تشخیص کاربران متقلب همدست در شبکه اجتماعی حراجی

مهیلا دادفرنیا[۱]، دانشجوی دکترا؛ فضل‌الله ادیب‌نیا[۲]، استادیار

۱- گروه مهندسی کامپیوتر - دانشگاه یزد - یزد - ایران - m-dadfarnia@stu.yazd.ac.ir
۲- گروه مهندسی کامپیوتر - دانشگاه یزد - یزد - ایران – fadib@yazd.ac.ir

چکیده: در طی سال‌های گذشته حراجی‌های برخط مورد توجه زیادی قرار گرفته است که با توجه به بعد مالی حراجی، منجر به افزایش تقلب در حراجی‌های برخط نیز شده است. به عنوان مثال، تقلب در حراج‌های اینترنتی از جمله شرکت eBay که یک شرکت مشهور در حراج‌های برخط می‌باشد، افزایش یافته است. به همین دلیل، پژوهش‌های مختلفی با رویکرد تشخیص خریداران و فروشندگان متقلب در سراسر دنیا انجام‌شده است. در حال حاضر در ایران، حراجی‌ها بیشتر در سازمان‌های دولتی بزرگ انجام می‌شود که قطعاً در صورت تقلب ضررهای مالی بزرگی را خواهد داشت. یکی از انواع تقلب، تقلب با روش همکاری و تبانی کاربران متقلب در حراجی می‌باشد که این نوع تقلب در صورت وقوع بسیار خطرناک می‌باشد و ضررهای مالی تاسف‌باری را خواهد داشت. در این مقاله الگوریتمی پیشنهاد می‌گردد که ابتدا ویژگی‌های مؤثر در یافتن افراد صادق را برای هر کاربر حراجی استخراج نماید و سپس با استفاده از طبقه‌بندی تک دسته‌بند OCSVM، برای هرکاربر یک نمره ناهنجاری را حساب نماید. در نهایت این الگوریتم، ارتباطات بین کاربران حراجی را با روش مارکف تصادفی مدل‌سازی می‌نماید تا باورهای کاربران را در مورد کاربران همسایه در گراف انتشار دهد و با این روش نمرات ناهنجاری را بازنگری نماید. نتایج این الگوریتم نشان می‌دهد که این تکنیک در هر سه نوع مختلف تقلب‌های اینترنتی می‌تواند کاربران متقلب همکار را با کارآیی بالا تشخیص دهد و در زمان سریع‌تر نسبت به الگوریتم‌های قبلی همگرا شده و پاسخ می‌دهد.

واژه‌های کلیدی: تشخیص تقلب همکارانه، پیشنهاد در حراجی، مدل مارکف تصادفی، انتشار باور، طبقه‌بندی تک دسته‌بند.

Detecting Collusive Fraud in Social Network of Online Auction

Mahila Dadfarnia[1], PhD student; Fazlollah Adibnia[2], Assistant Professor

1- Department of Computer Engineering, Yazd University, Yazd, Iran, Email: m-dadfarnia@stu.yazd.ac.ir
2- Department of Computer Engineering, Yazd University, Yazd, Iran, Email: fadib@yazd.ac.ir

Abstract: During last years, online auction has attracted many researchers. However, growing popularity in online auctions result in increasing fraud in online auctions. For instance, fraud increased in eBay, as a popular company in online auction. Therefore, many researches have done for detecting fraudulent buyers and sellers. One of the fraud type in online auctions is collusive auction fraud, in which multiple seller and bidders collude with each other. This kind of fraud is dangerous and caused catastrophic financial losses. Therefore, many techniques proposed to deal with this kind of fraud in online auctions. In this paper, we propose a novel detection technique in online auctions that use one-class to calculate an anomaly score for each unlabeled user. Then it models the users' interactions in the auctions as a pairwise markov random field (MRF). Next, our technique applies belief propagation to the MRF to revise anomaly scores. The results of our experiments show that our proposed technique is able to detect different types of collusive auction frauds within a reasonable detection time.

Keywords: Detecting Collusive Fraud, Bidding, Markov Random Field, Belief Propagation, One Class Classification

تاریخ ارسال مقاله: ۱۳۹۶/۰۹/۲۳
تاریخ اصلاح مقاله: ۱۳۹۶/۱۲/۲۰
تاریخ پذیرش مقاله: ۱۳۹۷/۰۳/۱۴
نام نویسنده مسئول: فضل‌الله ادیب‌نیا
نشانی نویسنده مسئول: ایران - یزد - صفاییه - دانشگاه یزد - گروه مهندسی کامپیوتر.

۱- مقدمه

در دنیای سایبری امروز، افزایش تعداد برنامه‌های تحت وب و پیچیدگی تهاجم روی شبکه‌های کامپیوتری باعث شده است که مسئله محافظت و مراقبت یکی از مسائل کلیدی در بحث امنیت کامپیوتر شود و جرائم رایانه‌ای به یکی از مباحث جدی و مهم تبدیل گردد. به همین دلیل، شناسایی حملات ممکن در انواع ارتباطات کاربران در زمینه‌های مختلف اجتماعی به منظور مقابله ضروری می‌باشد[۱]. از آنجا که یک شبکه اجتماعی از صدها فرد تشکیل شده که خود یک شبکه بزرگی را تشکیل می‌دهند، ارتباط کاربران در کنار تکنیک‌های فنی باعث چالش‌هایی در زمینه امنیت شده است. یکی از این شبکه‌های اجتماعی، شبکه حراجی می‌باشد که با توجه به منافع مالی زیادی که در مزایده‌های برخط وجود دارد، توجه کاربران کلاهبردار اینترنتی را به خود جلب کرده است. به طورکلی کاربران شبکه‌های حراجی به دو دسته تقسیم می‌شوند: فروشندگان و پیشنهاددهندگان. فروشنده کسی است که یک حراجی را برای فروش کالایی راه‌اندازی می‌نماید و پیشنهاددهنده کسی است که پیشنهاد خرید اقلام مورد نظرش را در آن حراجی می‌دهد و این دو کاربر در تعامل با یکدیگر ماهیت حراجی را تشکیل می‌دهند.

چندین نوع حراجی برخط وجود دارد. در حراجی انگلیسی که معمول‌ترین حراجی می‌باشد، حراجی با یک قیمت پیشنهادی از طرف فروشنده شروع می‌شود و به‌تدریج این قیمت از طرف پیشنهاددهندگان افزایش پیدا می‌کند[۲]. بنابراین، در زمان پایان یافتن حراجی بالاترین قیمت توسط برنده حراجی پیشنهاد داده شده است [۳].

