سال اول، شماره‌هاي ۱و ۲، پاییز و زمستان ۱۳۸۷
صص: ۹- ۱۸

فصلنامه علمي-پژوهشي
فنّاوری اطلاعات و ارتباطات ایران

طراحی و پیاده سازی سیستم هوشمند شناسایی رفتار مشکوک در بانکداری اینترنتی به کمک نظریهٔ مجموعه‌های فازی

لیلا ساروخانی* **غلامعلی منتظر*[۱]**

* دانشکده فنی ومهندسی، دانشگاه تربیت مدرّس

چکیده

یکی از مهمترین موانع برای استفاده از بانکداری اینترنتی عدم امنیت تراکنشها و برخی سوءاستفاده‌ها در مسیر انجام مبادلات مالی است.به همین دلیل جلوگیری از نفوذ غیرمجاز و تشخیص جرم از مسائل مهم در مؤسسات مالی وبانکهاست. دراین مقاله سیستمی هوشمندی طرح شده است که تشخیص رفتارهای مشکوک و غیرمعمول کاربران در سیستم بانکداری اینترنتی را امکان‌پذیر می‌سازد. از آنجا که رفتار کاربران مختلف همراه با ابهام و عدم قطعیت است این سیستم بر اساس نظریهٔ فازی طراحی شده تا رفتار کاربران را شناسایی کرده و رفتارهای مشکوک با شدتهای مختلف را دسته‌بندی کند. مدل طراحی شده در سیستم بانکداری اینترنتی بانک ملت به عنوان یکی از بزرگترین بانکهای برخط کشور آزمایش شده و نتایج آن نشان دهندهٔ موفقیت سیستم در شناسایی رفتارهای مشکوک، با درجهٔ صحت ۹۳٪ ، است .

کلید واژگان: بانکداری اینترنتی، رفتار مشکوک، سیستم هوشمند، تشخیص الگو، نظریهٔ فازی

۱- مقدمه

دسترسی آسان و گسترده به اینترنت آن را به یکی از معابر اصلی فروش برای خرده فروشی‌ها بدل کرده است. با توجه به وجود توان بالقوه در استفاده از اینترنت حجم مبادلات تجارت الکترونیکی در سالهای گذشته رشد زیادی داشته است[۱]. همچنین یکی از ابزارهای ضروری برای تحقق و گسترش تجارت الکترونیکی، وجود سیستم بانکداری الکترونیکی است که همگام با سیستم‌های جهانی مالی و پولی، عملیات و فعالیت‌های مربوط به تجارت الکترونیکی را تسهیل کند. در حقیقت می‌توان گفت پیاده‌سازی تجارت‌الکترونیکی، نیازمند تحقق بانکداری‌الکترونیکی است. به همین دلیل، استفاده از سیستم‌های الکترونیکی در موسسات مالی و اعتباری جهان به سرعت رو به گسترش بوده و شمار استفاده‌کنندگان از خدمات بانکداری الکترونیکی روز به روز در حال افزایش است از این رو صنعت بانکداری تلاش دارد تا با بکارگیری اینترنت به عنوان یکی از کانالهای اصلی ارائهٔ خدمات، برای نفوذ و قدرتمند ساختن کسب و کار خود استفاده کند[۲].

بانکداری اینترنتی در دو حالت به انجام فعالیتهای مرتبط با کسب و کار می‌پردازد. نخست حالت بدون تراکنش است که در آن بانک از اینترنت به عنوان ابزاری برای تبلیغات دربارهٔ خدمات خود استفاده می‌کند و دوم حالت تراکنشی است که در آن بانک به مشتریان خود اجازه می‌دهد از اینترنت برای انجام عملیات مالی خود شامل مشاهدهٔ ریزکارکرد، مانده، صورتحساب و همچنین انتقال وجه و پرداخت قبوض استفاده کنند. با این که هردو این حالات ریسک پذیر هستند ولی حالت تراکنشی در بانکداری اینترنتی که به آن بانکداری برخط[۱] نیز می‌گویند به علت داشتن

۱. نویسنده عهده‌دار مکاتبات (montazer@modares.ac.ir)

معماری ناهمگن و حضور اجزای مختلف و همچنین انجام فعالیتهای مالی توسط مشتریان ، مستعد پذیرش ریسک بالاتری است[۲].

