طراحی سیستم خبره به منظور تشخیص حمله‌های فیشینگ در بانکداری الکترونیکی

محمود مقیمی[۱] حسین اکبری پور[۲] محمدرضا امین‌ناصری[۳]

۱- دانشجوی کارشناسی ارشد فناوری اطلاعات، دانشگاه تربیت مدرس
m.moghimi@modares.ac.ir

۲- دانش‌آموخته‌ی کارشناسی ارشد مهندسی صنایع، دانشگاه تربیت مدرس
h.akbaripour@modares.ac.ir

۳- دانشیار- بخش مهندسی صنایع، دانشگاه تربیت مدرس
amin_nas@modares.ac.ir

چکیده: امروزه از مهمترین ریسک‌ها و چالشهای مورد توجه در تجارت الکترونیک و بانکدارای الکترونیکی، خطر کلاهبرداری آنلاین و حملات فیشینگ است. در این تحقیق سیستم خبره‌ای با استفاده از مشخصه‌های ظاهری صفحه، قابلیت‌های امنیتی و نیز اطلاعات موجود در دامنه وب سایت ارائه گردیده است که قادر به استدلال در خصوص میزان مشکوک بودن یک وب سایت به یک حمله فیشینگ در بانکداری الکترونیکی می‌باشد. در سیستم خبره پیشنهادی از شبکه عصبی مصنوعی جهت تشکیل پایگاه دانایی سیستم استفاده شده است. ورودی های سیستم خبره، ۲۷ پارامتر مختلف قابل ارزیابی هستند که از یک صفحه وب استخراج می‌شوند. فرآیند استنتاج نیز با استفاده از موتور استنتاج موجود در پوسته سیستم خبره به ترتیب برای هر کدام از بخش‌های پارامترهای ورودی بصورت مجزا انجام می‌گیرد. درنهایت نیز نتیجه هر بخش در مقایسه با نتیجه سایر بخش ها ارزیابی و خروجی حاصل بعنوان استنتاج نهایی سیستم ارائه می‌گردد. به منظور اعتبارسنجی سیستم پیشنهادی، خروجی حاصله بر اساس مقادیر واقعی مورد ارزیابی قرار گرفت که نتایج قابل قبولی در شناسایی این نوع حملات در مقایسه با سایر سیستم‌های موجود در ادبیات، ارائه نموده است.

کلمات کلیدی: فیشینگ، بانکداری الکترونیکی، سیستم خبره، شبکه عصبی مصنوعی.

تاریخ ارسال مقاله : ۱۳۹۱/۸/۲۹

تاریخ پذیرش مشروط : ۱۳۹۳/۱/۳۰

تاریخ پذیرش مقاله : ۱۳۹۳/۲/۱۳

نام نویسنده‌ی مسئول: محمدرضا امین ناصری

نشانی نویسنده‌ی مسئول: تهران- بزرگراه جلال آل احمد- پل نصر- دانشگاه تربیت مدرس- گروه مهندسی صنایع، شماره تماس: ۸۲۸۸۳۳۴۴

۱- مقدمه

امروزه اینترنت نقش قابل توجهی در فعالیت های تجاری و کسب و کار دارد. امنیت ضعیف اینترنت و دستاوردهای بزرگ مالی ارائه شده در این محیط، انگیزه ای قوی برای مهاجمان به منظور ارتکاب به کلاهبرداری در این عرصه را فراهم نموده است. در تجارت الکترونیک[۱] و بانکدارای الکترونیکی[۲] از مهمترین ریسک ها و چالشهای مورد توجه، خطر کلاهبرداری آنلاین است که معمولاً از طریق روش‌هایی همانند حملات فیشینگ[۳] و سرقت هویت رخ می‌دهد. اگر چه بسیاری از بانک‌ها با ارائه طیف گسترده‌ای از سیستم‌های امنیتی و کنترلی سعی در حفاظت از اطلاعات مشتریان دارند اما امروزه تقلب و کلاهبرداری در بانکداری الکترونیکی به طور چشمگیری افزایش یافته است [۱]. تشخیص و جلوگیری از این نوع حملات، گامی مهم در جهت حفاظت از داده های شخصی و اطلاعات مالی محسوب می‌شود [۲].

