط ضعف موجود در رویه‌ها، اقدام به کلاهبرداری از مشتریان و بانک‌ها کنند. اجرای این‌گونه اعمال کلاهبردارانه به‌دلیل آنکه توسط شخص ثالث و با هدف کسب منفعت انجام می‌شود، نظم رفتار مالی مشتری را برهم‌زده و باعث ایجاد ناهنجاری-هایی در رفتار کلی مشتری می‌شوند.

هدف کلی از انجام این تحقیق کشف ناهنجاری[2]های ایجادشده در رفتار مالی مشتریان با استفاده از استخراج رفتار مشتری بر اساس روند تبادلات مالی وی است. تمرکز این تحقیق برروی یافتن تراکنش‌هایی است که در آن‌ها مشتری مورد کلاهبرداری قرار گرفته است و بدون اطلاع و آگاهی وی مبالغی از حساب او کسر شده است.

ناهنجاری به الگویی[3] گفته می‌شود که به هیچ یک از رفتار[4]ها و الگوهای نرمال[5] تعلق نداشته باشد[1] . همچنین در تعریف دیگری، نمونه‌هایی ناهنجار تلقی می‌شوند که با دیگر داده‌ها سازگاری[6] نداشته باشند[2] . نمونه‌هایی که به اشتباه توسط فرد خبره طبقه‌بندی[7] شده باشند نیز در دسته نمونه‌های ناهنجار قرار می‌گیرند. برای مثال اگر نمونه A به دسته X نسبت داده شده باشد در حالیکه از منظر ماهوی به دسته Y تعلق داشته باشد، آنگاه این طبقه بندیِ از پیش تعیین شده نیز نشانگر یک نوع ناهنجاری است.

شکل ۱.۱نمونه‌هایی از ناهنجاری را نشان می‌دهد[2] :

[1] Online
[2] Anomaly
[3] Pattern
[4] Behavior
[5] Normal
[6] Consistency
[7] Classification



شکل ۱.۱ نمونه هایی از انواع ناهنجاری [2]

در شکل ۱.۱ نقاط o_1 و o_2 نقاطی ناهنجار و مجموعه نقاط موجود در ناحیه O_3 ناحیه[1]ای ناهنجار را تشکیل می‌دهند.

کشف ناهنجاری نقش بسزایی در تامین امنیت و تضمین صحت عملکرد سیستم های مالی دارد. برای مثال در سیستم‌های بانکی، پایش نحوه و تاریخچه عملکرد کارت‌های اعتباری به دلیل جلوگیری از اعمال تغییرات ناگهانی در میزان موجودی کارت‌ها توسط شخصی غیر از صاحب کارت، از اهمیت بالایی برخوردار است. از کاربردهای دیگر کشف ناهنجاری در سیستم‌های بانکی می‌توان به بررسی سوابق و عملکرد مالی و اعتباری درخواست‌کنندگان وام اشاره کرد[2] . بدین ترتیب که سوابق فرد در مقایسه با سوابق سایر درخواست کنندگان وام که بازپرداخت وام خود را به‌خوبی انجام داده‌اند سنجیده می‌شود و درصورتیکه سوابق او نسبت به این دسته از افراد نوعی ناهنجاری تلقی شود، آنگاه بانک با در نظر گرفتن ریسک مربوطه، اقدام به اعطای وام به درخواست‌کننده می‌کند. مشابه رویه اعطای وام برای درخواست‌های صدور بیمه نامه نیز وجود دارد.

[1] Region

۲.۱ چالش های پیش‌رو

با افزایش روزافزون درگاه‌های اتصال الکترونیک[1] و به تبع آن اضافه‌شدن مشکلات امنیتی مرتبط با هریک از درگاه‌ها، پتانسیل اجرای موفقیت‌آمیز اعمال کلاهبردارانه به‌مرور افزایش می‌یابد [3]. به کارگیری هریک از درگاه‌ها موجب می‌شود تا راه‌های جدیدی برای کلاهبرداری ایجاد شود که مقابله و پیش‌گیری از هریک از راهکارهای جدید نیازمند صرف وقت و هزینه است. برای مثال در زمینه کلاهبرداری‌های کارت اعتباری با اضافه شدن درگاه‌هایی نظیر پایانه‌های فروش[2] ، فروشگاه‌های اینترنتی و فروشگاه‌های متصل به دستگاه‌های دستی[3] مانند تلفن‌های همراه، به ازای هریک از این پایانه‌ها روش‌های جدید کلاهبرداری ایجاد شده است. متخصصان شرکت VISA در سال ۲۰۰۶ پیش‌بینی کردند که تا سال ۲۰۰۸ میزان کلاهبرداری در کارت‌های اعتباری ۶۵٪ افزایش پیدا کرده و این رقم به متوسط ۱۱ دلار به ازای هرکارت در سال برسد [3].