امروزه تقلب در حراج‌های اینترنتی، یکی از حملات مطرح جرائم سایبری به حساب می‌آید که رشد سریعی داشته است. حراجی اینترنتی یک تجارت بزرگ برخط است که حتی قراردادهای دولتی نیز توسط حراجی بسته می‌شود و بسیاری از معاملات اقتصادی از طریق حراجی انجام می‌گیرد. به عنوان مثال، شرکت eBay که یک شرکت پیشرو در حراج‌های اینترنتی است، درآمد ۸/۹۷ بیلیون دلاری در سال ۲۰۰۶ داشته است و در حال حاضر دارای یک جامعه ۲۲۱ میلیون نفری از کاربران فعال در سراسر جهان است[۴, ۵]. شرکت‌های مختلف نیز در این حراج‌ها مشارکت می‌کنند و تقلب در آن می‌تواند تأثیر قابل توجهی در بازار میلیارد دلاری در سراسر جهان بگذارد[۶]. گزارش سالانه مرکز شکایات اینترنتی نشان می‌دهد که تقلب در حراج‌های اینترنتی برخط به عنوان یکی از جرائم جدی اینترنتی در سال‌های اخیر مطرح گردیده است[۷]. با وجودی که تنها تعداد کمی از قربانیان، جرائم مربوطه را اعلام می‌نمایند، گزارش FBI نشان می‌دهد که در سال ۲۰۱۵ ضررهای مالی از طریق حراجی برخط از ۱۸/۹ میلیون بیشتر شده است[۸]. بنابراین همواره یک کاربر که می‌خواهد یک کالا را در حراجی خرید کند، در مورد امنیت حراجی نگرانی دارد.

متاسفانه، با توجه به اینکه تقلب در حراجی بسیار پیچیده و مرموزانه است، تشخیص رفتارهای متقلبانه مشکل می‌باشد [۹، ۱۰، ۱۱]. روش های تقلب در حراجی انواع مختلفی دارد که در مرجع [۱۲، ۱۳] آمده است. در حراجی هر کاربری می‌تواند به صورت منحصر به فرد تقلب کند؛ به عنوان مثال، با توجه به اینکه، ایجاد چندین هویت از یک کاربر ساده می‌باشد، یک فروشنده می‌تواند چندین حساب کاربری متقلبانه ایجاد نماید و قیمت‌های پیشنهادی را با پیشنهادهای خودش دستکاری نماید. علاوه بر این، امکان تقلب با همکاری کاربران متقلب دیگر در حراجی می‌باشد. به عنوان مثال در روش شیلینگ رقابتی رفتار پیشنهاد مصنوعی توسط پیشنهاد دهنده صورت می‌گیرد که بدون قصد خرید کالا می‌باشد و فقط به منظور تحریک دیگر شرکت کنندگان می‌باشد که در نتیجه باعث بالا رفتن قیمت کالا می‌شود. در این روش بعد از مدتی پیشنهاددهنده شیلینگ، قیمت بالاتری پیشنهاد می‌دهد و امیدوار است که دیگر پیشنهاددهندگان مشروع رقم بالاتری نیز پیشنهاد کنند. هدف از این کار این است که برنده قانونی، مبلغ بیشتر پرداخت کند که این کار باعث بالا رفتن سود فروشنده می‌شود. در واقع در این روش، یک حراجی متقلبانه داریم که فروشنده با هدف تقلب آن را ایجاد نموده و با همکاری تعدادی خریدار متقلب می‌خواهد قیمت کالا را بالا ببرد تا خریدار قیمت بیشتری برای آن کالا پرداخت نماید. این نوع تقلب چندان معمول نیست، اما بسیار خطرناک می‌باشد؛ زیرا تشخیص آن مشکل‌تر می‌باشد و اغلب باعث ضرر مالی زیادی می‌گردد.

با وجودی که در دنیا به امنیت حراجی‌ها اهمیت زیادی داده شده است، در ایران تحقیقات زیادی در این زمینه انجام نشده است. در این مقاله روشی پیشنهاد می‌گردد که بتوانیم در زمان منطقی و سریع رفتارهای همکارانه متقلب‌ها را تشخیص دهیم. یکی از روش‌ها برای تشخیص، طبقه‌بندی می‌باشد. روش‌های مختلفی برای طبقه‌بندی وجود دارد که به عنوان مثال در [۱۴] از روش طبقه‌بندی دو کلاسی و در [۱۵] از روش طبقه‌بندی چند کلاسی استفاده شده است. در این مقاله از روش تک کلاسی استفاده می‌شود که با توجه به اطلاعات یک کلاس عمل طبقه‌بندی انجام می‌پذیرد. در این روش، ابتدا یکسری ویژگی را برای کاربرانی که برچسب ندارند، بر مبنای رفتار آنها در حراجی تعریف می‌نماییم و سپس یک عدد ناهنجاری با استفاده از طبقه‌بندی تک کلاسی برای آن کاربران محاسبه می‌نماییم. در ادامه، یک متغیر تصادفی دودویی را به هر کاربر نسبت می‌دهیم و ارتباطات بین کاربران مختلف را با یک گراف در روش تصادفی مارکف (MRF) مدل سازی می‌نماییم. درنهایت روش انتشار باور را بر روی این گراف انجام می‌دهیم تا نمرات ناهنجاری به دست آمده را بازبینی نموده و افراد متقلب همکار را تشخیص دهیم. با توجه به اینکه اغلب کاربرانی که با افراد متقلب در خرید و فروش شرکت می‌نمایند، خود نیز متقلب می‌باشند، این روش درواقع با در نظر گرفتن همسایه‌ها در گراف به یافتن افراد همکار متقلب کمک می‌نماید. در ادامه این روش را به صورت جزئی‌تر بررسی می‌نماییم.

این مقاله در ادامه به صورت زیر سازماندهی شده است:

در ادامه در بخش ۲، به کارهای مرتبط پرداخته می‌شود. با توجه به اینکه مدل تصادفی مارکف و انتشار باور در این مقاله استفاده شده است،

در بخش ۳ به پیش‌زمینه این مباحث می‌پردازیم. در بخش ۴، مدل پیشنهادی برای شبکه حراجی برخط تعریف می‌نماییم و سپس در بخش ۵، به الگوریتم پیشنهادی برای تشخیص افراد متقلب در حراجی می‌پردازیم که از سه گام تشکیل شده و آن را توضیح می‌دهیم. سپس در بخش ۶ به آزمایشات پرداخته می‌شود و در بخش ۷ نیز به نتیجه گیری پرداخته می‌شود.

۲- کارهای مرتبط

با توجه به افزایش تقلب در حراجی‌ها و ضرورت شناسایی این افراد، همواره تشخیص افراد متقلب در حراجی‌ها یکی از مسائل مورد علاقه محققان بوده است. به عنوان مثال، تعداد زیادی از موارد مورد تحقیق از ویژگی‌های پروفایل کاربران استفاده نمودند تا کاربران متقلب را تشخیص دهند[۱۶، ۱۷، ۱۸]. اما این روش‌ها روابط بین کاربران را درنظر نگرفته و به همین دلیل برای یافتن کاربران متقلب همکار مفید نیستند.

علاوه بر این روش، Chau و سایر همکارانش روش 2LFS را پیشنهاد دادند که در دو سطح، کشف الگوی غیرنرمال در حراجی را انجام می‌دهد[۱۹]. این الگوریتم در سطح اول، از ویژگی‌های سطح کاربر (اطلاعاتی مانند سن کاربر، تعداد و قیمت کالاهای فروخته و خریداری شده، زمان‌های انجام تراکنش و غیره) استفاده می‌نماید و در سطح دوم از ویژگی‌های سطح شبکه ارتباطی بین کاربران استفاده می‌نماید. سپس این کاربران و روابط بین آنها را با روش تصادفی مارکف مدل می‌نماید. در نهایت هدف این است که وضعیت هر گره را با سه برچسب <مورد اعتماد، متقلب، همدست متقلب[۱]> مشخص نماید. در واقع با توجه به مدل تصادفی مارکف این وضعیت‌ها، مقدار مشاهده شده مخفی برای هر کاربر می‌باشند که ممکن است خطا داشته باشند. ارتباط بین دو کاربر با یک لبه[۲] مشخص می‌شود که این گره‌های مخفی را به هم وصل می‌نماید. هر گره مخفی نیز با یک گره مشاهده شده مرتبط می‌باشد که در واقع مقدار اصلی آن را مشخص می‌نماید. اما این روش همه انواع حملات را درنظر نمی‌گیرد و فقط حملات تقلب در اعتبار را در نظر می‌گیرد و به همین دلیل در انواع دیگر حملات ستاره‌ای و حلقه‌ای دارای کارآیی پایین می‌باشد.