روشهای متعددی با بکارگیری الگوهای مبتنی بر سرور و یا وب سایت، سعی در جلوگیری از بروز اینچنین حملاتی در محیط‌های تحت وب دارد [۳]. در الگوهای مبتنی بر سرور، سرویس دهنده سعی می‌کند با استفاده از مکانیزم‌های امنیتی، تصدیق هویت مشتریان را انجام داده و مانع از بروز حملاتی نظیر حملات فیشینگ یا استراق سمع شود [۴] [۵]. روشهایی نظیر استفاده از رمز پویا، سرویس پیام‌کوتاه و یا سرویس های چند لایه تصدیق هویت از این نوع می باشند [۶]. در تحقیق [۷] نشان داده است که با بکارگیری استانداردهایی در زمان طراحی نرم‌افزارهای مربوطه، می توان مانع از بروز این حملات شد. ولی با این وجود، متقاعد کردن تمامی طراحان وب به پیروی از چنین مقرراتی کار نسبتاً دشواری است. حملات فیشینگ با توجه به کانال‌های مختلف مورد استفاده، در انواع مختلفی نظیر نرم‌افزارهای مخرب، ایمیل های فیشینگ، وب سایت های جعلی و سرقت هویت دسته‌بندی می‌گردند [۸]. طبق تقسیم بندی گروه کاری ضد سرقت هویت (APWG)[۴]، به طور کلی مکانیسم‌های دفاعی در مقابل حملات فیشینگ و تقلب، به سه دسته روشهای شناسایی، روشهای پیشگیری و روشهای اصلاحی تقسیم می‌شوند [۹]. برای شناسایی و پیشگری از این نوع حملات تکنیک‌های مختلفی ارائه گردید است که معمولاً در قالب یک مرورگر مستقل، یک پلاگین[۵] برای مرورگر و یا فیلترهای قابل نصب در سرورهای ایمیل، پیاده‌سازی گردیده‌اند. در سیستم‌های ارائه شده در قالب یک پلاگین برای مرورگر و یا یک ابزار شناسایی، سعی شده تا از طریق بررسی اطلاعات وب سایت نظیر دامنه، تصویر وب سایت، موارد امنیتی لحاظ شده و سایر ویژگیهای صفحه‌ی وب اقدام به شناسایی حملات فیشینگ شود [۱۰]. ابزارهای زیادی با استفاده از این روش طراحی شده‌اند که بعنوان نمونه می توان به فیلتر ضد فیشینگ موجود در مرورگر اینترنت اکسپلورر (IE)[۶] اشاره نمود [۱۱]. برخی از ابزارها نیز از طریق یک مرجع و با استفاده از لیست‌های سیاه اقدام به تایید و یا رد یک وب سایت می نمایند. ابزار آنتی‌فیش[۷] که یک پلاگین برای مرورگر است، با استفاده از این رویکرد به منظور محافظت از کاربران در برابر حملات فیشینگ مبتنی بر وب سایتهای جعلی مورد استفاده قرار می‌گیرد [۱۲]. در تحقیق [۸] نیز یک نمونه اولیه مرورگر وب، پیاده سازی گردیده که از طریق پردازش ایمیل‌های ورودی قادر به شناسایی حملات فیشینگ می‌باشد. ضعف روش‌هایی که به یک مرجع بعنوان لیست سیاه متکی هستند، عدم مقیاس پذیری و امکان بهنگام سازی سریع این مراجع است. درواقع این رویکرد به تنهایی از کارایی چندانی در جلوگیری از این نوع حملات برخوردار نیست، چراکه ایجاد سایت‌های فیشینگ اغلب ارزان و آسان بوده و متوسط طول عمر آنها نیز تنها چند ساعت می‌باشد [۱۳] [۳].

استفاده از مشخصه های ظاهری صفحات و دامنه ثبتی وب سایت‌ها، یکی از بهترین و ساده‌ترین روشهای تشخیص حملات فیشینگ است که تحقیقات زیادی در خصوص آن انجام گردیده است [۱۴] [۱۵]. بسیاری از این روشها از الگوریتم های هوش مصنوعی و یادگیری ماشین نظارت شده[۸] استفاده می کنند که در تحقیق [۹] به نمونه‌هایی از آن اشاره گردیده است. پن و دینگ در تحقیق [۳]، روشی جهت شناسایی حملات فیشینگ بر اساس مدل شئی یک سند وب (DOM)[۹] و تشخیص ناهنجاری‌ها، پیشنهاد نمودند. چاندراسکاران و همکاران در مقاله خود، با استخراج ۲۵ ویژگی مختلف شامل سبک بکاررفته در محتوا و امنیت آن و سپس انتخاب و دسته‌بندی آنها از طریق روشهای شبیه سازی تبرید تدریجی (SA)[۱۰] و ماشین بردار پشتیبان (SVM)[۱۱] ، تکنیکی جهت طبقه‌بندی حملات فیشینگ بر اساس خواص ساختاری ایمیل های ارسالی ارائه نمودند [۱۶]. چن و گو نیز الگوریتم جدیدی با عنوان لینک گارد (LinkGuard) برای مقابله با حملات فیشینگ ارائه نمودند که با استفاده از ویژگیهای عمومی این نوع حملات، آنها را شناسایی می نمود [۱۷]. به مرور در روشهای پیشنهادی سعی شد تا بیشتر از تکنیکهای هوش مصنوعی در شناسایی این نوع حملات استفاده گردد که از آن جمله می‌توان به سیستم خبره فازی ارائه

شده توسط آباروس و همکاران اشاره نمود [۱۳]. این سیستم با بکارگیری تکنیک های داده‌کاوی و منطق فازی قادر به شناسایی مشخصه های وب سایت و نهایتاً تعیین میزان امنیت آن سایت می باشد [۱۳] [۱۸]. از تکنیکهای بکارگرفته شده دیگر در این زمینه می‌توان به استفاده از الگوریتم کلونی مورچگان و یا ارزیابی محتوای وب سایت‌ها اشاره نمود [۱۰] [۱۹]. مشکل روش ارائه شده در تحقیق [۱۳] تعداد زیاد قواعد است. در تحقیق [۲۰]، سایر روش‌های ارائه شده با یکدیگر مورد مقایسه قرار گرفته است. مشکل این روش‌ها نیز میزان بالای خطای خروجی الگوریتم‌ها است که باعث عدم شناسایی صحیح این نوع حملات می‌گردد.