از سوی دیگر باید اشاره کرد که همراه با افزایش امکانات و خدمات بانکها در اینترنت و رشد روزافزون انجام تراکنشهای برخط توسط مشتریان، میزان بروز جرائم مالی در صنعت بانکداری اینترنتی نیز به سرعت در حال رشد است، به طوری که آهنگ رشد جرایم برخط بین ۸ تا ۹ درصد در سال تخمین زده می‌شود[۱]. آمارها نشان می‌دهند که میزان ضرر مالی بانکها از این جرایم در انگلستان در سال ۲۰۰۷ میلادی بالغ بر ۲۲/۶ میلیون پوند بوده است[۳] از این رو بانکها به سرعت در حال بهبود بخشیدن و سرمایه‌گذاری بر روی سیستم‌های ضد سرقت خود هستند. با توجه به رشد تهدیدات و حملات کامپیوتری که با انگیزه‌های مالی انجام می‌شود امنیت صنعت بانکداری باید به عنوان یک موضوع مهم شناخته شود[۴] چراکه از دست‌دادن اعتماد مشتریان به دلیل وقوع کلاهبرداری در این نوع خدمات در نهایت باعث به خطرافتادن اقتصاد عمومی خواهد شد. بدین‌لحاظ روشهای مختلفی برای شناسایی جرم، رفتارهای غیرمعمول و مغایر با قانون کاربران وجود دارد و تحقیقات مشابهی نیز در صنایع مختلف مانند صنایع بهداشتی[۵]، ارتباطات راه دور[۶]، بخش‌مالی [۷]، پست الکترونیکی[۸] وغیره انجام شده است،علاوه براین روشهای مختلفی برای مواجهه بارفتارهای فریب آمیز در بانکداری اینترنتی موردتوجه قرارگرفته که در ادامه شرح مختصری از آنها ذکر می‌شود. بانکهای فراهم‌کنندهٔ خدمات اینترنتی، روشهای مختلفی را برای تشخیص جرم و غربال کردن تراکنشهای مشتریان به کار می‌برند. روشهایی که اخیراً استفاده می‌شود شامل مشاهدهٔ تراکنشها از طریق سیستمهای تصدیق نشانی[۲](AVS)، روش تصدیق کارت[۳](CVM)، شمارهٔ شناسایی شخصی(PIN)[۴] و روشهای زیست‌سنجی[۵] است. AVS شامل شناسایی نشانی از طریق کدهای زیپ شدهٔ مشتریان است در حالی که CVM و PIN شامل بررسی عددی است که مشتری به عنوان رمزعبور برای خود در نظر گرفته است. زیست سنجی نیز شامل شناسایی و تصدیق امضا یا اثرانگشت مشتری است[۱]. روشهای مبتنی بر قواعد، نگهداری سیاههٔ مشتریان بدحساب و خوش‌حساب و تقسیم بندی آنها برحسب مناطق جغرافیایی مختلف نیز از جمله اطلاعاتی است که برای کمک به تشخیص جرم استفاده شده است. همچنین یکی دیگر از روشهای مورد استفاده در تشخیص جرم، استفاده از روشهای داده‌کاوی است که بر تحلیلهای آماری و کشف رفتار مشتریان و استفاده از الگوها برای شناسایی جرم تمرکز دارند [۱]. این روشها مبتنی بر یادگیری قواعدی خاص هستند و قادرند شاخص‌های رفتارهای فریب‌آمیز را از پایگاه داده‌های بزرگ تراکنشهای کاربران کشف کنند. این شاخصها برای ایجاد سیستمهای پایشگر[۱] استفاده می‌شوند تا رفتارهای غیرمعمول مشتریان را ثبت کرده و رفتارهای مشکوک را از میان آنها شناسایی کنند. در نهایت خروجی این سیستم‌ها می‌تواند برای اعلام هشدار و اخطار درخصوص کاربران متخلف استفاده شود[۹]. قواعد وابستگی[۲] نیز یکی از بهترین روشهای داده‌کاوی برای خلق چنین مدلهایی است. در[۱۰] از این روش برای استخراج دانش و به‌دست آوردن الگوهای رفتاری غیرمتعارف از مجموعهٔ بزرگ تراکنشهای کاربران در پایگاه‌داده‌های تراکنش کارت اعتباری برای تشخیص و جلوگیری از وقوع جرم استفاده شده است. روش دیگری که تاکنون برای شناسایی و تشخیص جرم استفاده شده شبکه‌های عصبی مصنوعی است که قابلیت استخراج الگو از پایگاه‌داده‌های حاوی تراکنشهای گذشتهٔ مشتریان را دارند. این شبکه‌ها آموزش پذیر بوده و قابلیت انطباق با شکلهای جدید جرم را دارا هستند [۱]. در[۴] از یک شبکهٔعصبی برای تشخیص جرم در بانکداری اینترنتی با کمک فرایند یادگیری بانظارت و استفاده از مجموعه‌های یادگیرنده برای ساختن مدلهایی از تراکنشهای فریب‌آمیز بانکداری اینترنتی ارائه شده است. شبکهٔ مطرح شده با کمک مجموعهٔ بزرگی از تراکنشها که می توانند طی فرایندی الگوی فعالیتهای غیرقانونی و مجرمانه را شناسایی کنند طراحی شده است. علاوه بر این در [۱۱] یک سیستم نروفازی[۳] برای یافتن حسابهای بی‌اعتبار و پیش‌بینی درمورد این حسابها ارائه شده که با دقت نسبتاً بالایی شناسایی این گونه حسابها را امکانپذیر ساخته است در این مقاله از تاریخچهٔ تخلفات انجام شده در سیستم و همچنین اطلاعات مربوط به افرادی که در بازپرداخت وام خود دچار مشکل شده‌اند استفاده شده است. از منطق فازی نیز در صنعت بیمه، شبکه‌های رایانه‌ای، ارتباطات راه

۱.Online Banking
۲.Address Verification systems(AVS)
۳.Card Verification Method(CVM)
٤.Personal Identification Number(PIN)
٥.Biometrics

۱.monitoring
۲.Association Rules
۳.Neuro fuzzy

دور و غیره برای شناسایی جرائم الکترونیکی استفاده شده است. به عنوان مثال در [۱۲] از یک سیستم فازی سلسله‌مراتبی برای تشخیص تغییرات غیرعادی رفتارترافیکی پست‌های الکترونیکی استفاده شده است در این مقاله اطلاعات مربوط به فرستنده و دریافت‌کننده، تاریخ و زمان ارسال نامه، استخراج و براساس آن مقیاسهای مختلفی برای اندازه گیری رفتار استاندارد پست‌الکترونیکی کاربران تعیین می‌شود. آنگاه برپایهٔ این رفتار استاندارد وبه کمک یک سیستم‌فازی سلسله مراتبی رفتار غیرعادی کاربر شناسایی می‌شود. در [۱۳] برای تشخیص حمله در یک شبکهٔ رایانه‌ای از منطق فازی استفاده شده است. در این مقاله یک سیستم هوشمند برای شناسایی حملات طراحی شده که درآن ازدو الگوریتم داده کاوی دسته بندی و قوانین همبستگی فازی به صورت ترکیبی برای پیش‌بینی رفتارهای مختلف در شبکه‌ها استفاده شده است. به این ترتیب که قواعد فازیِ وزن‌دار در هر دسته با یکدیگر تجمیع شده تا تصمیم گیری در مورد اجزای آن دسته انجام شود.

هدف اصلی در این مقاله تشخیص رفتارهای مشکوک مشتریان و دسته بندی آنها با استفاده از نظریهٔ فازی در سیستم بانکداری اینترنتی بانک ملت است. بانک ملت در حال حاضر یکی از بزرگترین بانکهای ارائه دهندهٔ خدمات بانکداری اینترنتی در ایران است و بیش از یکصد هزار کاربر اینترنتی دارد که روزانه حدود ده هزار نفر از این کاربران وارد سیستم شده و مبادلات مالی خود را از طریق اینترنت انجام می‌دهند. میزان تراکنشهای انجام شده در این سیستم به بیش از پنجاه هزار تراکنش در هرروز می‌رسد که میزان رشدی برابر با چهل درصد درسال را نشان می دهداز این رو با توجه به افزایش اقبال عمومی برای استفاده از این نوع خدمات توسط کاربران و همچنین سیاست جدید بانکها و دولت ایران برای تشویق مردم در جهت استفادهٔ بیشتر از خدمات بانکداری‌الکترونیکی و اینترنتی، فراهم‌کردن امکاناتی برای امن ساختن این‌گونه سیستم‌ها و افزایش اطمنیان مردم برای استفاده از این خدمات ضروری است. مقالهٔ حاضر شامل بخشهای زیر است: بخش ۲ به مروری اجمالی بر نظریهٔ فازی و نکات مهم برای استفاده از این نظریه می‌پردازد پس از آن معماری کلی سیستم شناسایی رفتار براساس نظریهٔ فازی مطرح شده است. سپس نحوهٔ پیاده‌سازی آن در سیستم بانکداری اینترنتی بانک ملت و نیز نتایج کاربرد سیستم تشریح شده است و درنهایت بخش آخر به نتیجه گیری اختصاص یافته است.