در ادامه Pandit و همکارانش یک سیستم تشخیص تقلب در حراج اینترنتی به نام NetProbe را طراحی و اجرا کردند که ایده اصلی آن یافتن خصوصیات یک کاربر خاص می‌باشد که با توجه به روابط با سایر کاربران تعریف می‌گردد[۲۰]. در این روش یک گراف، روابط بین کاربران حراجی را نشان می‌دهد و احتمال متقلب بودن یک کاربر با توجه به رفتار کاربران همسایه‌اش در گراف مشخص می‌شود. در این ایده، هر حراجی به عنوان یک گراف درنظر گرفته می‌شود که خریداران و فروشندگان با گره نشان داده می‌شوند و معاملات انجام شده توسط آنها نیز با یال نشان داده می‌شود. همچنین الگوریتم تصادفی مارکف و الگوریتم انتشار باور برای به دست آوردن الگوهای مشکوک و یافتن افراد متقلب احتمالی در معاملات ایجاد شده توسط متقلبان بکار می‌رود. در واقع در این روش تنها ارتباط بین دو کاربر فروشنده و خریدار در حراجی برای تشخیص افراد متقلب به کار می‌رود و سایر ویژگی‌ها نادیده گرفته شده است.

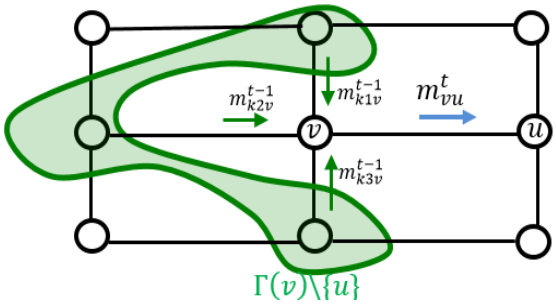
همچنین Tsang و همکارانش الگوریتم SPAN را پیشنهاد داده‌اندکه این الگوریتم ۲ مرحله دارد [۲۱]: در مرحله اول نمرات رفتار غیرعادی[۳] برای هر کاربر محاسبه می‌شود. محاسبه نمرات رفتار غیرعادی برای هر کاربر شامل ۳ مرحله انتخاب ویژگی‌ها، محاسبه نمرات LOF برای هر ویژگی و ترکیب مجموعه ویژگی‌ها برای هر کاربر می‌باشد. اما این روش همگرایی پایینی دارد و سرعت اجرای این روش بسیار کند می‌باشد.

۳- مباحث مرتبط پیش زمینه

در این بخش، روش تصادفی مارکف و روش انتشار باور را توضیح خواهیم داد که در ادامه از آنها استفاده می‌نماییم.

۳-۱- میدان تصادفی مارکف (MRF[۴])

مدل مارکف یکی از مدل‌های آماری مطرح و شناخته شده می‌باشد و می‌تواند رفتارهای پیچیده را مدل نماید. میدان‌های تصادفی مارکف نوعی از مدل‌های گرافیکی هستند که ارتباط بین متغیرهای تصادفی را به کمک یک گراف ساده بدون جهت مدل می‌کنند[۲۲]. به همین منظور، برای مدل نمودن رفتارهای کاربر و تخمین عملیات کاربر از این مدل استفاده می‌گردد. این مدل دارای ویژگی می‌باشد که یک کاربر از همه گره‌های دیگر به غیر از همسایه خود، می‌تواند مستقل شرطی باشد. به صورت فرمال، MRF یک گراف غیرجهت دار G=(U,E) می‌باشد که U مجموعه کاربران می‌باشد و E مجموعه یال‌ها می‌باشد که هر یال نمایش دهنده ارتباط بین کاربران می‌باشد. دو گره $u,v\epsilon U$ را همسایه تعریف می‌کنیم، اگر توسط یک یال $(u,v)\epsilon E$ به هم مرتبط شده باشند.



شکل۱- انتشار باور

اگر $X=(x_u)_{u\in U}$ مجموعه‌ای از متغیرهای تصادفی مرتبط با G باشد، برای هر گره $u\in U$ مدل MRF بر روی G با رابطه ۱ تعریف می‌شود:

$$P(X=x)=P(X_u=x) \qquad (۱)$$

در میدان‌های تصادفی مارکف، به هر گره u مقداری به نام پتانسیل گره۵ φ_u نسبت داده می‌شود که باور پیشین (حالت قابل‌مشاهده) آن گره را

نمایش می‌دهد. همچنین، به هر یال (u, v) مقداری به نام پتانسیل یال۶ ψ_{vu} نسبت داده می‌شود که وابستگی میان u و v را نمایش می‌دهد. توزیع احتمال توأم متغیرهای تصادفی با استفاده از میدان تصادفی مارکف به‌صورت رابطه ۲ محاسبه می‌شود:

$$P(X = x) = \frac{1}{Z}\prod_{u\in U}\varphi_u(X_u)\prod_{(v,u)\in E}\psi_{vu}(X_v, X_u) \qquad (۲)$$

$X = (x_u)_{u\in U}$ مجموعه متغیرهای تصادفی در X و متغیر z نیز ثابت نرمال‌سازی می‌باشد.

۳-۲- انتشار باور (BP[۷])

انتشار باور، یک الگوریتم برای انتشار پیغام در مدل‌های گرافیکی مثل شبکه‌های بیزین و میدان تصادفی مارکف می‌باشد[۲۳، ۲۴]. در این روش برای هر گره مشاهده نشده، باور آن گره محاسبه می‌گردد و این باور نشان می‌دهد که چقدر احتمال دارد یک گره مقدار خاصی را داشته باشد. یک گره مقدارش را با توجه به پیغام‌هایی که از گره‌های همسایه می‌گیرد، دریافت می‌نماید. به طور خلاصه در این روش، در یک فرآیند تکراری هر گره پیغام‌ها را به همسایگانش می‌فرستد و از پیغام‌های ورودی نیز برای به‌روزسانی باورها استفاده می‌نماید. اگر $m_{vu}^t(X_u)$ را پیغامی درنظر بگیریم که از گره v به گره u در زمان t ارسال می‌شود، به صورت زیر با رابطه ۳ می‌توان آن را محاسبه نمود:

$$m_{vu}^t(X_u) = \sum_{X_v}\varphi_v^t(X_v)\psi_{vu}(X_v, X_u)\prod_{k\in\Gamma(v)\setminus\{u\}} m_{kv}^{t-1}(X_v) \qquad (۳)$$

در این رابطه $\Gamma(v)$ مجموعه همسایگان گره v است و پیغام $m_{vu}^t(X_u)$ ، از ضرب همه پیام‌ها از همسایگان v به غیر از u محاسبه می‌شود. شکل۱ مثالی از انتشار باور را نشان می‌دهد که $m_{vu}^t(X_u)$ پیغامی است که از گره v به گره u در زمان t ارسال می‌شود و با توجه به ضرب پیغام‌هایی که همسایه‌های v به غیر از u در زمان t-1 به v می‌فرستند (فلش سبز) بدست می‌آید.