در این تحقیق سیستم خبره‌ای به منظور تشخیص حملات فیشینگ در بانکداری الکترونیکی ارائه گردیده است. در سیستم ارائه شده از شبکه‌های عصبی مصنوعی برای کاهش تعداد قواعد و افزایش دقت خروجی سیستم استفاده شده است. قواعد سیستم خبره پیشنهادی با استفاده از مشخصه‌های متمایزکننده وب سایتهای اصلی از تقلبی در تحقیق [۱۳] ارائه گردیده است. این قواعد تشکیل دهنده تمامی حالات ممکن مشخصه‌های فوق را مورد ارزیابی قرار می‌دهند و در نهایت بر اساس مقادیر ورودی، استنتاج لازم صورت گرفته و بعنوان خروجی سیستم ارائه می‌گردد.

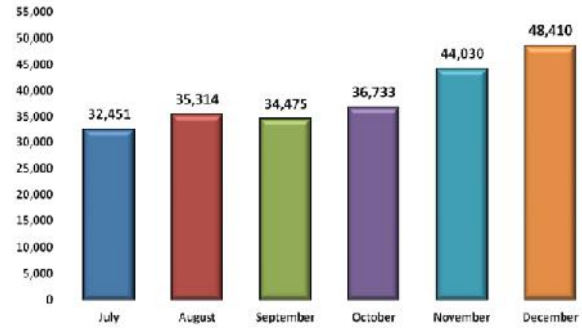
ادامه این مقاله بدین صورت ارائه خواهد شد: در بخش دوم ابتدا توضیحاتی در خصوص حملات فیشینگ ارائه خواهد شد. سپس در بخش سوم و چهارم، با معرفی سیستم خبره و اجزاء آن، سیستم خبره پیشنهای مورد بررسی قرار گرفته و چگونگی پیاده سازی آن تشریح می‌گردد. در بخش پنجم جامعیت سیستم پیشنهادی مورد بررسی قرار گرفته و در پایان در بخش ششم نیز، جمع بندی مطالب و کارهای آتی ارائه گردیده است.

۲- حملات فیشینگ

در دهه گذشته حملات فیشینگ به طور فزآینده ای برای کسب سود آسان از طریق انجام معاملات غیر قانونی مالی افزایش یافته است. در یک تعریف کلی می توان گفت که فیشینگ شکلی از سرقت هویت الکترونیکی است که در آن ترکیبی از مهندسی اجتماعی و روش‌های ساخت وب سایت های جعلی برای فریب کاربر به منظور آشکار کردن اطلاعات محرمانه و ارزشمند مورد استفاده قرار می‌گیرد [۱۸]. در حملات فیشینگ از روشهای خاصی مانند استفاده از وب‌سایت های مخرب و یا پست الکترونیک، برای ربودن این اطلاعات استفاده می‌گردد. در این حملات به نوعی سعی در این است تا مرجع سایت به عنوان یک سازمان قابل اعتماد نشان داده شود [۱]. اغلب کاربران نیز با این نوع حملات آشنا نبوده و در بیشتر موارد با دیدن ظاهر مشابه سایت اصلی، بدون توجه به سایر اطلاعات تکنیکی و امنیتی سایت اقدام به وارد نمودن اطلاعات محرمانه خود در آن سایت می‌نمایند. بدین صورت، مهاجم می‌تواند به راحتی از طریق این اطلاعات اقدام به جعل هویت نماید. شکل (۱) یک نمونه صفحه فیشینگ که دقیقاً مشابه سایت PayPal در زمان حمله است را نشان می‌دهد.

شکل (۱): یک نمونه صفحه فیشینگ برای سایت PayPal

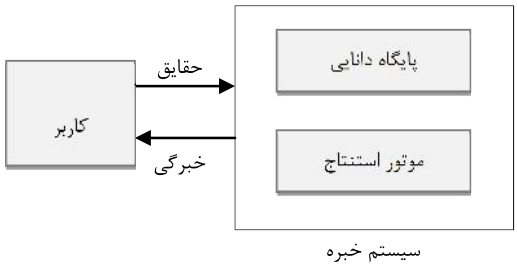
متاسفانه حملات فیشینگ چه در تعداد و چه در پیچیدگی در حال رشد روزافزونی هستند [۱۱]. طبق گزارش گروه APWG در ماه دسامبر سال ۲۰۱۱، بیش از ۴۸ هزار سایت فیشینگ کشف شده که میزبانی ۶۳٪ این سایت ها نیز در ایالات متحده بوده است (شکل(۲)) [۲۱]. رمز موفقیت این نوع از حملات بر قدرت جلب اعتماد مردم استوار است و مهاجمان از هر چیزی که بتواند آنان را قانونی جلوه نماید، استقبال خواهند کرد [۱۱]. مهاجمان فیشینگ، با استفاده از تکنیکهای مختلف از طریق طعمه گذاری، یک کاربر را به سمت بازدید از یک سایت جعلی هدایت می‌کنند [۱۱].