۲- اجمالی برنظریهٔ سیستم‌های فازی

نظریهٔ مجموعه‌های فازی در سال ۱۹۶۵ میلادی توسط عسکر لطفی زاده مطرح شد. نظریهٔ مجموعه‌های فازی روشی را برای محاسبهٔ داده‌ها و اطلاعات غیر قطعی و مبهم ارائه می‌کند ضمن اینکه سازوکار استنتاج، برای استدلال را براساس مجموعه‌ای از قواعد "اگر-آنگاه" فراهم می سازد. این قواعد به کمک مجموعه‌های فازی تعریف می‌شوند که در آنها هریک از اعضای مجموعه درجهٔ تعلقی بین صفر و یک دارند. یک نمونه واقعی از عدم قطعیت وجود ابهام در زبان طبیعی انسانهاست [۱۴]. سیستمهای فازی مفاهیم نظریهٔ مجموعه فازی و منطق فازی را با یکدیگر تلفیق و چارچوبی برای ارائهٔ دانش زبانی همراه با عدم قطعیت فراهم می‌کنند و دو مشخصهٔ اصلی دارند که محبوبیت آنها را بیشتر کرده است: یکی اینکه آنها برای استدلال تقریبی[۱] به ویژه برای سیستمهایی که استخراج یک مدل ریاضی از آنها کار دشواری است، مناسب بوده و دیگری اینکه منطق فازی اجازه می‌دهد تصمیم‌گیری با استفاده از اطلاعات ناکامل و غیرقطعی با کمک متغیرهای زبانی[۲]، که به راحتی توسط انسانها قابل درک هستند، انجام شود.

سیستم‌های مبتنی بر منطق فازی شامل چهار جزء اصلی هستند و همانطور که ذکر شد می‌توانند راه حل های عملی و مناسبی را در شرایط مختلف ارائه دهند هریک از این اجزا به طور خلاصه در ذیل تشریح شده است[۱۱]:

الف- فازی‌ساز[۳]: در فرایند فازی‌سازی روابط بین ورودی‌ها و متغیرهای زبانی با استفاده از توابع عضویت تعریف می‌شود. در این مرحله مقادیر ورودی به درجهٔ تعلق متغیرهای زبانی متناظر تبدیل می‌شوند. در واقع متغیرهای ورودی از طریق واحد فازی‌ساز به اعداد فازی تبدیل می‌شوند. در این مقاله هر متغیر ورودی به علت محاسبات ساده تر به یک عددفازی مثلثی تبدیل شده است. هر عدد فازی مثلثی با سه تایی $\tilde{A}=(a_1,a_2,a_3)$ نشان داده می‌شود که در آن $a_1 \le a_2 \le a_3$ است و تابع عضویت آن به شکل زیر نمایش داده می شود:

$$\mu_{\tilde{A}}=\begin{cases}0 & x<a_1\\ (x-a_1)/(a_2-a_1) & a_1\le x\le a_2\\ (x-a_3)/(a_2-a_3) & a_2\le x\le a_3\\ 0 & x>a_3\end{cases} \qquad (۱)$$

۱.Approximate reasoning
۲.Linguistic variables
۳.Fuzzifier

ب- پایگاه‌دانش[۱]: پایگاه دانش از ترکیب دانش خبرگان حوزهٔ مورد بحث به وجود می‌آید و به شکل قواعدی از متغیرهای زبانی تشکیل می‌شود. این قواعد برای بیان ارتباط میان مجموعه‌های فازی ورودی و خروجی استفاده می‌شود. قالب گرامری یک قانون فازی به شکل زیر بیان می‌شود:

اگر (شرایط ورودی برقرار باشد) آنگاه (مجموعه نتایج خروجی قابل استنتاج است)

برای ایجاد پایگاه قواعد فازی در سیستم حاضر، نظریات خبرگان جمع‌آوری و با استفاده از آنها و عملگرهای سه گانهٔ فازی (شامل "یا"، "و"، "نه") در میان هفت متغیر ورودی ، ۵۰ قاعدهٔ "اگر- آنگاه" ایجاد شد. جزئیات این مرحله در بخش سوم تشریح شده است.

ج- موتور استنتاج[۲]: این بخش واحد تصمیم‌گیر سیستم فازی است. یک موتور استنتاج قابلیت استنتاج خروجی‌ها با استفاده از قواعد و عملگرهای فازی را داراست بدین معنا که عملگرهایی مانند: کمینه، بیشینه و یا مجموع را ترکیب و خروجی فازی را از مجموعه‌های فازی ورودی و روابط فازی استخراج کرده و از این طریق توانایی تصمیم‌گیری در انسان را شبیه‌سازی می‌کند. دراین مقاله موتوراستنتاج ممدانی[۳] به عنوان هستهٔ سیستم فازی انتخاب شد که از طریق روابط زیر فرایند بکارگیری ورودیها را بر اساس قواعد تعریف شده اعمال می‌کند [۱۵].

$$\mu_{A^k \to B^k}(x,y) = \min\{\mu_{A^k}(x), \mu_{B^k}(y)\} \qquad (۲)$$

$$\mu_{A^k \to B^k}(x,y) = \mu_{A^k}(x), \mu_{B^k}(y) \qquad (۳)$$

د- نافازی‌ساز[۴]: این مرحله عکس فرایند فازی‌سازی را انجام می‌دهد. نافازی‌ساز، یک خروجی با مقدار قطعی از مجموعه‌های فازی که خروجی موتور استنتاج هستند تولید می‌کند. روش‌های زیادی برای نافازی سازی مطرح شده است که در این مقاله از روش گرانیگاه[۵] با کمک رابطهٔ (۴) استفاده شده است [۱۵].

$$y' = \frac{\int_s yi\mu_B(y)dy}{\int_s \mu_B(y)dy} \qquad (۴)$$

همچنان که اشاره شد در این مقاله از یک سیستم خبرهٔ استنتاج فازی برای تشخیص رفتارهای مشکوک مشتریان بانکداری اینترنتی استفاده شده است. نحوهٔ طراحی و معماری سیستم فازی در بخش بعد تشریح شده است.