سپس هر گره باور خود را با توجه به همه پیغام‌های دریافت شده از همسایگانش، به‌روزرسانی می‌نماید که به صورت رابطه ۴ محاسبه می‌شود:

$$b_u^t(x_u) = K\varphi_u\,(x_u)\prod_{v\in\Gamma(u)} m_{vu}^t(x_u) \qquad (۴)$$

در این رابطه K ثابت نرمال‌سازی است. $b_u^t(x_u)$ احتمال این است که متغیر تصادفی X_u مقدار x_u را داشته باشد. این روش زمانی باورهای صحیح را محاسبه می‌نماید که شامل حلقه نباشد و در غیر اینصورت همگرا نمی‌شود. در صورت وجود حلقه از الگوریتم کامل‌تر انتشار باور به نام LBP استفاده می‌شود که در این صورت همگرا می‌شود. الگوریتم LBP هنگامی‌خاتمه می‌یابد که تغییرات پیغام‌ها ناچیز شده یا تعداد تکرارها از یک حد آستانه مشخص عبور کند.

۴- مدل پیشنهادی برای شبکه حراجی برخط

هر شبکه حراجی از مجموعه کاربران فروشنده و پیشنهاد دهنده خریدار، مجموعه حراجی، مجموعه پیشنهادها و قوانین حراجی تشکیل شده است. در واقع روابط بین کاربران با توجه به قوانین حراجی و پیشنهادهای پیشنهاددهندگان در حراجی‌ها تعریف می‌گردد و هرکاربر می‌تواند به صورت همزمان در یک حراجی یا در حراجی‌های مختلف، هم فروشنده و هم خریدار باشد. فروشنده حراجی، یک حراجی را ایجاد می‌نماید و کالایی را برای فروش با قیمت اولیه پیشنهاد می‌دهد. پیشنهاددهندگان یا شرکت‌کنندگان در حراجی نیز کالای مورد نظر خود را با قیمت پیشنهادی برای خرید به فروشنده پیشنهاد می‌نمایند.

ما شبکه حراجی را به صورت $N = (U, A, B)$ تعریف می‌نماییم که U مجموعه کاربران، A مجموعه حراجی‌ها و B مجموعه پیشنهادها می‌باشند. برای ادامه مجموعه‌هایی را با کمک سه مجموعه اصلی بالا تعریف می‌نماییم. اگر در شبکه حراجی، حراجی $a \in A$ و کاربر $u \in U$ داشته باشیم، $A_u^{\rightarrow}$ را مجموعه حراجی‌هایی تعریف می‌کنیم که توسط کاربر u ایجاد شده و به عبارتی u فروشنده می‌باشد. همچنین B_a را به عنوان پیشنهادهای در حراجی a و U_a را مجموعه کاربران پیشنهاددهنده در حراجی a تعریف می‌نماییم. در واقع مجموعه حراجی‌هایی که توسط کاربران مجموعه U ایجاد می‌شوند، مجموعه کل حراجی‌ها می‌باشند، یعنی:

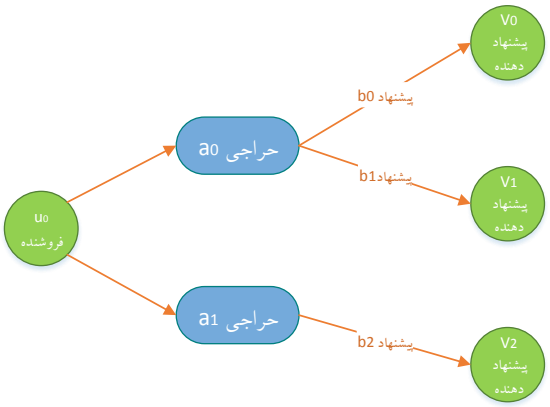
$$\bigcup_{u\in U} A_u^{\rightarrow} = A \qquad (۵)$$

علاوه بر این، $A_u^{\leftarrow}$ را مجموعه حراجی‌هایی تعریف می‌نماییم که کاربر u در آنها پیشنهاد داده است و $A_u^{\leftarrow+}$ مجموعه حراجی‌هایی که کاربر u در آن حراجی‌ها برنده شده است. $A_u^{\leftarrow-}$ را نیز مجموعه حراجی‌هایی تعریف می‌کنیم که کاربر u در آن‌ها بازنده شده است. بنابراین با توجه به اینکه در هر حراجی کاربر u یا برنده است یا بازنده، خواهیم داشت:

$$A_u^{\leftarrow} = A_u^{\leftarrow+} \cup A_u^{\leftarrow-} \qquad (۶)$$

مجموعه $B_{u,a}$، نیز به صورت مجموعه پیشنهادهایی که توسط u در حراجی a داده شده است تعریف می‌کنیم. علاوه بر این $U_u^{\leftarrow}$ مجموعه فروشندگان در حراجی‌هایی تعریف می‌کنیم که کاربر u در همان حراجی‌ها پیشنهاد داده است. همچنین $U_u^{\rightarrow}$ را مجموعه پیشنهاددهندگان در حراجی‌هایی تعریف می‌کنیم که کاربر u آن حراجی‌ها را ایجاد نموده است.

به عنوان مثال، فرض کنیم مجموعه پیشنهادها $B = \{b_0, b_1, b_2\}$ و حراجی‌های $A = \{a_0, a_1\}$ و کاربران $U = \{u_0, v_0, v_1, v_2\}$ را داشته باشیم و روابط شکل ۲ را بین این مجموعه‌ها داشته باشیم، یعنی کاربر فروشنده u_0 حراجی a_0, a_1 را ایجاد نماید؛ در حراجی a_0 ، کاربر v_0 پیشنهاد b_0 و کاربر v_1 پیشنهاد b_1 را مطرح نماید و در حراجی a1 ، کاربر v_2 پیشنهاد b_2 را مطرح نماید.



شکل۲- مثالی از روابط بین حراجی‌ها و کاربران و پیشنهادها

در این صورت، روابط ۷ الی ۱۱ را برای مجموعه‌های $U_{v1}^{\leftarrow}, U_{a_0}, B_{v1,a0}, B_{a_0}, U_{u0}^{\rightarrow}$ خواهیم داشت:

$$B_{a_0} = \{b_0, b_1\} \quad (۷)$$

$$B_{v1,a0} = \{b_1\} \quad (۸)$$

$$U_{a_0} = \{v_1, v_0\} \quad (۹)$$

$$U_{u0}^{\rightarrow} = \{v_0, v_1, v_2\} \quad (۱۰)$$

$$U_{v1}^{\leftarrow} = \{u_0\} \quad (۱۱)$$

۵- الگوریتم پیشنهادی

در این بخش الگوریتم پیشنهادی توضیح داده می‌شود که روشی برای تشخیص افراد متقلب همکار در شبکه حراجی می‌باشد. این الگوریتم در سه گام انجام می‌شود: گام اول استخراج ویژگی‌ها در شبکه حراجی می‌باشد. گام دوم محاسبه نمره ناهنجاری و گام سوم بهبود نمرات ناهنجاری با استفاده از روش انتشار باور می‌باشد. شبه کد این الگوریتم در شکل ۳ آورده شده است. این گام‌ها در ادامه توضیح داده می‌شوند.