شکل(۲): تعداد سایت‌های فیشینگ کشف شده در نیمه دوم سال ۲۰۱۱ [21]

۳- سیستم خبره و اجزای آن

سیستم خبره یک برنامه کامپیوتری هوشمند است که قادر به شبیه سازی قضاوت و رفتار یک انسان دارای دانش تخصصی و یا تجربی در یک زمینه خاص، می باشد. سیستم های خبره برای حل مسائل پیچیده مانند یک فرد خبره، از طریق استدلال در مورد دانش طراحی می شوند [۲۲]. هر سیستم خبره مطابق شکل (۳)، از سه بخش اصلی پایگاه دانایی[۱۲]، موتور استنتاج[۱۳] و رابط کاربری[۱۴] تشکیل شده است.



شکل (۳): مدل پایه یک سیستم خبره

- *پایگاه دانایی*؛ در یک سیستم خبره پایگاه دانایی دربرگیرنده دانشی است که جهت ارائه استنتاج مورد استفاده قرار می-گیرد. درواقع این پایگاه دانایی از طریق کسب حقایق و مهارت های یک فرد خبره ساخته می‌شود و سپس می-بایست توسط روشی جهت انجام استنتاج بازنمایی گردد [۲۲].
- *موتور استنتاج*؛ برای دستیابی به نتایج مورد نظر، سیستم خبره با جستجو در پایگاه دانایی و بر اساس منطق استنتاج و پردازش قواعد، استنتاج لازم را انجام می‌دهد. این موتور با استفاده از یک روش استدلال که به صورت استدلال پیشرو یا پسرو و یا مخلوطی از هر دو می باشد، عمل استنتاج را انجام می‌دهد.
- *واسط کاربر*؛ این بخش فراهم کننده ارتباط بین کاربر و سیستم خبره می‌باشد. واسط کاربر یک سیستم خبره نه تنها کاربر را قادر می‌سازد تا به سوالات پاسخ دهد بلکه کاربر را مجاز می‌سازد عملیات اجرایی سیستم را با پرسش در مورد توضیحات داده شده قطع نماید. در پایان عملیات استنتاج نیز خروجی سیستم خبره از طریق واسط کاربر نمایان می‌گردد.

۴- سیستم خبره پیشنهادی به منظور تشخیص حملات فیشینگ

سیستم خبره پیشنهادی با استفاده از مشخصه‌های متمایز کننده یک وب سایت معتبر از یک وب سایت جعلی، قادر به ارائه استدلال لازم در جهت میزان فیشینگ بودن یک وب سایت می باشد. ایده اصلی در سیستم خبره پیشنهادی استفاده از یک شبکه عصبی مصنوعی برای کاهش تعداد قواعد و همچنین افزایش سرعت استنتاج است. در ادامه اجزای مختلف سیستم خبره پیشنهادی و نحوه پیاده سازی آن توضیح داده خواهد شد.

۴-۱- پایگاه دانایی

در تحقیق [۱۳] خصوصیات قابل ارزیابی جهت شناسایی وب سایت های فیشینگ، در قالب ۲۷ پارامتر و در شش بخش مختلف ارائه شده است. این شش بخش شامل مشخصه‌های دامنه وب سایت، قابلیت‌های امنیت و رمزنگاری مورد استفاده، کدهای اسکریپتی موجود در صفحه، ظاهر صفحه و محتوای آن، آدرس صفحه وب و ویژگیهای رفتاری سایت می‌باشند. ترکیب مقادیر مختلف قابل استخراج برای هریک از این خصوصیات، تشکیل پایگاه دانایی سیستم را می دهد. سیستم خبره نیز می تواند بر اساس این دانش ذخیره شده، استنتاج لازم در خصوص میزان مشکوک بودن وب سایت مورد نظر به یک وب سایت فیشینگ را انجام دهد. نحوه جمع آوری مقادیر هر یک از این خصوصیات صفحه نیز از اهمیت بسیاری برخوردار است. صحت و تأیید پایگاه دانش نقش بسیار مهمی در استنتاج درست و تصمیم گیری صحیح یک سیستم هوشمند را ایفا می‌نماید. روشهای مختلف استخراج این مقادیر از محتوای صفحات وب، در تحقیق [۳] ارائه شده است.

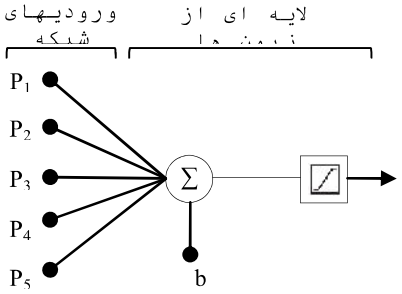
در سیستم پیشنهادی، مشخصه‌های قابل ارزیابی یک صفحه بعنوان پارامتر ورودی برای سیستم خبره درنظر گرفته شده‌اند. لیست کلی این پارامترها در قالب شش بخش و سه لایه مختلف در جدول (۱) نشان داده شده است. لایه‌ی اول فقط حاوی پارامترهای مورد ارزیابی بخش مربوط به مشخصه های دامنه وب سایت می‌باشد. لایه‌ی دوم شامل دو بخش امنیت و رمزنگاری و کدهای اسکریپتی موجود در صفحه می‌باشد. لایه‌ی سوم نیز دربرگیرنده‌ی پارامترهای واقع در بخش‌های مربوط به ظاهر صفحه و محتوای آن، آدرس صفحه وب و ویژگیهای رفتاری سایت می‌باشد [۱۲]. پارامترهای مذکور نیز با توجه به کاربرد هر یک بصورت دو یا سه مقداری (که نشان دهنده وجود یا عدم