۳- معماری سیستم خبرهٔ شناسایی رفتار در بانک ملت

برای استفاده از قابلیت سیستم فازی در شناسایی رفتارهای مشکوک، نخست کلیهٔ رفتارهای کاربران در پنج سطح مختلف دسته‌بندی و سپس سیستم خبرهٔ فازی برای استنتاج این خروجی‌ها طراحی شده است و برای ارزیابی قابلیت سیستم طراحی شده، این مدل در سیستم بانکداری اینترنتی بانک‌ملت به‌کارگرفته شده است. در ادامه، ابتدا شیوهٔ دسته‌بندی رفتارهای کاربران، تشریح و سپس به توضیح سیستم خبره فازی و نحوهٔ تشخیص رفتارهای مشکوک مشتریان در سامانهٔ بانکداری بانک ملت پرداخته خواهد شد.

۳-۱- دسته‌بندی رفتارهای مشتریان

همچنان که در بخش۱عنوان شد پژوهشهای قبلی غالباً به شناسایی جرم بدون درنظر گرفتن شدت و ضعف آن بسنده می‌کنند [۱، ۱۰]. لیکن هدف اصلی در این مقاله شناسایی رفتارهای مشکوک مشتریان است. بدین لحاظ، برای رسیدن به تعریف واضحی از "کاربر مشکوک" ابتدا باید تعاریف مشخصی از رفتارهای مختلف کاربران بانکداری اینترنتی و رده‌های مختلف آن به‌وجود آید، بدیهی است این دسته‌بندی کمک شایانی به اتخاذ راهبرد صحیح برای نوع برخورد با هریک از آن دسته‌ها می‌نماید. از این رو پنج دسته رفتار جداگانه به شرح زیر تعریف شده است:

الف- رفتارعادی: شامل کاربرانی است که عملیات آنها به صورت عادی، بدون اشتباه و کامل انجام شده است.

ب- رفتارکمی مشکوک: شامل کاربرانی است که در هنگام ورود خطا داشته و به تلاش برای ورود غیرمجاز مظنون هستند.

ج- رفتار مشکوک: شامل کاربرانی است که در هنگام ورود خطاهای پی‌درپی داشته و به تلاش برای ورود غیرمجاز مظنون هستند و همچنین بدون انجام عملیات خاصی مرتباً وارد سامانه شده‌اند.

د- رفتار بسیارمشکوک: شامل کاربرانی است که به تلاش برای ورود غیرمجاز مشکوک و همچنین عملیات خاصی را خارج از عرف معمول تکرار کرده‌اند.

---

۱. Knowledge base
۲. Inference engine
۳. Mamdani
۴. Defuzzifier
۵. Center of gravity (COG)

هـ- *رفتار خطرناک*: شامل کاربرانی است که از مرورگرهای ناشناخته استفاده کرده و رفتارهای یک کاربر مشکوک را نیز انجام داده‌اند.

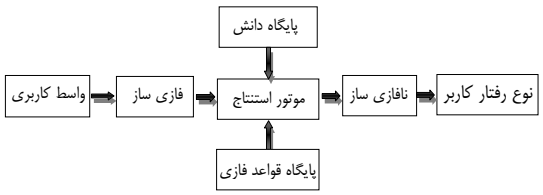
همانطور که از معانی واژه‌ها برمی‌آید شدت غیرمعمول بودن رفتارها به ترتیب از حالت عادی به خطرناک زیاد می‌شود. این رفتارها به کمک اعداد فازی مدل شده و به عنوان متغیرهای زبانی مطابق شکل۱ در خروجی سیستم فازی به کار گرفته شده‌اند بدین معنا که نتیجهٔ استنتاج سیستم فازی تخصیص کاربر به یکی از این پنج دسته خواهد بود.

۳-۲- طراحی سیستم فازی

مدل سیستم خبرهٔ فازی با استفاده از نرم افزار متلب (۷/۰۶) طراحی شده و شامل پنج بخش است:

۱) واسط کاربر که اطلاعات مربوط به متغیرهای ورودی سیستم را از یک پایگاه داده دریافت می‌کند.
۲) پایگاه قواعد فازی
۳) واحد فازی ساز
۴) موتور استنتاج فازی
۵) واحد نافازی ساز

شکل۲ معماری کلی این سیستم را نمایش می‌دهد.



شکل۱: معماری سیستم خبرهٔ فازی

۳-۲-۱- تعیین پارامترهای ورودی

در مرحلهٔ اول واسط کاربری، اطلاعات مربوط به متغیرهای ورودی سیستم را به شکل اعدادی قطعی دریافت می‌کند. برای تعیین پارامترهای ورودی، سیستم به طور کامل بررسی شد این بررسی از دیدگاه کاربری و سیستمی صورت گرفت. از آنجا که پارامترهای ورودی می‌باید از فایل ثبت ورود[۱] رفتار کاربران استخراج شود ابتدا جدولهایی که رفتار کاربران با ذکر جزئیات آنها ثبت می‌کند، به طور کامل بررسی شد. بدین منظور دو جدول شناسایی شد که اطلاعات مختلفی شامل نوع مرورگر و شمارهٔ شناسایی اتصال به اینترنت، نوع فعالیت و تراکنشی که انجام داده به همراه جزئیات آن و همچنین تاریخ و زمان انجام هر فعالیت و تراکنش از ورود تا پرداخت قبض، انجام حواله، گرفتن صورتحساب برای هر کاربر در آن ثبت شده است. پس از تحلیل آنها و با کمک خبرگان بانک، پارامترهایی استخراج شد که لزوماً در جدولهای اولیه موجود نبود و با انجام محاسبات مختلف به‌دست آمدند. به تعبیر دیگر این اقلام اطلاعاتی به صورت ساخت‌نایافته در جدولها درج شده بود لذا عملیات پیش‌پردازش بر روی این جدولها صورت گرفت و از این طریق رکوردهای حاوی اطلاعات مربوط به انجام هر تراکنش برای هر کاربر تجزیه شدند و هریک از اقلام اطلاعاتی برای هرکاربر در هر روز به صورت جداگانه استخراج شد. پس از این مرحله برای استخراج اقلام‌اطلاعاتی که در طراحی سیستم فازی مورد نیاز بود از نظر بیست خبرهٔ بانکداری اینترنتی استفاده شد که در نتیجهٔ آن برخی پارامترها حذف و برخی اضافه شدند. بعد از تحلیل نهایی، هفت پارامتر که در تعیین رفتار کاربر در سامانهٔ بانکداری‌اینترنتی نقش دارند به عنوان متغیرهای ورودی و یک پارامتر (رفتار کاربر) به عنوان متغیر خروجی در سیستم تعیین شدند که نام و مفهوم هریک از آنها مطابق جدول۱ است. بدیهی است، این متغیرها متناسب با سامانهٔ بانکداری اینترنتی بانک ملت در ایران بوده و ممکن است برخی از موارد آن در کشورهای دیگر تغییر کند.