Algorithm 1 : Fraud Detection in Auction
Input: U, A, B
Output: belief Matrix b
01: initialize $A_u^{\rightarrow}, A_u^{\leftarrow}, A_u^{\leftarrow -}, U_a, U_u^{\leftarrow}, B_{u,a}, B_a$
02: for each user $u \in U$ calculate f_u
03: calculate decision function η using OCSVM algorithm
04: calculate belief of each node b_u using belief propagation algorithm
Return b

شکل۳ – شبه کد الگوریتم پیشنهادی

گام اول: استخراج ویژگی‌ها

در این مرحله، ابتدا اطلاعات مجموعه‌ای از کاربران عادی به‌عنوان ورودی دریافت می‌کنیم و سپس برای هر کاربر $u \in U$ با توجه به مجموعه حراجی‌هایی که به‌عنوان پیشنهاددهنده یا فروشنده در آن‌ها مشارکت داشته، یک بردار ویژگی $f_u = \{f_{u1}, f_{u2}, .., f_{u10}\}$ محاسبه می‌نماییم.

این ویژگی‌ها به دو دسته ویژگی‌های پیشنهاددهندگان و فروشندگان تقسیم می‌شوند:

- **ویژگی‌های پیشنهاددهندگان:**

برای هر کاربر u با توجه به حراجی‌هایی که در آن‌ها به‌عنوان پیشنهاددهنده مشارکت داشته است، پنج ویژگی زیر استخراج می‌شود:

ویژگی اول: تعداد کل حراجی‌هایی که کاربر u در آن‌ها مشارکت داشته است.

$$f_{u1} = |\overleftarrow{A}_u| \quad (۱۲)$$

ویژگی دوم: تعداد پیشنهاددهندگان متمایزی که با کاربر u در حراجی‌های یکسان مشارکت داشته‌اند.

$$f_{u2} = \left|\bigcup_{a \in \overleftarrow{A}_u} U_a \setminus \{u\}\right| \quad (۱۳)$$

ویژگی سوم: تعداد حراجی‌های با فروشندگان متفاوت که پیشنهاددهنده u در آنها شرکت کرده است.

$$f_{u3} = |U_u^{\leftarrow}| \quad (۱۴)$$

ویژگی چهارم: متوسط تعداد حراجی‌هایی که در آن‌ها پیشنهاد داده ولی بازنده شده است.

$$f_{u4} = \frac{|\overleftarrow{A}_u^-|}{|\overleftarrow{A}_u|} \quad (۱۵)$$

ویژگی پنجم: متوسط تعداد پیشنهادهای کاربر u در حراجی‌هایی که در آن‌ها بازنده شده است.

$$f_{u5} = \frac{1}{|\overleftarrow{A}_u^-|} \sum_{a \in \overleftarrow{A}_u^-} |B_{u,a}| \quad (۱۶)$$

- **ویژگی‌های فروشندگان:**

برای هر کاربر u با توجه به حراجی‌هایی که به‌عنوان فروشنده راه‌اندازی کرده است و رفتارهای کاربران فروشنده، پنج ویژگی زیر استخراج می‌شود:

ویژگی اول: تعداد کل حراجی‌هایی که کاربر u راه‌اندازی کرده است.

$$f_{u6} = |\vec{A}_u| \quad (۱۷)$$

ویژگی دوم: تعداد کل پیشنهاددهندگانی که در حراجی‌های کاربر فروشنده u مشارکت داشته‌اند.

$$f_{u7} = \sum_{a \in \vec{A}_u} |U_a| \quad (۱۸)$$

ویژگی سوم: متوسط تعداد پیشنهاددهندگان متمایز در حراجی‌هایی که کاربر u راه‌اندازی کرده است.

$$f_{u8} = \frac{|U_u^{\leftarrow}|}{|\vec{A}_u|} \quad (۱۹)$$

ویژگی چهارم: متوسط تعداد پیشنهاددهندگان بازنده در حراجی‌هایی که کاربر u راه‌اندازی کرده است.

$$f_{u9} = \frac{|\{v \in U | \vec{A}_u \cap \overleftarrow{A}_v^- \neq \emptyset\}|}{|\vec{A}_u|} \quad (۲۰)$$

ویژگی پنجم: متوسط تعداد پیشنهادها در حراجی‌هایی که کاربر u راه‌اندازی کرده است.

$$f_{u10} = \left| \bigcup_{a \in A_u} B_a \right| \quad (۲۱)$$

گام دوم: محاسبه نمرات ناهنجاری

در این مرحله، نمره ناهنجاری را برای هر کاربر $u \in U$ برمبنای رفتار فروشنده در هنگام ایجاد حراجی برای فروختن کالا و یا رفتار پیشنهاد دهنده را در هنگام پیشنهاد دادن قیمت برای خرید کالا محاسبه می نماییم. به این منظور، ابتدا یک مدل از رفتارهای کاربران می‌سازیم. چنانچه یک مجموعه از کاربران برچسب خورده صادق داشته باشیم، ابتدا بردار ویژگی f_u را برای آن دسته از کاربران محاسبه می‌نماییم و سپس از تابع تصمیم‌گیری $\eta: U \to \mathbb{R}$ برای یادگیری استفاده می‌نماییم. به این صورت که در واقع با توجه به بردار ویژگی f_u برای کاربران صادق، می‌توانیم کاربران دیگر را نیز برچسب‌گذاری نماییم. برای این منظور، الگوریتم طبقه‌بندی تک دسته‌بند OCSVM برای تابع تصمیم‌گیری استفاده می‌شود. خروجی این الگوریتم طبقه‌بندی برای کاربران مشکوک و احتمالاً مخرب مقدار مثبت بوده و برای افراد صادق، عدد منفی را می باشد. فرآیند یادگیری به شیوه‌ای است که با توجه به بردار ویژگی که در بخش قبل برای هر کاربر محاسبه گردید، افراد صادق برچسب‌خورده و غیر برچسب‌خورده را در مجموعه +L طبقه‌بندی می‌نماید.