داده‌های کشف کلاهبرداری[4] به مقدار زیاد نامتوازن[5] و ناهمگون هستند[4] [5] . بدین معنی که تعداد نمونه‌های ناهنجار نسبت به نمونه‌های نرمال بسیار کم است و درصد کمی از کل مجموعه نمونه‌های موجود در یک مجموعه داده[6] را تشکیل می‌دهند. این عدم توازن درحدی است که اگر سیستمی تمام نمونه‌ها را نرمال تشخیص دهد هم خطای چندانی از نظر درصد خطا به وقوع نمی‌پیوندد اما به دلیل میزان ضرری که هر نمونه کلاهبرداری متوجه مشتری و موسسات مالی می‌کند کشف این نمونه‌ها از اهمیت بالایی برخوردار است. لازم است تا علاوه بر شناسایی تک‌تک این موارد، نمونه‌ها را بر اساس میزان و حجم ضرری که متوجه موسسه مالی می‌کند، وزن دهی کرد. در نتیجه نمی‌توان سیستمی داشت که در مقابل این خطا بی‌تفاوت بوده و همه‌ی نمونه‌ها را به یک چشم نگاه کند. برای مثال در صورتی که در کل ۱ میلیون تراکنش[7] یک موسسه مالی در بازه مشخص، که مقدار سر جمع آنها به یک میلیارد تومان می‌رسد، تنها ۲ تراکنش ناهنجار وجود داشته باشد (که این میزان برابر تنها ۰.۰۰۰۲

[1] Electronic Access Point
[2] Point Of Sale (POS)
[3] Handheld
[4] Fraud Detection
[5] Skew
[6] Dataset
[7] Transaction

درصد از کل تراکنش‌ها در بازه مشخص است) که اگر فرض کنیم مبلغ این ۲ تراکنش در حدود ۱۰۰ میلیون تومان باشد، آنگاه از منظر حجم مالی، این مقدار ناهنجاری در حدود ۱۰ درصد از کل تراکنش های مالی آن موسسه در بازه زمانی مشخص را شامل می‌شود که مقداری قابل توجه بوده و لازم است تا این گونه تراکنش ها هر-چه سریع‌تر کشف و در صورت امکان از وقوع آن‌ها جلوگیری شود. برای غلبه بر مشکل نسبت کم نمونه‌های ناهنجار به کل نمونه‌ها از روش‌های متفاوتی استفاده می‌شود که یکی از آنها عبارت است از تغییر توزیع نمونه‌های ناهنجار در کل نمونه‌های موجود در مجموعه داده. بدین معنی که با تکثیر نمونه‌های ناهنجار در کل جمعیت، تعداد و سهم این نوع نمونه‌ها در کل جمعیت بالاتر برده شود تا به مقدار معقولی رسیده و بتوان از این طریق بر مشکل ناچیز بودن خطا غلبه کرد. توجه به این نکته ضروری است که انجام این عمل منجر به اضافه‌شدن دانش جدید به مجموعه‌داده نمی‌شود بلکه تنها به تصحیح توزیع نمونه ها کمک می‌کند[4] . در این تحقیق برای حل مشکل کمبود داده‌های ناهنجار از معیار میزان ضرر تحمیلی برای تشخیص میزان کارکرد صحیح مدل در تعیین وضعیت نهایی نمونه استفاده شده است. در مثال بالا میزان ضرر تحمیلی برای مشتری برابر ۱۰ درصد خواهد بود.

استخراج الگوهای ناهنجار و کلاهبرداری از داده‌ها چالش‌هایی به همراه دارد که تعدادی از آن‌ها به شرح زیر هستند[2] :

۱- معمولا مرز بین رفتار ناهنجار و رفتار نرمال به درستی و شفافیت مشخص نیست. به همین دلیل رفتارهایی که به مرز نزدیک هستند می‌توانند جزو هر دو دسته ناهنجار و نرمال قرارگیرند.

۲- در بسیاری از حوزه ها رفتار نرمال در طی زمان تغییر می‌کند و با استفاده از رفتار نرمال کنونی نمی‌توان در مورد رفتارهای آینده تصمیم‌گیری‌کرد. برای مثال اعمال محدودیت‌های جدید از سوی مراکز قانون‌گذاری و مراکز تصمیم‌گیری در مورد نحوه تعامل موسسات مالی با مشتریان می‌تواند بر روی رفتار و الگوی خرج‌کردن مشتریان بسیار تاثیرگذار باشد.