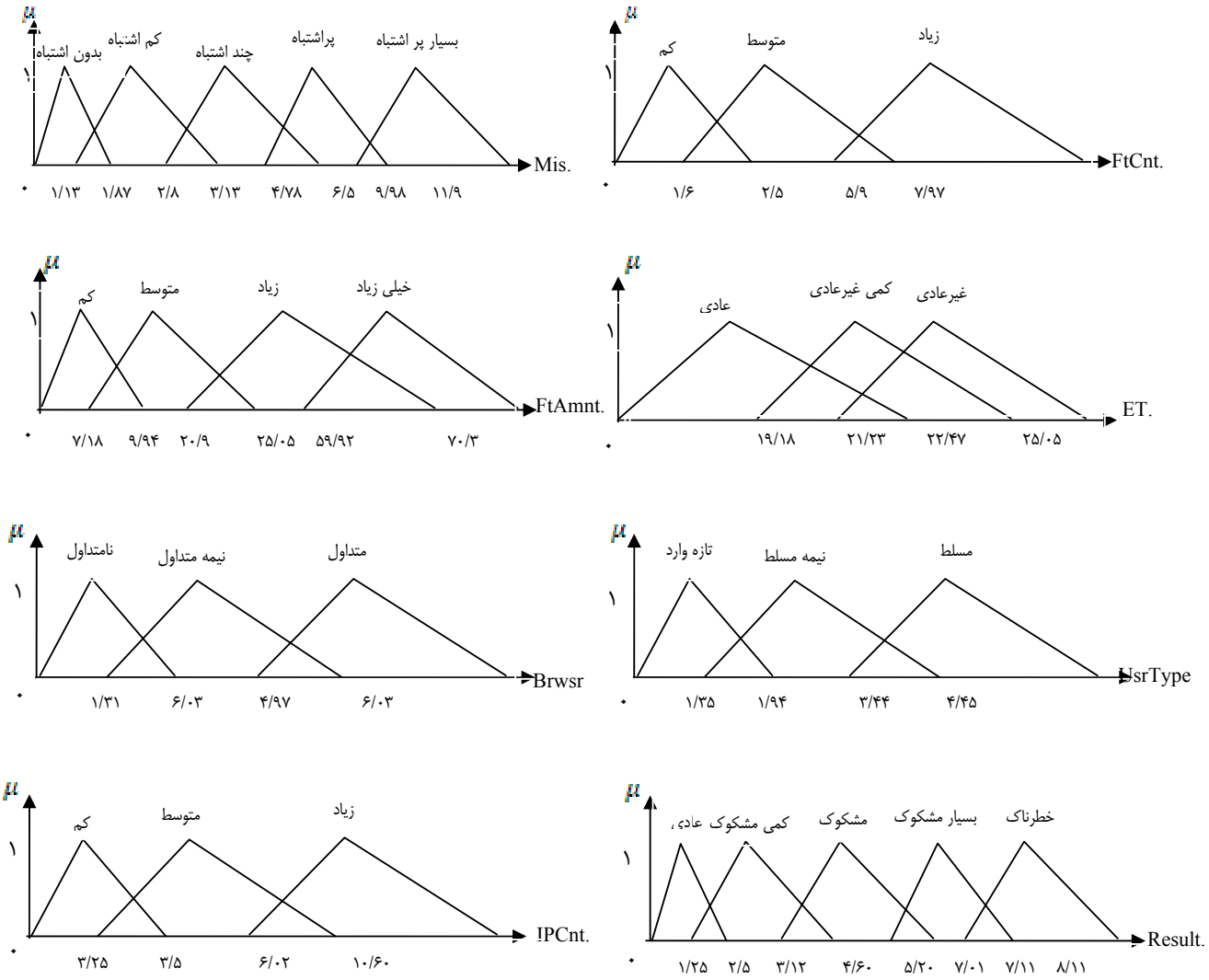
۳-۲-۲- تعریف توابع عضویت متغیرهای ورودی

در این بخش برای هر متغیر ورودی مجموعه‌ای از واژه‌های زبانی تعریف شدند که تعداد آنها از سه تا پنج واژه برای هر پارامتر ورودی متغیر است. سپس از طریق پرسشنامه، از خبرگان بانکداری اینترنتی درخواست شد تا مقادیر و بازه‌های این واژه‌ها را تعیین کنند. در این پرسشنامه، هر پارامتر ورودی و مفهوم آن، واژه‌های زبانی هریک از متغیرها به همراه مقادیر عددی آنها، در قالب پنج گزینهٔ متفاوت برای هر واژه در اختیار خبرگان قرار گرفت. پس از جمع‌آوری نتایج، میانگین نظر خبرگان به شکل اعداد فازی مثلثی به عنوان مقادیر ورودی سیستم تعیین شد که نتایج آن در جدول ۲ آمده است. این بخش از آن جهت اهمیت دارد که این متغیرها و مقادیر آنها در گام بعدی و در پایگاه قواعد فازی استفاده شده‌اند. علاوه بر این نمایش متغیرهای زبانی ورودی و خروجی در شکل۲ نشان داده شده است.

۱.Log file

جدول۱: نام و مفهوم متغیرهای زبانی ورودی وخروجی

ردیف	نام متغیرها	نماد	واحد	مفهوم
۱	تعداد اشتباه	Mis	بار	تعداد خطاهای کاربر هنگام ورود به سیستم
۲	تعداد حواله	FtCnt	بار	تعداد حواله‌های اینترنتی که کاربر انجام داده است
۳	مبلغ حواله	FtAmnt	صدهزاریال	مبلغ حواله‌های اینترنتی که کاربر انجام داده است
۴	تعداد IP	IPCnt	–	تعداد IPهای مختلفی که در هنگام ورود کاربر به سامانه ثبت شده است.
۵	زمان ورود	ET	ساعت	ساعاتی از شبانه روز که کاربر از سامانه استفاده کرده است
۶	قدمت کاربر	UsrType	ماه	مدت زمان آشنایی کاربر با سامانه
۷	نوع مرورگر	Brwsr	–	نوع مرورگر کاربر از لحاظ متداول بودن
۸	رفتار کاربر	Result	–	خروجی – رفتاری که به کاربر تخصیص می‌یابد