گام سوم: تشخیص افراد متقلب

در این مرحله برای هر کاربر $u \in U$ یک متغیر تصادفی دودویی $X_u \in \{-1, +1\}$در نظر می‌گیریم که عدد 1+ نشان‌دهنده این است که کاربر u صادق می‌باشد و عدد 1- نشان‌دهنده متقلب بودن کاربر می‌باشد. با توجه به موارد مطرح شده در بخش ۳-۱ می‌دانیم که مدل تصادفی مارکف از اطلاعات قبلی برای تابع پتانسیل گره برای هر کاربر استفاده می‌نماید و از تابع پتانسیل لبه برای ارتباط همبستگی بین دوکاربر همسایه استفاده می‌نماید. چنانچه $\eta(u)$ نمره ناهنجاری باشد که برای کاربر u در مرحله قبل محاسبه نمودیم و $\beta > 0$ پارامتر ثابت نرمال‌سازی درنظر گرفته شود، θ_u را با توجه به رابطه ۲۲ می‌توان محاسبه نمود:

$$\theta_u = \begin{cases} 0 & \text{if } u \in L^+ , \\ \dfrac{1}{1 + \exp(-\eta(u)/\beta)} & \text{if } u \notin L^+ , \end{cases} \quad (۲۲)$$

که با توجه به آن، تابع پتانسیل گره را از رابطه ۲۳ و تابع پتانسیل لبه $(u, v) \in E$ را از رابطه ۲۴ محاسبه می‌نماییم:

$$\varphi_u^{L^+}(x_u) = \begin{cases} 1 - \theta_u & \text{if } x_u = +1 , \\ \theta_u & \text{if } x_u = -1 , \end{cases} \quad (۲۳)$$

و

$$\psi_{uv}(x_u, x_v) = \begin{cases} \vartheta_{uv} & \text{if } x_u x_v = +1 , \\ 1 - \vartheta_{uv} & \text{if } x_u x_v = -1 , \end{cases} \quad (۲۴)$$

در اینجا ϑ_{uv} را برابر با ۰/۶۵ گذاشتیم تا نشان دهیم که دو کاربر u, v ارتباط مثبت با یکدیگر دارند و احتمالاً دارای برچسب‌های یکسان می باشند. یعنی چنانچه یکی صادق باشد، به احتمال زیاد همسایه صادق دارد و در غیر اینصورت چنانچه یکی متقلب باشد، با همسایه‌های متقلب خود تبانی خواهد نمود.

سپس از روش انتشار باور استفاده می‌نماییم تا توزیع پسین ۸ برای هر متغیر تصادفی از کاربران برچسب نخورده محاسبه نماید. همان‌طور که قبلاً در بخش ۲-۳ اشاره نمودیم، این روش با توجه به ارتباط پیغام‌ها بین همسایه‌ها، باور هر گره را از همسایه خود به دست می‌آورد. پیغامی که از کاربر $v \in U$ به کاربر $u \in U$ در زمان t فرستاده می‌شود با رابطه ۲۵ به دست می‌آید:

$$m_{vu}^{(t)}(x_u) = \sum_{x_v} \varphi_v^{L^+}(x_v) \psi_{uv}(x_u, x_v) \prod_{w \in \Gamma_v \setminus \{u\}} m_{wv}^{(t-1)}(x_v) \quad (۲۵)$$

سپس باور هر کاربر برچسب نخورده $u \in U \setminus L^+$ را با رابطه ۲۶ محاسبه می‌نماییم و با توجه به خروجی این رابطه هر کاربر $u \in U \setminus L^+$ را با صادق و متقلب برچسب‌گذاری می‌نماییم.

$$b_u^{(t)}(x_u) = K \varphi_u^{L^+}(x_u) \prod_{v \in \Gamma_u} m_{vu}^{(t)}(x_u) . \quad (۲۶)$$

چنانچه σ را پارامتر حد آستانه تعریف شده توسط کاربر درنظر بگیریم، اگر در آخرین تکرار (زمان t_{max})، مقدار $b_u^{(t_{max})}(+1) > \sigma$ باشد، کاربر را صادق در نظر می‌گیریم و در غیر این صورت متقلب خواهد بود.

۶- نتایج و آزمایشات

در این قسمت به ارزیابی الگوریتم ارائه شده برای تشخیص کاربران متقلب پرداخته شده است. برای این منظور الگوریتم پیشنهادی را با الگوریتم‌های قبلی 2LFS [۱۹] و SPAN [۲۱] مقایسه می‌نماییم و از بانک اطلاعاتی در [۲۱] استفاده می‌نماییم. این بانک اطلاعاتی از ۶۰ شبکه حراجی تشکیل شده که شامل کاربران صادق و متقلبی می‌باشد که با هم تعامل می‌نمایند. بانک‌اطلاعاتی مذکور، سه نوع رفتار متقلبانه زیر را در خود دارد که در آن از هر کدام از این رفتارهای متقلبانه ۲۰ شبکه حراجی موجود می‌باشد:

تقلب در اعتبار[۹]: این تقلب زمانی رخ می‌دهد که تعدادی از پیشنهاددهندگان با یک فروشنده خاص تبانی می‌نماید تا اعتبار خود را بالا ببرد.

تقلب ستاره‌ای[۱۰]: این رفتار متقلبانه نیز در زمانی می‌باشد که چندین پیشنهاددهنده متقلب، پیشنهادهای خود را در حراجی می‌دهند که یک کاربر خاص راه‌اندازی نموده است.

تقلب حلقه‌ای[۱۱]: این رفتار متقلبانه زمانی می‌باشد که یک گروه از کاربران متقلب هم به عنوان پیشنهاددهنده و هم به عنوان فروشنده با هم تبانی می‌نمایند. جدول ۱ میانگین مشخصات این مجموعه داده‌ها را نشان می‌دهد.

جدول۱- میانگین مشخصات مجموعه داده‌ها

	Reputation Fraud	Collusive Star	Collusive Clique
$\|U\|$	۲۴۰۹۹	۲۳۹۱۰	**۲۳۹۱۵**
$\|A\|$	۲۰۱۹۵	۱۹۰۸۵	**۱۸۸۸۷**
$\|B\|$	۲۳۴۰۱۰	۲۵۱۰۱۵	**۲۳۸۹۶۰**
Bidders In each Auction	۲/۷	۲/۹	**۲/۹**
Bids In each Auction	۱۱/۶	۱۳/۲	**۱۲/۷**

به منظور مقایسه کارآیی روش پیشنهادی با روش‌های قبلی از دو معیار TPR[۱۲] و FPR[۱۳] استفاده می‌گردد. معیار TPR نشان می‌دهد که دقت تشخیص کلاس مثبت چه مقدار است و معیار FPR نرخ هشدار غلط را با توجه به دسته منفی بیان می‌کند. این معیارها از روابط ۲۷ و ۲۸ محاسبه می‌گردند:

$$TPR = \frac{TP}{TP + FN}\ , \qquad (۲۷)$$

$$FPR = \frac{FP}{FP + TN}\ , \qquad (۲۸)$$

در این معادله‌ها، متغیر TN بیانگر تعداد رکوردهایی است که کلاس واقعی آنها منفی بوده و الگوریتم دسته‌بندی نیز دسته آنها را به‌درستی منفی تشخیص داده است. متغیر TP بیانگر تعداد رکوردهایی است که دسته واقعی آنها مثبت بوده و الگوریتم دسته‌بندی نیز دسته آنها را به‌درستی مثبت تشخیص داده است. متغیر FP بیانگر تعداد رکوردهایی است که دسته واقعی آنها منفی بوده و الگوریتم دسته‌بندی دسته آنها را به اشتباه مثبت تشخیص داده است. متغیر FN نیز بیانگر تعداد رکوردهایی است که دسته واقعی آنها مثبت بوده و الگوریتم دسته‌بندی دسته آنها را به اشتباه منفی تشخیص داده است.

جهت ارزیابی دقیق‌تر آزمایش‌ها، درصد کمی (ده درصد) از کل داده‌ها را به عنوان داده‌های برچسب خورده آموزشی در نظر گرفتیم. با توجه به اینکه از روش تک دسته‌بند استفاده نمودیم، این داده‌های آموزشی از مجموعه داده‌های نرمال باید انتخاب شوند. برای آزمایش، ده بار، ده درصد از کل داده‌ها را به عنوان داده آموزشی در نظر گرفته و نتایج آن را بین ده بار آزمایش میانگین می‌گیریم.