وجود معیار مورد نظر در وب سایت است) ارزیابی می‌گردد. در سیستم پیشنهادی از روش مبتنی بر قاعده جهت بازنمایی دانایی استفاده شده است. پیاده‌سازی قواعد نیز در قالب جدول تصمیم انجام گردیده است. در این جدول با توجه به مقادیر مختلف پارامترهای ورودی، خروجی هر قسمت تعیین می‌گردد. در تحقیق‌های [۱۳] و [۱۸] میزان وجود هر یک از ۲۷ پارامتر ذکر شده در وب سایت‌های مختلف تعیین گردیده و ترکیب هر یک از حالات مختلف آنها با یکدیگر ارائه شده است. با استفاده از این حالات می‌توان پایگاه دانایی سیستم پیشنهادی را تشکیل داد. در اینجا وزن کلیه پارامترهای ورودی یکسان در نظر گرفته شده است.

جدول (۱): مشخصه های قابل ارزیابی در تشخیص سایت فیشینگ

ردیف	پارامتر	بخش بندی	لایه
۱	استفاده از آدرس IP در آدرس وب سایت	مشخصه های دامنه وب سایت	اول
۲	میزان غیرعادی بودن آدرس درخواست		
۳	میزان وجود اختلال در آدرس دهی پایه		
۴	وجود اطلاعات غیر عادی ثبت شده در سرور		
۵	میزان غیرعادی بودن آدرس وب سایت		
۶	استفاده از گواهینامه امنیت سایت	امنیت و رمزنگاری	دوم
۷	اعتبار داشتن گواهینامه		
۸	وجود فایل های کوکی غیرعادی		
۹	وجود اطلاعات شناسایی گواهینامه		
۱۰	وجود صفحات انتقال	کدهای اسکریپتی موجود در صفحه	
۱۱	وجود حمله استرادلینگ[۱۵]		
۱۲	وجود حمله فارمینگ[۱۶]		
۱۳	مخفی نمودن لینک در زمان حرکت موس		
۱۴	استفاده از فرم هندلر در سرور		
۱۵	اشکالات املایی در محتوا	ظاهر صفحه و محتوای آن	سوم
۱۶	سایت کپی شده از سایت اصلی		
۱۷	استفاده از فرم با دکمه ارسال[۱۷]		
۱۸	استفاده از پنجره پاپ آپ[۱۸]		
۱۹	غیرفعال شدن کلیک راست موس		
۲۰	وجود آدرس وب طولانی	آدرس صفحه وب	
۲۱	جایگزاری کاراکترهای مشابه در آدرس		
۲۲	افزودن پسوند و پیشوند در آدرس		
۲۳	استفاده از کاراکتر @ در آدرس		
۲۴	استفاده از کاراکترهای هگزادسیمال[۱۹]		
۲۵	میزان تاکید بر امنیت	ویژگیهای رفتاری سایت	
۲۶	نوع خوشامدگویی عمومی مورد استفاده		
۲۷	صرف زمان بیش از حد در دریافت اطلاعات		

در لایه‌ی اول یعنی مشخصه‌های دامنه وب سایت، تعداد ۵ پارامتر وجود دارد که هرکدام از این پارامترها شامل سه حالت کم، متوسط و زیاد است. بنابراین تعداد حالات ممکن در لایه‌ی اول برابر 3^5 و یا ۲۴۳ حالت متفاوت خواهد شد. خروجی هر لایه بصورت سه حالت اصلی، مشکوک و تقلبی[۲۰] نشان داده می شود. در سیستم خبره پیشنهادی برای محاسبه خروجی لایه‌ی اول، از شبکه های عصبی مصنوعی (ANN)[۲۱] جهت محاسبه وزن هر یک از پارامترهای ورودی لایه‌ی اول استفاده گردیده است. شکل (۴)، نمونه شبکه عصبی تک لایه با یک نرون و پنج ورودی، جهت ارزیابی مقادیر مختلف این لایه را نشان می‌دهد.



شکل (۴): شبکه عصبی تک لایه با یک نرون به منظور تعیین مقدار لایه‌ی اول

برای محاسبه مقدار وزن ها و مقدار بایاس، لازم است تا شبکه عصبی بر اساس تعدادی قاعده بعنوان حقایق سیستم، آموزش ببینند. در تحقیق [۱۳] تعدادی از قوانین این بخش لیست گردیده که چند نمونه آن در جدول (۲) نشان داده شده است. در اینجا از این قوانین برای آموزش شبکه عصبی فوق استفاده گردیده است.

پس از محاسبه مقدار وزن ها و مقدار بایاس می‌توان قواعد مربوطه را در سیستم بر اساس خروجی طراحی نمود که در این صورت تعداد قوانین به طور چشمگیری کاهش خواهد یافت. برای بخش دوم یعنی پارامترهای مربوط به امنیت و رمزنگاری، برای هر پارامتر دوحالت ورودی در نظر گرفته شده که دراینصورت تعداد حالات ممکن در جدول تصمیم برابر 2^4 یا ۱۶ حالت مختلف خواهد شد.