شکل۲: مقادیر زبانی متغیرهای ورودی و خروجی

جدول۲: مقادیرزبانی متغیرهای ورودی و خروجی

ردیف	نام متغیر	مقادیر زبانی	اعداد فازی
۱	تعداد اشتباه هنگام ورود به سیستم (Mis)	بدون اشتباه	[۰ ,۱ ,۱/۸۷]
		کم اشتباه	[۱/۱۳, ۲ ,۳/۱۳]
		چند اشتباه	[۲/۸۰, ۴/۵۵ , ۶/۵۰]
		پراشتباه	[۴/۷۸ , ۸/۳۶ , ۱۱/۹۰]
		بسیار پراشتباه	[۹/۹۸ , ۱۶/۰۳ , ۲۴/۳۳]
۲	تعداد حوالهٔ اینترنتی (FtCnt)	کم	[۰ ,۱ ,۲/۵۰]
		متوسط	[۱/۶۰, ۴/۷۳ ,۷/۹۷]
		زیاد	[۵/۹ ,۱۳/۲۳ ,۱۸/۸۰]
۳	مبلغ حوالهٔ اینترنتی (FtAmnt)	کم	[۰, ۱ ,۹/۹۴]
		متوسط	[۷/۱۸, ۱۶/۰۹ ,۲۵/۰۵]
		زیاد	[۲۰/۹ , ۴۶/۵ , ۷۰/۳۰]
		خیلی زیاد	[۵۹/۹۲ ,۹۲/۹۲, ۱۳۵/۶۰]
۴	زمانهای ورود به سیستم (ET)	عادی	[۰, ۱۳/۹۰ ,۲۲/۴۷]
		کمی غیرعادی	[۱۹/۱۸ ,۱۶/۰۹ ,۲۵/۰۵]
		غیرعادی	[۲۱/۲۳ ,۲۵/۳۳ ,۳۰/۴۷]
۵	نوع‌کاربر از لحاظ قدمت (UsrType)	تازه وارد	[۰, ۱, ۱/۹۴]
		نیمه مسلط	[۱/۳۵, ۲/۸۰ , ۴/۴۵]
		مسلط	[۳/۴۴, ۹/۹۷ ,۱۶/۸۰]
۶	نوع مرورگر (Brwsr)	نامتداول	[۰/۰, ۰/۹۹ ,۱/۹۴]
		نیمه متداول	[۱/۳۱, ۳/۷۲, ۶/۰۳]
		متداول	[۴/۹۷ ,۷/۷۴, ۱۰/۵۶]
۷	تعداد IP کاربر (IPCnt)	کم	[۰ ,۱ ,۳/۵۰]
		متوسط	[۳/۲۵, ۷/۱۲ ,۱۰/۶۰]
		زیاد	[۶/۰۲ , ۹/۶۶ ,۱۵/۳۰]
۸	خروجی رفتار کاربر (Result)	عادی	[۰ ,۱ ,۲/۵۰]
		کمی مشکوک	[۱/۲۵, ۲/۳۲ ,۴/۶۰]
		مشکوک	[۳/۱۲, ۵/۲۱ ,۷/۰۱]
		بسیار مشکوک	[۵/۲۰, ۶/۶۲ ,۸/۱۱]
		خطرناک	[۷/۱۱, ۹/۲۲ ,۹/۸۵]

۳-۲-۳- تولید پایگاه قواعد فازی

در این مرحله پایگاه قواعد فازی با استفاده از متغیرهای زبانی ورودی و نظریات خبرگان با ۵۰ قاعدهٔ "اگـر- آنگـاه" مطـابق آنچه در بخش دوم بیان شـده اسـت، ایجـاد شـد. تعـدادی از مهمترین قواعد حاصل از نظرات خبرگـان در جـدول ۳ آمـده است.

جدول۳: نمونه‌هایی از قواعد پایگاه قواعد فازی

ردیف	شرح قاعده
۱	اگر کاربر، کم اشتباه و تعداد حواله متوسط و تعدادآی پی متوسط و زمان ورود غیرعادی و نوع مرورگر متداول باشد آنگاه رفتار کمی مشکوک است.
۲	اگر کاربر، کم اشتباه و تعداد حواله متوسط و تعدادآی پی متوسط و زمان ورود غیرعادی و نوع مرورگر غیرمتداول باشد آنگاه رفتار مشکوک است.
۳	اگر کاربر، کم اشتباه و تعداد حواله متوسط و تعدادآی پی متوسط و زمان ورودعادی و نوع مرورگر غیرمتداول باشد آنگاه رفتارکمی مشکوک است.
۴	اگر کاربر، چند اشتباه و تعداد حواله متوسط و مبلغ حواله زیاد و تعدادآی پی متوسط و زمان ورود کمی غیر عادی و نوع مرورگر غیرمتداول باشد آنگاه رفتارمشکوک است.
۵	اگر کاربر، چند اشتباه و تعداد حواله متوسط و مبلغ حواله زیاد و تعدادآی پی متوسط و زمان ورود کمی غیر عادی و نوع مرورگر نیمه متداول باشد آنگاه رفتارکمی مشکوک است.
۶	اگر کاربر، بی اشتباه و تعداد حواله متوسط و مبلغ حواله زیاد و تعدادآی پی متوسط و زمان ورود کمی غیر عادی و نوع مرورگر متداول باشد آنگاه رفتارعادی است.
۷	اگر کاربر، بی اشتباه و تعداد حواله زیاد و مبلغ حواله خیلی زیاد و تعدادآی پی کم و زمان ورود عادی و نوع مرورگر متداول باشد آنگاه رفتار عادی است.
۸	اگر کاربر، بی اشتباه و تعداد حواله زیاد و مبلغ حواله خیلی زیاد و تعدادآی پی کم و زمان ورود غیرعادی و نوع مرورگر متداول باشد آنگاه رفتار عادی است.
۹	اگر کاربر، بسیار پر اشتباه و تعدادآی پی زیاد و زمان ورود غیرعادی و نوع مرورگر غیرمتداول باشد آنگاه رفتار بسیار مشکوک است.
۱۰	اگر کاربر، بسیار پر اشتباه و تعدادآی پی زیاد و زمان ورود غیرعادی و نوع مرورگر غیرمتداول و نوع کاربر مسلط باشد آنگاه رفتار خطرناک است.
۱۱	اگر کاربر، پر اشتباه و نوع کاربر نیمه وارد و مبلغ حواله زیاد و تعدادآی پی متوسط و زمان ورود کمی غیر عادی و نوع مرورگر نیمه متداول باشد آنگاه رفتاربسیار مشکوک است.
۱۲	اگر کاربر، پر اشتباه و نوع کاربر مسلط و مبلغ حواله زیاد و تعدادآی پی متوسط و زمان ورود غیر عادی و نوع مرورگر متداول باشد آنگاه رفتاربسیار مشکوک است.
۱۳	اگر کاربر، پر اشتباه و نوع کاربر مسلط و مبلغ حواله زیاد و تعدادآی پی متوسط و زمان ورود غیر عادی و نوع مرورگر غیرمتداول باشد آنگاه رفتارخطرناک است. است.
۱۴	اگر کاربر، پر اشتباه و نوع کاربر ناوارد و مبلغ حواله متوسط و تعدادآی پی متوسط و زمان ورود عادی و نوع مرورگر متداول باشد آنگاه رفتار کمی مشکوک است.
۱۵	اگر کاربر، بسیار پر اشتباه و نوع کاربر ناوارد و مبلغ حواله متوسط و تعدادآی پی زیاد و زمان ورود عادی و نوع مرورگر متداول باشد آنگاه رفتار کمی مشکوک است.
۱۶	اگر کاربر، چند اشتباه و نوع کاربر ناوارد و مبلغ حواله متوسط و تعدادآی پی متوسط و زمان ورود عادی و نوع مرورگر متداول باشد آنگاه رفتار عادی است.
۱۷	اگر کاربر، چند اشتباه و نوع کاربر مسلط و مبلغ حواله کم و تعداد حواله کم و تعدادآی پی زیاد و زمان ورود غیرعادی و نوع مرورگر نیمه متداول باشد آنگاه رفتار خطرناک است.
۱۸	اگر کاربر، بی اشتباه و نوع کاربر مسلط و مبلغ حواله خیلی زیاد و تعداد حواله کم و تعدادآی پی زیاد و زمان ورود عادی و نوع مرورگر متداول باشد آنگاه رفتار کمی مشکوک است.
۱۹	اگر کاربر،کم اشتباه و نوع کاربر مسلط و مبلغ حواله خیلی زیاد و تعداد حواله کم و تعدادآی پی زیاد و زمان ورود غیرعادی و نوع مرورگر متداول باشد آنگاه رفتار کمی مشکوک است.
۲۰	اگر کاربر،پر اشتباه و نوع کاربر مسلط و مبلغ حواله خیلی زیاد و تعداد حواله کم و تعدادآی پی زیاد و زمان ورود عادی و نوع مرورگر متداول باشد آنگاه رفتار خطرناک است.
۲۱	اگر کاربر،کم اشتباه و نوع کاربر مسلط و مبلغ حواله زیاد و تعداد حواله زیاد و تعدادآی پی زیاد و زمان ورود عادی و نوع مرورگر متداول باشد آنگاه رفتار کمی مشکوک است.