در ادامه، پارامتر حد آستانه را تغییر دادیم و تأثیر آن را بر روی کارآیی محاسبه نمودیم. همانطور که از شکل ۴ مشخص می‌باشد، با افزایش حد آستانه هر دو مقدار TPR, FPR تمایل به سیر صعودی دارند. دلیل آن این است که با افزایش حد آستانه در واقع تعداد زیادی از کاربران متقلب به صورت صادق و تعداد زیادی از کاربران متقلب، صادق دسته‌بندی می شوند.

برای نتیجه‌گیری، پارامتر حد آستانه σ را برابر $\sigma = 0.40$ قرار می‌دهیم و پارامتر ثابت نرمال‌سازی را برابر $\beta = 0.50$ قرار می‌دهیم. جدول ۲ کارآیی روش CFD را در مقایسه با روش‌های دیگر با توجه به معیارهای TPR و FPR نشان می‌دهد. همانطور که مشخص می‌باشد روش CFD در دو بانک اطلاعاتی Collusive Star و Collusive Clique بهتر از روش‌های SPAN و 2LFS عمل می‌نماید و عمل تشخیص افراد صادق و متقلب را با دقت بیشتری انجام می‌دهد.

جدول۲- مقایسه کارآیی روش پیشنهادی با روش‌های دیگر

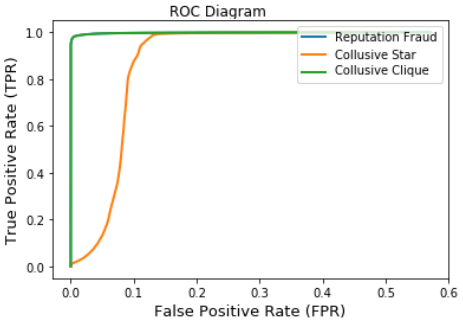
Technique	Reputation Fraud		Collusive Star		Collusive Clique	
	TPR	*FPR*	*TPR*	*FPR*	*TPR*	*FPR*
2LFS	۱۳/۳	۱/۰	۱۴/۴	۱/۰	۳۱/۲	۱/۰
SPAN	۹۸/۹	۱/۰	۴۷/۶	۱/۰	۹۴/۴	۱/۰
Proposed Method	۹۷/۸	۲/۱	۴۸/۷	۱/۳	۹۷/۲	۱/۸

شکل ۴ - مقادیر متفاوت حد آستانه (σ) در تأثیر کارآیی الگوریتم پیشنهادی

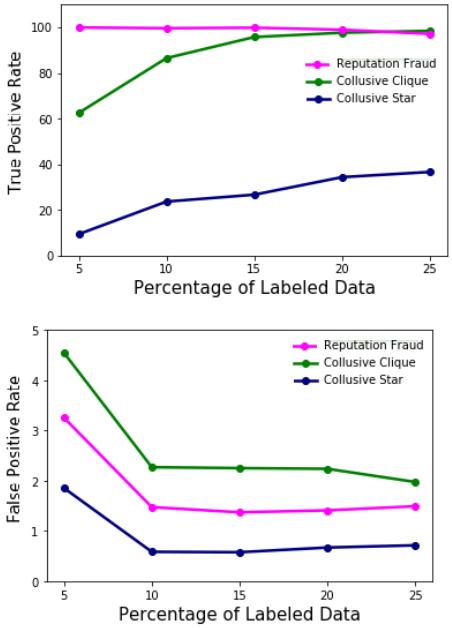
شکل ۵ کارآیی روش پیشنهادی را در سه مجموعه بانک داده‌ای ذکر شده، نشان می‌دهد. این آزمایش با در نظر گرفتن نرخ برچسب ۰/۱ درصد و نرخ حد آستانه ۰/۳ انجام شده است. همانطور که مشخص می‌باشد، نتایج تشخیص در مجموعه داده‌ای Collusive Clique و Reputation Fraud نزدیک به هم بوده و در مجموعه داده‌ای Collusive Star نرخ تشخیص پایین‌تری خواهیم داشت.

شکل ۶ کارآیی روش پیشنهادی را در سه مجموعه بانک داده‌ای ذکر شده با درنظر گرفتن درصد داده‌های برچسب خورده متفاوت نشان می‌دهد. همانطور که این شکل نشان می‌دهد، هرچه درصد داده‌های

برچسب خورده زیادتر می‌شود، مقدار TPR تمایل به سیر صعودی داشته و مقدار FPR کمتر خواهد شد و به این ترتیب درصد نتایج تشخیص بیشتر خواهد شد.



شکل ۵ – کارآیی الگوریتم پیشنهادی در سه مجموعه بانک داده‌ای



شکل ۶ – کارآیی الگوریتم پیشنهادی در سه مجموعه بانک داده‌ای با توجه به درصد داده‌های برچسب خورده

علاوه بر این روش پیشنهادی سریع‌تر همگرا می‌شود. همانطور که نتایج در جدول ۳ نشان می‌دهند، روش پیشنهادی CFD برای هر سه نوع تقلب، در کمتر از ۷ تکرار همگرا می‌شود، در صورتی که روش SPAN بعد از بیشتر از ۳۰ بار تکرار همگرا می‌شود. بنابراین هزینه زمانی صرف شده برای این روش بسیار کمتر از روش SPAN می‌باشد و این یک مزیت مهم نسبت به روش SPAN می‌باشد.

در ادامه کارآیی روش‌ها را بدون گام دوم یعنی انتشار نمرات ناهنجاری مقایسه نمودیم. همانطور که نتایج در جدول ۴ نشان می‌دهند، TPR در روش پیشنهادی بدون گام دوم (یعنی انتشار نمرات ناهنجاری) بسیار بهتر از SPAN می‌باشد. اما FPR در این روش بالا می‌باشد که مقایسه جدول ۲ و ۴ نشان می‌دهد که روش پیشنهادی با وجود گام سوم و استفاده از فرضیه همبستگی برچسب همسایه‌ها تأثیر مناسبی در نتیجه خروجی دارد و FPR را به مقدار کمی تقلیل می‌دهد. همانطور که مشاهده می‌شود FPR در جدول ۴ بسیار کمتر از جدول ۲ می‌باشد، یعنی با این روش تعداد افراد متقلبی که به اشتباه صادق تشخیص داده شده را کاهش دادیم.