جدول (۲): نمونه ای از قوانین سیستم به منظور آموزش شبکه عصبی

استفاده از **IP** آدرس	میزان غیرعادی بودن آدرس درخواست	میزان وجود اختلال در آدرس دهی	وجود اطلاعات غیر عادی ثبت شده در سرور	میزان غیرعادی بودن آدرس وب سایت	نتیجه ارزیابی لایه اول
کم	کم	کم	کم	کم	اصلی
کم	کم	کم	متوسط	زیاد	مشکوک
کم	کم	متوسط	متوسط	زیاد	تقلبی
متوسط	متوسط	متوسط	کم	زیاد	تقلبی

خروجی سایر بخش‌ها نیز با توجه به مقادیر ورودی های هر یک در قالب جدول تصمیم ارزیابی می گردد. در مرحله بعدی لازم است طبق جدول (۱)، خروجی هر لایه بر اساس نتیجه هر

بخش مشخص گردد تا بتوان در نهایت با مقایسه حالات مختلف هر لایه با دو لایه‌ی دیگر در خصوص میزان فیشینگ بودن وب سایت، استنتاج نمود. بنابراین لایه‌ی دوم حاصل مقایسه خروجی دو بخش امنیت و رمزنگاری و کدهای اسکریپتی موجود در صفحه می باشد. با توجه به بازه مقادیر ورودی، جدول تصمیم لایه‌ی دوم نیز با ۹ قاعده مطابق جدول (۳) ایجاد می‌گردد.

جدول (۳): جدول تصمیم برای لایه‌ی دوم

ردیف	امنیت و رمزنگاری	کدهای اسکریپتی صفحه	خروج لایه‌ی دوم	شماره قاعده
۱	اصلی	اصلی	قانونی	۱
۲	اصلی	مشکوک	قانونی	
۳	اصلی	تقلبی	نا مطمئن	۲
۴	مشکوک	اصلی	نا مطمئن	۳
۵	مشکوک	مشکوک	نا مطمئن	
۶	مشکوک	تقلبی	نا مطمئن	
۷	تقلبی	اصلی	نا مطمئن	۴
۸	تقلبی	مشکوک	تقلبی	۵
۹	تقلبی	تقلبی	تقلبی	

با توجه به مقدار خروجی جداول تصمیم سیستم خبره، می توان طی یک فرآیند نرمال‌سازی، سطرهایی از جدول که دارای خروجی مشابه هستند را با توجه به مقادیر خاص ظاهر شده در ورودی، کاهش داد. بعنوان مثال در جدول (۳)، خروجی لایه‌ی دوم به ازاء مقادیر موجود در سطرهای ۴، ۵ و ۶ بعنوان نامطمئن در نظر گرفته شده است، لذا این سه قاعده می‌تواند در قالب یک قاعده بدین صورت بیان گردد: "*اگر نتیجه بخش امنیت و رمزنگاری، مشکوک بود خروجی لایه‌ی دوم نامطمئن است.*" بنابراین با اجرای این فرآیند بر روی تمامی جداول تصمیم، تعداد قواعد سیستم خبره کاهش خواهد یافت. برای بخش دوم تعداد قواعد از ۱۶ مورد به ۴ مورد، برای بخش سوم (کدهای اسکریپتی موجود در صفحه) تعداد قواعد از ۳۲ مورد به ۱۳ مورد، برای بخش چهارم (ظاهر صفحه و محتوای آن) تعداد قواعد از ۳۲ مورد به ۱۱ مورد، برای بخش پنجم (آدرس صفحه وب) تعداد قواعد از ۳۲ مورد به ۸ مورد و برای بخشش ششم نیز تعداد قواعد از ۸ مورد به ۵ مورد کاهش خواهد یافت. به همین ترتیب پس از انجام فرآیند نرمان سازی برای لایه ها، تعداد قواعد لایه‌ی دوم از ۹ مورد به ۵ مورد و لایه‌ی سوم نیز از ۲۷ مورد به ۱۴ مورد کاهش می یابد. استنتاج نهایی نیز بر اساس خروجی سه لایه، نشان دهنده میزان فیشینگ بودن وب سایت مورد آزمایش می باشد. قوانین لازم برای این سطح نیز طبق خروجی‌های حاصل از سه لایه ارزیابی سیستم پس از طی فرآیند نرمال سازی تعداد ۲۰ قانون خواهد شد.

۴-۲- موتور استنتاج

موتور استنتاج با استفاده از قواعد منطق و دانش و حقایق موجود در پایگاه دانایی در خصوص مسائل مطرح شده، استنتاج می‌نماید. هنگامی که قواعد توسط موتور استنتاج مورد بررسی قرار می‌گیرند، دستورات لازم در صورتی که اطلاعات ارائه شده توسط کاربر توسط شرایط موجود در قوانین ارضاء شود، اجرا خواهند شد [۲۲]. در سیستم پیشنهاد شده، فرآیند استنتاج با استفاده از موتور استنتاج موجود در پوسته سیستم خبره برای هر کدام از بخش‌ها بصورت مجزا انجام می‌گیرد. درنهایت نیز نتیجه هر بخش در مقایسه با نتیجه سایر بخش‌ها ارزیابی و خروجی بعنوان استنتاج نهایی سیستم ارائه خواهد گردید.