۳- اجرای موفق یک روش در یک حوزه‌ی مشخص تضمین نمی‌کند که اجرای آن روش در حوزه های دیگر نیز موفقیت‌آمیز باشد. این امر از آنجا ناشی می‌شود که پیچیدگی‌های فنی و مفهومی موجود در هر حوزه به طور کلی متفاوت از سایر حوزه‌هاست. برای مثال نمی‌توان اطمینان داشت که اعمال روش‌های استخراج الگوی موفق در کلاهبرداری‌های شرکت‌های مخابراتی همان نتیجه و موفقیتی

را درپی داشته باشد که اعمال آنها در بانک‌ها به همراه دارد این امر به دلیل متفاوت بودن معیارها و حدود تصمیم‌گیری در دو حوزه متفاوت است.

۴- معمولا داده‌ها دارای مقداری بی‌نظمی[1] هستند که این امر باعث مشکل‌تر شدن تصمیم‌گیری و جداسازی نمونه‌های نرمال از نمونه‌های ناهنجار می‌شود. وجود بی‌نظمی در بسیاری از موارد منجر به ایجاد خطاهای محسوس در تصمیم‌گیری می‌شود. در بسیاری موارد عدم حذف آن، موجب تشکیل قوانین و محدوده‌های نامناسب در سیستم‌های استخراج الگو می‌شوند و به این ترتیب سیستم دچار ضعف در تصمیم‌گیری شده و ممکن است الگوهای نرمال را ناهنجار تشخیص دهد و بالعکس.

۵- تصمیم‌گیری در مورد وضعیت نمونه‌ی جدیدی که به سیستم کشف کلاهبرداری وارد می‌شود باید در کمترین زمان ممکن انجام شود. اهمیت این موضوع زمانی مشخص می‌شود که در سیستم‌های کشف کلاهبرداری‌های کارت‌های اعتباری و سیستم‌های مالی، لازم است تا بلافاصله پس از وقوع جرم اقدامات لازم - برای جلوگیری از ضرررسانی بیشتر عمل مجرمانه به مشتری و موسسه‌ی مالی[2]- به سرعت انجام‌پذیرند [6].

۶- سیستم‌های کشف کلاهبرداری در هر لحظه با تعداد زیادی درخواست روبرو هستند و لازم است با به‌ازای هریک از درخواست‌های ورودی به سیستم به‌صورت آنی[3] اقدام به پاسخگویی مناسب کنند [6].

---

[1] Noise
[2] Financial Institute
[3] Real Time

۳.۱ تعاریف

۱.۳.۱ انواع ناهنجاری

انواع رایج ناهنجاری به دو دسته‌ی ناهنجاری‌های نقطه‌ای[1] و ناهنجاری‌های محتوایی[2] تقسیم‌بندی می‌شوند[1] . در ادامه به‌شرح هریک از این موارد پرداخته‌خواهدشد.

۱. ناهنجاری‌های نقطه ای

اگر نمونه‌ای به تنهایی در مقابل سایر داده‌ها ناهنجار تشخیص داده شود این ناهنجاری را نقطه‌ای گویند [1] [2]. در شکل ۱.۱ نقاط o_1 و o_2 نقاطی از نوع ناهنجاری نقطه‌ای هستند. این گونه ناهنجاری‌ها معمولا به شکل نویز در مجموعه‌داده دیده می‌شوند.