جدول ۳– مقایسه تعداد حلقه تکرار LBP روش پیشنهادی CFD با روش‌های دیگر

Technique	Reputation Fraud	Collusive Star	Collusive Clique
SPAN	۴۰-۳۰	۴۰-۳۰	۹۰-۸۰
CFD	۶/۱	۶/۲	۶/۳

جدول ۴– مقایسه کارآیی روش پیشنهادی بدون گام دوم (محاسبه نمرات ناهنجاری) با روش SPAN

Technique	Reputation Fraud		Collusive Star		Collusive Clique	
	TPR	*FPR*	*TPR*	*FPR*	*TPR*	*FPR*
SPAN-AD	۹/۱	۱/۰	۲۰/۴	۱/۰	۷۵/۸	۱/۰
روش پیشنهادی بدون گام دوم	۹۵/۵	۶/۲	۷۲/۹	۷/۳	۹۸	۵/۸

۷– نتیجه

تبانی در تقلب زمانی رخ می‌دهد که چندین فرد متقلب پیشنهاددهنده و خریدار با یکدیگر همکاری کنند تا افراد دیگر از جمله پیشنهاددهنده‌ها را فریب دهند. در طول سال‌های قبل چندین روش برای تشخیص تبانی در تقلب پیشنهاد شده که اغلب آنها کارآیی لازم را نداشته و کند می باشند و یا نرخ تشخیص پایینی دارند. در این مقاله یک مدل ریاضی برای حراجی مطرح می‌گردد و سپس الگوریتمی پیشنهاد می‌گردد که ترکیب طبقه‌بندی تک دسته‌بند OCSVM و روش‌های تصادفی مارکف و انتشار باور می‌باشد. الگوریتم پیشنهادی شامل سه مرحله اصلی استخراج ویژگی، تشخیص و انتشار می‌باشد. در مرحله استخراج ویژگی، با توجه به رفتار کاربران متقلب با سایر افراد ویژگی‌های هر کاربر استخراج می‌گردد. در مرحله تشخیص، برای هر کاربر یک نمره ناهنجاری اولیه محاسبه می‌شود. در مرحله انتشار، نمرات ناهنجاری کاربران با استفاده از یک روش انتشار باور، بازبینی شده و کاربران مشکوک به تقلب شناسایی می‌شوند. کارآیی الگوریتم پیشنهادی با توجه به سه نوع رفتار همکاری متقلبانه بررسی شده است. نتایج نشان می دهند که نرخ FP و TP در بسیاری موارد بهتر از الگوریتم‌های قبلی عمل می‌نمایند و این روش سریع‌تر از روش‌های قبلی همگرا می‌گردد.

مراجع

[1] C. Bauner, "Mechanism choice and the buy-it-now auction: A structural model of competing buyers and sellers", International Journal of Industrial Organization. pp. 19-31, Elsevier 2015.

[2] F. Dong , S. M. Shatz and H. Xu, "Combating online in-auction fraud: Clues, techniques and challenges", *Computer Science Review* vol. 3, pp. 245–258, Elsevier 2009.
[3] F. M. Menezes and P. K. Monteiro, A*n Introduction to Auction Theory*, New York, NY, USA: Oxford University Press, 2005.
[4] MI. Melnik, "Confronting the Challenges of Asymmetry of Information and Competition: The Rise of eBay", InTrends and Innovations in Marketing Information Systems, pp. 293-307, IGI Global Book 2015.
[5] DataStax. *eBay Engages Customers with Personalized Recommendations*. Accessed on: Aug. 22, 2017. [Online]. Available:https://www. datastax.com/resources/casestudies/ebay
[6] M. M. Flax, *Economic Crimes*, San Clemente, CA, USA: LawTech Publishing Group, 2005.
[7] J. S. Thomas and F. A. Jose, "E-Auction Frauds - A Survey", vol. 61, pp. 41–45, IJCA 2013.
[8] Center, N. W. C. C., the Federal Bureau of Investigation, 2015. Ic3 2015 internet fraud crime report. https://pdf.ic3.gov//2015IC3Report.
[9] S. Ganguly and S. Sadaoui, "Classification of Imbalanced Auction Fraud Data", Canadian Conference on Artificial Intelligence, pp. 84-89, Springer 2017.
[10] CH. Yu, "A Fuzzy Genetic Approach for Optimization of Online Auction Fraud Detection", Frontier Computing , pp. 965-974, Springer 2016.
[11] DH. Chau and C. Faloutsos, "Fraud Detection Using Social Network Analysis", a Case Study, Encyclopedia of Social Network Analysis and Mining, pp. 547-552, Springer 2014.
[12] J. Li, KF. Tso and F. Liu, Profit earning and monetary loss bidding in online entertainment shopping: the impacts of bidding patterns and characteristics. Electronic Markets. pp. 77-90, 2017.
[13] D. H. Chau and C. Faloutsos, "Fraud detection in electronic auction", European Web Mining Forum Proceeding, pp. 87–97, 2005.
[۱۴] مرتضی خرّم کشکولی، مریم دهقانی، «تشخیص، شناسایی و جداسازی عیب توربین گاز پالایشگاه دوم پارس جنوبی با استفاده از روش‌های ترکیبی داده‌کاوی، k-means، تحلیل مؤلفه‌های اصلی (PCA) و ماشین بردار پشتیبان (SVM) »، مجله مهندسی برق دانشگاه تبریز، شماره (۲)۴۷، صفحه ۵۱۵-۵۰۱، سال ۱۳۹۶.
[۱۵] میلاد رفیعی، مهدی عباسی، محمد نصیری، «روشی کارا برای پیاده‌سازی موازی الگوریتم دسته بندی بسته درخت سلسله‌مراتبی بر روی واحد پردازش گرافیکی» .مجله مهندسی برق دانشگاه تبریز، شماره (۳)۴۶، صفحه ۱۹۶-۱۸۱، سال ۱۳۹۵.
[16] KM. Dolan and S. Agent, "Internet auction fraud: the silent victims", Journal of Economic Crime Management. 2004;2(1):1-22.
[17] M. Jenamani, Y. Zhong and B. Bhargava, Cheating in online auction–Towards explaining the popularity of English auction, Electronic Commerce Research and Applications. 2007;6(1):53-62.
[18] TD. Kavu, T. Rugube, F. Kawondera, and N. Chifamba, "A fraud detection tool in E-auctions", African Journal of Mathematics and Computer Science Research, pp. 1–11, Academic Journals 2016.
[19] D. H. Chau, S. Pandit, and C. Faloutsos, "Detecting Fraudulent Personalities in Networks of Online Auctioneers", pp. 103–114, Springer 2006.
[20] S. Pandit, D. Chau, S. Wang, and C. Faloutsos, "Netprobe: a Fast and Scalable System for Fraud Detection in Online Auction Networks", Conference on World Wide Web, vol. 42, pp. 201–210, ACM 2007.
[21] S. Tsang, Y. S. Koh, G. Dobbie, and S. Alam, "SPAN: Finding Collaborative Frauds in Online Auctions", Knowledge-Based Syst., vol. 71, pp. 389–408, Elsevier 2014.
[22] S. Z. Li, Markov Random Field Modeling in Image Analysis, 3rd ed., London, UK: Springer-Verlag London, 2009.
[23] K. P. Murphy, Y. Weiss, and M. I. Jordan, "Loopy belief propagation for approximate inference: An empirical study", 15th Conference Uncertainty in Artificial Intelligence, pp. 467–475, ACM 1999.
[24] J. S. Yedidia, W. T. Freeman, and Y. Weiss, "Understanding belief propagation and its generalizations", Exploring Artificial Intelligence in the New Millennium, pp. 239–269, 2003.

زیرنویس‌ها

[1] Honest, Fraud, Accomplice
[2] Edge
[3] Anomaly
[4] Markov Random Field
[5] Node potential
[6] Edge potential
[7] Loopy Belief Propagation
[8] Posterior distribution
[9] Reputation Fraud
[10] Collusive Star
[11] Collusive Clique
[12] True positive rate
[13] False positive rate