۴-۳- پیاده سازی سیستم

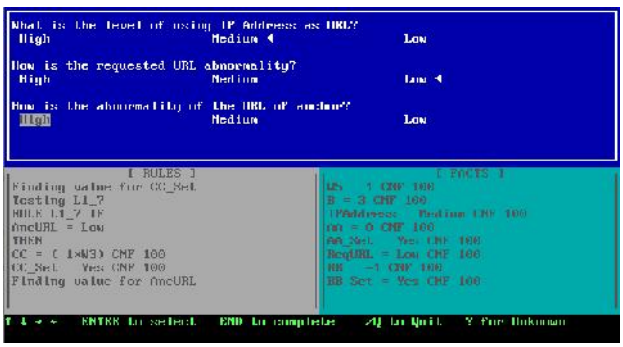
سیستم‌های خبره را می‌توان با استفاده زبانهابی مختلف برنامه نویسی و یا با استفاده از محصولاتی با نام *پوسته سیستم خبره*[۲۲] ، پیاده سازی نمود. پوسته سیستم خبره برای توسعه یک سیستم عملی بسیار مفید است. معمولا، یک پوسته از موتور استنتاج، بخش ارائه دانش، امکانات رابط کاربری و ویرایشگر دانش تشکیل شده است [۲۲]. به منظور طراحی سیستم خبره پیشنهادی از پوسته خبره VP-Expert استفاده شده است. مقادیر مورد بررسی معیارهای استنتاج سیستم در قالب قواعد سیستم بصورت "اگر.....آنگاه....."[۲۳] پیاده‌سازی شدند. مقادیر وزن ها و بایاس شبکه عصبی مصنوعی مربوط به بخش اول، بصورت مقادیر ثابت در پوسته VP-Expert درنظر گرفته شد. سپس با استفاده از دستورات سیستم خبره همانند سایر قواعد پیاده سازی گردید. در مجموع این سیستم خبره با تعداد ۹۹ قاعده طبق جدول (۴) پیاده سازی شد.

جدول (۴) : تعداد قواعد بخش‌های مختلف سیستم خبره پیشنهادی

بخش / لایه	تعداد قاعده
بخش اول	۱۹
بخش دوم	۴
بخش سوم	۱۳
بخش چهارم	۱۱
بخش پنجم	۸
بخش ششم	۵
لایه دوم	۵
لایه سوم	۱۴
استنتاج نهایی	۲۰
کلیه قواعد سیستم خبره	**۹۹**

۴-۴- ارتباط با کاربر

در بخش ارتباط با کاربر، سیستم خبره سؤالاتی را در خصوص هر یک از پارامترهای ورودی، مطرح می‌کند و کاربر باید پاسخ مناسب را وارد نماید تا سیستم بتواند با توجه به پاسخ های داده شده استنتاج لازم را انجام دهد. در شکل (۵)، صفحه ارتباط با کاربر سیستم خبره نشان داده شده است. عمل استنتاج بر اساس قواعد تنظیم شده در سیستم انجام می‌گیرد. نتیجه هر بخش بصورت مجزا بر اساس مقادیر ورودی محاسبه می‌گردد. خروجی هر لایه نیز طبق قواعد پیاده‌سازی شده در سیستم خبره محاسبه گردیده و درنهایت نتیجه استنتاج در خروجی سیستم بصورت یک عبارت شامل *کاملاً قانونی* (Very-Legitimate)، *قانونی* (Legitimate)، *مشکوک* (Suspicious)، *فیشینگ* (Phishy) و *کاملاً فیشینگ* (Very-Phishy)، نشان داده می‌شود. هر یک از این عبارات نشان دهنده میزان فیشینگ بودن وب سایت مورد نظر است. دقت خروجی سیستم خبره به نحوه‌ی جمع‌آوری مقادیر پارامترهای ورودی نیز بستگی فراوانی دارد. در بین سه لایه پارامترهای ورودی، لایه‌ی دوم که شامل امنیت و رمزنگاری و کدهای اسکریپتی موجود در صفحه می باشد از اهمیت بیشتری برخوردار است.



شکل (۵): محیط ارتباط با کاربر سیستم خبره

۵- اعتبار سنجی سیستم خبره پیشنهادی

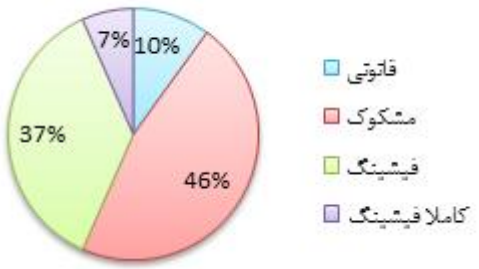
جهت بررسی اعتبار و کارایی سیستم خبره، خروجی سیستم بر اساس مقادیر واقعی کنترل می‌گردد. برای این منظور از اطلاعات موجود سایت PhishTank استفاده شده است. سایت PhishTank یکی از سرویس های معروف مربوط به OpenDNS می باشد که در زمینه ثبت و نگهداری اطلاعات سایتهای فیشینگ فعالیت می‌نماید. برای بررسی میزان اعتبار سیستم تعداد ۱۵۰ صفحه فیشینگ از بانک اطلاعات سایت مذکور استخراج گردید. این صفحات تماماً مربوط به سایت های بانکداری الکترونیکی یا تجارت الکترونیکی بوده و سعی شده تا فرآینده ارزیابی بصورت کامل بر روی صفحات فیشینگ مربوطه انجام پذیرد. لیست برخی از این سایتها در جدول (۵) نشان داده شده است.