۴- پیاده سازی سیستم

در این مرحله سیستم خبرهٔ فازی با استفاده از اطلاعاتی که از محیط واقعی سیستم بدست آمده بود به مرحلهٔ اجرا درآمد. برای ارزیابی سیستم چندین نمونه از اقلام اطلاعاتی مربوط به کاربران مختلف، که در پایگاه دادهٔ بانکداری اینترنتی بانک ملت وجود داشت به سیستم داده شد که دو نمونه از آنها با جزئیات بیشتر در ذیل تشریح شده است.

الف- مورد کاوی اول

کاربری، دریک روز، ۷ بار برای ورود به سیستم تلاش کرده است که ۲ بار آن بین ساعت ۹ تا ۱۲ صبح، ۳ بار در فاصلهٔ ساعت ۴ تا ۶ بعدازظهر و ۲ مرتبه نیز بین ساعت ۲ تا ۳ نیمه شب بوده است. کاربر فوق در ورود صبح و نیمه شب بدون اشتباه بوده ولی در ورود بعدازظهر خود، در هر ۳ مرتبه رمز عبور را اشتباه وارد کرده و نتوانسته وارد سیستم شود. در ورود صبح و بعدازظهر با دو IP و در نیمه شب در هر بار از یک IP متفاوت استفاده کرده است و همچنین مرورگری که در نیمه شب برای کاربر ثبت شده برای سیستم ناشناخته بوده است. مبلغ ده میلیون ریال در صبح آن روز به حساب کاربر وارد شده که در نیمه شب طی حواله‌ای اینترنتی از حساب وی کسر شده است. این کاربر بیش از ۱۱ ماه است که از سامانهٔ بانکداری اینترنتی استفاده می‌کند. پس از ورود اطلاعات این کاربر، سیستم طراحی شده رفتار کاربر را "بسیار مشکوک" شناسایی می‌کند که از نظر خبرگان بانکداری اینترنتی بانک ملت تأیید شده است.

ب- موردکاوی دوم

کاربری که بیش از ۶ ماه است از سیستم بانکداری اینترنتی استفاده می‌کند در فاصله زمانی ساعت پنج تا شش صبح، پس از ۲ بار اشتباه در درج رمز ورود، در بارسوم وارد سیستم شده و ۵ حوالهٔ اینترنتی با مبالغ خرد و بسیار کم انجام داده است. کاربر فوق، در طول روز، ۳۷ بار، بدون اشتباه وارد سیستم شده ولی هیچ نوع عملیات مالی انجام نداده است. در پایان شب نیز دو مرتبه برای ورود به سیستم تلاش کرده ولی موفق نشده است. همچنین در تلاش نیمه‌شب کاربر، از یک IP استفاده شده ولی برای سیستم ناشناخته بوده است. پس از ورود اطلاعات این کاربر سیستم طراحی شده رفتار کاربر را "مشکوک" شناسایی می‌کند که از نظر خبرگان بانکداری اینترنتی بانک ملت تأیید شده است.

درادامه و طی جدول ۴ نمونه‌های دیگری از عملکرد سیستم به همراه نظر خبرگان در مورد آنها، آمده است.

بدیهی است هرچه خروجی سیستم به ۱ نزدیکتر باشد رفتار کاربر با شدت بیشتری مشکوک است بدین معنا که اعداد خروجی ۰/۵۰ و کمتر از آن رفتار عادی برای کاربر محسوب شده و رفتارهای مشکوک با خروجی بزرگتر از ۰/۵۰ مشخص می‌شوند. همانطور که در جدول نیز مشاهده می شود در ردیف ۱ سیستم تشخیص کاملاً دقیقی نداشته است. طبق نظر کارشناسان نتایج تا حدود ۹۳٪ با واقعیت همخوانی دارد.