۲. ناهنجاری محتوایی

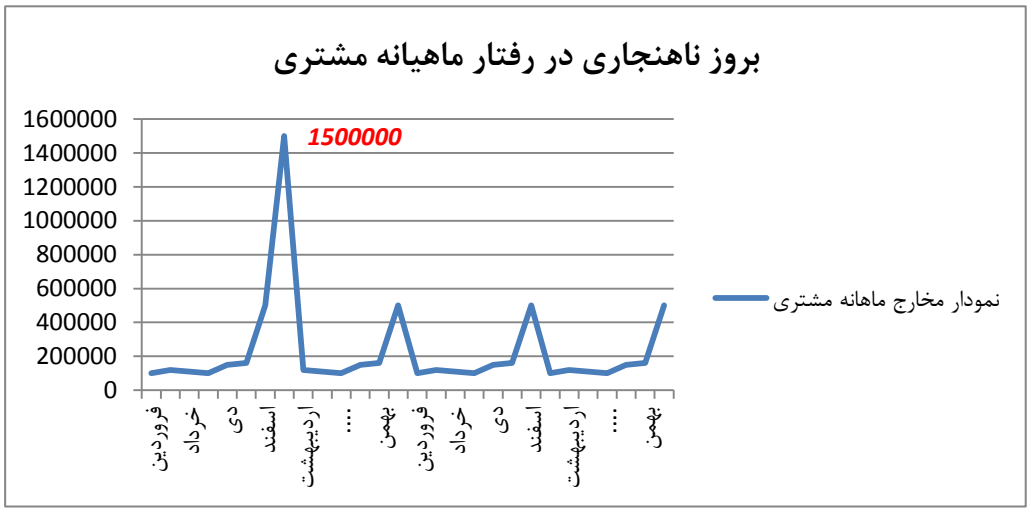
اگر نمونه بر اساس شرایط زمانی یا مکانی ناهنجار تشخیص داده شود، به این نمونه یک ناهنجاری محتوایی گفته می‌شود [2] [1] چون ممکن است که آن نمونه در شرایط زمانی و مکانی متفاوت، ناهنجاری تشخیص داده نشود.

در این حالت صفات هر نمونه به صفات محتوایی و صفات رفتاری تقسیم می‌شوند:

- **صفات محتوایی:** این صفات برای تعیین مکان یا زمان رخداد نمونه به کارمی‌روند [1]. برای مثال طول و عرض جغرافیایی، یک صفت محتوایی هستند که از کنار هم قراردادن آنها مکان وقوع نمونه بدست می‌آید. همچنین صفت زمان نیز یک صفت محتوایی است.

[1] Point Anomalies
[2] Contextual Anomalies

- **صفات رفتاری:** صفات رفتاری، صفاتی هستند که در موقعیت مشخص شده توسط صفات محتوایی مقدار می‌گیرند [2] [1]. برای مثال میزان موجودی حساب مشتری در تاریخ مشخص یک صفت رفتاری است. در حوزه کارت اعتباری، زمان خرید، یک صفت محتوایی و مبلغ خرید یک صفت رفتاری است. اگر فرض کنیم که یک مشتری هر هفته صد هزار تومان خرج می‌کند و در هفته منتهی به عید نوروز این مقدار به پانصد هزار تومان می‌رسد، در این‌صورت یک خرید یک میلیون و پانصد هزار تومانی در ماه‌های میانی سال *می‌تواند* به عنوان یک ناهنجاری محسوب شود چون با الگوهای ثابت خرید مشتری سازگاری ندارد.



شکل ۲.۱ نمونه‌ای از بروز ناهنجاری در رفتار مشتریان در ماه فروردین

۲.۳.۱ ناهنجاری های تجمعی[1]

اگر مجموعه ای از داده‌های مربوط به‌هم، نسبت به کل داده‌ها، دارای ناهنجاری باشند آنگاه این ناهنجاری، تجمعی خوانده می‌شود [2] [1]. در این نوع ناهنجاری، هر کدام از نمونه‌ها به تنهایی ناهنجاری

[1] Collective Anomalies

محسوب نمی‌شوند، اما از کنار هم قرار دادن آنها یک رفتار ناهنجار بدست می‌آید. بر این اساس نقاط موجود در محدوده O_3 در شکل ۱.۱ یک ناهنجاری تجمعی را تشکیل می‌دهند.

این نوع ناهنجاری‌ها معمولا به خاطر تغییر شرایط و قوانین اجرایی و همچنین به‌خاطر اعمال برخی اعمال خرابکارانه‌ی سازمان‌یافته رخ می‌دهند. برای مثال در حوزه‌ی بانک‌ها و موسسات مالی، در صورتیکه عده‌ای از مشتریان اقدام به پولشویی جهت پاکسازی مسیر ورود پول به حساب‌هایشان کنند، آنگاه در برخورد با حساب‌های این مشتریان شاهد ناهنجاری‌هایی خواهیم بود که این ناهنجاری‌ها زمانی خود را نشان می‌دهند که تراکنش‌های مالی این مشتریان با تراکنش‌های سایر مشتریان (مشتریان عادی) مقایسه شود. اما درصورتیکه تراکنش‌های آن‌ها با سایر مشتریانی که اقدام به پولشویی کرده‌اند مقایسه شود شاهد هیچ‌گونه ناهنجاری به‌دلیل یکسان بودن نوع رفتار آن‌ها نخواهیم بود.