جدول۴: نتیجهٔ اجرای سیستم به همراه نظر خبرگان

ردیف	تعداد اشتباه	تعداد IP	تعداد حواله	مبلغ حواله	زمان ورود	نوع مرورگر	نوع کاربر	خروجی سیستم	نظرخبرگان
۱	۳	۵	۲	۸۰	۳	۶	۶	۰/۵۳	مشکوک
۲	۷	۵	۲	۸۰	۸	۶	۱۱	۰/۷۹	بسیار
۳	۲	۴	۱۱	۷۸	۱۱	۷	۸	۰/۵۱	کمی
۴	۵	۲	–	–	۵	۸	۸	۰/۶۳	مشکوک
۵	۲	۲	۲	۶۷	۲۰	۶	۱۰	۰/۱۲	عادی
۶	۲	۲	۱۱	۳۰	۴	۶	۱۰	۰/۶۲	کمی
۷	۱	۴	۸	۹۲	۲	۳	۹	۰/۶۲	مشکوک
۸	۱۴	۵	۱	۱۰۰	۱۳	۵	۶	۰/۸۸	خطرناک
۹	۱۰	۵	۸	۵۰	۱۳	۶	۶	۰/۸۳	خطرناک
۱۰	۱	۱	۴	۲۳	۱۵/۵	۸	۸	۰/۵۰	عادی
۱۱	۱	۱	۱۲	۴۷	۲	۸	۸	۰/۶۲	کمی
۱۲	۲	۲	۱	۶۸	۱۳	۵	۶	۰/۰۸	عادی
۱۳	۱۱	۳	–	–	۲۰	۳	۷	۰/۸۱	خطرناک
۱۴	۲	۸	۳	۱۲	۱۹	۳	۶	۰/۶۲	مشکوک
۱۵	۱	۱	۱	۵۹	۱۵	۷	۴	۰/۵۰	عادی

۵- نتیجه گیری

امروزه تشخیص جرم و بهبود سطح امنیت در صنعت بانکداری الکترونیکی بسیار مهمتر از گذشته شده است. در این مقاله یک سیستم خبرهٔ فازی برای تشخیص رفتارهای مشکوک کاربران بانکداری اینترنتی طراحی شده است. در این سیستم نوع عملکرد کاربر در مواجهه با سیستم بانکداری اینترنتی، به عنوان ورودی سیستم فازی در نظر گرفته شده و خروجی، یکی از پنج دسته رفتار عادی، کمی مشکوک، مشکوک، بسیار مشکوک و خطرناک مشتری خواهد بود. مهمترین مزیت این سیستم نسبت به روشهای به کار رفته در سایر مقالات، نخست امکان مدلسازی رفتار کاربران در پنج دستهٔ مختلف است که با دقت بیشتری نوع رفتار کاربر را پیش بینی می‌کند و دیگر آنکه در نظرگرفتن حیطهٔ وسیعی از متغیرهای ورودی، امکان پوشش جامع‌تری از عوامل شناسایی کنندهٔ رفتار و عملکرد کاربر را مهیا می‌سازد. علاوه بر این پیاده سازی واقعی این سیستم در محیط یکی از بزرگترین بانکهای ارائه دهندهٔ خدمات اینترنتی در کشور حاکی از صحت عملکرد سیستم با درجهٔ ۹۳٪ است که نشان دهندهٔ قوت عملکرد آن است.

مراجع

[۱] J. T. S. Quah and M. Sriganesh, "Real-time credit card fraud detection using computational intelligence," Expert Systems with Applications,۲۰۰۷, pp.۹-۱۷

[۲] D. P. Dube and S. Ramanarayanan, "Internet Banking – A Layered Approach to Security," in Intelligent Information Technology, ۲۰۰۵, pp.۱۹۰-۱۹۷.

[۳] T. Young, "Lords call for e-crime shakeup ." , available online at http://www.computing.co. uk/computing/news/۲۲۲۱۰۲۰/lords-say-crime-nust- ,last visited October ۲۰۰۸.

[٤] K. B. Bignell, "Authentication in an Internet Banking Environment; Towards Developing a Strategy for Fraud Detection," in Internet Surveillance and Protection, ۲۰۰٦. ICISP '۰٦. International Conference on, Cote d'Azur, ۲۰۰٦, pp.۲۳-۲۳.

[٥] P. Alexopoulos, X. Benetou, T. Tagaris, P. Georgolios, and K. Kafentzis, "IWEBCARE: AN ONTOLOGICAL APPROACH FOR FRAUD DETECTION IN THE HEALTHCARE DOMAIN," in International Conference on Information Technologies (InfoTech-۲۰۰۷), BULGARIA, ۲۰۰۷, pp.۱-۸.

[٦] P. A. Estévez, C. M. Held, and C. A. Perez, "Subscription fraud prevention in telecommunications using fuzzy rules and neural networks," Experts Systems with Application, vol. ۳۱, pp. ۳۳۷-۳٤٤, ۲۰۰٦.

[۷] E. Kirkos, C. Spathis, and Y. Manolopoulos, "Data Mining techniques for the detection of fraudulent financial statements," Experts Systems with Application, vol. ۳۲, pp.۹۹٥-۱۰۰۳. ۲۰۰۷.

[۸] K. Kerremans, Y. Tang ,R. Temmerman, and G. Zhao, "Towards Ontology-based E-mail Fraud Detection," in Artificial intelligence, ۲۰۰٥. epia ۲۰۰٥. portuguese conference on: IEEE, ۲۰۰٥, pp. ۱۰٦-۱۱۱.

[۹] L. Fang, M. Cai, H. Fu, and J. Dong, "Ontology-Based Fraud Detection," in Computational Science – ICCS ۲۰۰۷, ۲۰۰۷, pp.۱۰٤۸-۱۰٥٥.

[۱۰] D. Sanchez, M. A. Vila, L. Cerda, and J. M. Serrano, "Association rules applied to credit card fraud detection," Expert Systems with Applications, ۲۰۰۸, pp.۱-۱٤.

[۱۱] Y.-P. Huang, C.-C. Lu, and T.-W. Chang , "An Intelligent Approach to Detecting the Bad Credit Card Accounts," in ۲٥th IASTED International Multi-Conference Artificial Intelligence and Applications, Innsbruck, Austria, ۲۰۰۷, pp. ۱-٦.

[۱۲] M. N. Mark Jyn-Huey Lim, Jacky Hartnett, "Detecting Abnormal Changes in E-mail Traffic Using Hierarchical Fuzzy Systems," IEEE, ۲۰۰۷, pp.۱-۷

[۱۳] R. A. Tansel Ozyera, Ken Barker, "Intrusion detection by integrating boosting genetic fuzzy classifier and data mining criteria for rule pre-screening," Journal of Network and Computer Applications, vol. ۳۰, pp. ۹۹-۱۱۳, ۲۰۰۷.

[۱٤] L. A. Zadeh, "The concept of a linguistic variable and its application to approximate reasoning," Information Sciences, vol.۸, pp. ۱۹۹-۲٤۹,۱۹۷٥.

[۱٥] L. Rutkowski, "Fuzzy Inference Systems," in Flexible Neuro-Fuzzy Systems, ۲۰۰٤, pp. ۲۷-٥۰.