به دلیل اینکه بطور متوسط وبسایت‌های فیشینگ برای چند ساعت فعال هستند لذا امکان استخراج داده‌های لازم برای بعضی از پارامترهای ورودی امکان پذیر نبود که در این حالت مقادیر آنها بصورت پیشفرض کمترین مقدار در نظر گرفته شد. برای آزمایش داده‌های استخراج شده در سیستم خبره، این داده‌ها بصورت یک پایگاه داده برای سیستم معرفی گردید تا فرآیند ارزیابی با سرعت و دقت بیشتری انجام پذیرد.

جدول (۵) : لیست برخی از سایتهای مورد هدف حملات فیشینگ

ردیف	سایت هدف	تعداد
۱	Bank of America Corporation	۱۴
۲	Royal Bank of Canada	۲
۳	Co-operative Bank	۱۰
۴	NatWest Bank	۳۶
۵	Banco De Brasil	۲۴
۶	Capitec Bank	۵
۷	Allied Bank Limited	۳
۸	Royal Bank of Scotland	۵
۹	Barclays Bank PLC	۵

طبق تقسیم‌بندی تعیین شده برای نمایش نتیجه خروجی که در بخش (۴-۴) بدان اشاره شد، از تعداد ۱۵۰ صفحه فیشینگ مورد ارزیابی، تعداد ۸۲ صفحه بعنوان صفحات مشکوک، ۵۵ صفحه بعنوان فیشینگ و ۱۰ صفحه نیز بصورت کاملاً فیشینگ ارزیابی گردید. نتیجه این ارزیابی در شکل (۶) نشان داده شده است.



شکل (۶) : نتیجه ارزیابی سیستم از ۱۵۰ سایت فیشینگ

سیستم خبره پیشنهادی با نرخ خطای حدودی ۱۰٪ قادر به شناسایی صفحات فیشینگ است. جهت ارزیابی بیشتر سیستم، طبق روش ارائه شده در تحقیق [۱۳]، سیستم پیشنهادی در حالات خاصی مورد بررسی قرار گرفت. در نمونه اول مطابق جدول (۶)، سعی شد تا فقط پارامترهای لایه‌ی اول برای سیستم تعیین گردد و سایر پارامترها نادیده گرفته شود. نتیجه این مقادیر بصورت مشکوک ارزیابی شده که نشان از ارزیابی دقیق سیستم دارد. این در حالی است که اگر مقادیر این پارامترهای بجای متوسط، کم در نظر گرفته شود خروجی سیستم نشان دهنده قانونی بودن وب سایت خواهد بود.

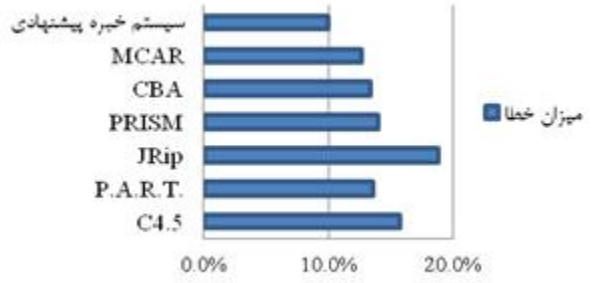
در نمونه دوم نیز، فقط مقادیر مورد نظر برای لایه‌ی دوم در نظر گرفته شد و سایر پارامترها بصورت عادی فرض شدند. با این وجود، خروجی سیستم نیز مجدداً مشکوک در نظر گرفته شد که در اینصورت مطمئناً در نمونه های واقعی نتیجه درستی را ارائه نموده و کاربر را نیز وادار به تأمل بیشتری در خصوص وب سایت مورد نظر خواهد نمود.

طبق تحقیق [۲۰]، در ارزیابی صورت گرفته با استفاده از الگوریتم‌های مختلف داده‌کاوی جهت دسته‌بندی صفحات، الگوریتم CBA [۲۴] با نرخ خطای ۱۳/۴% دارای کمترین خطا در ارزیابی صفحات فیشینگ با استفاده از ۲۷ ویژگی ذکر شده است [۲۰].

جدول (۶): مقادیر آزمایشی پارامترها برای ارزیابی سیستم خبره

پارامتر	مقدار
استفاده از آدرس IP در آدرس دهی وب سایت	متوسط
میزان غیرعادی بودن آدرس درخواست	متوسط
میزان وجود اختلال در آدرس دهی پایه	متوسط
وجود اطلاعات غیر عادی ثبت شده در سرور	متوسط
میزان غیرعادی بودن آدرس وب سایت	متوسط
سایر پارامترها	-
نتیجه ارزیابی	مشکوک

خروجی سیستم خبره پیشنهادی در این تحقیق در مقایسه با سایر الگوریتم ها در شکل (۷) نشان داده شده است.



شکل (۷) : مقایسه میزان خطای سیستم خبره پیشنهادی با سایر روش ها

٦- نتیجه گیری

در این تحقیق بر اساس معیارهای ارائه شده در مقالات مرتبط سیستم خبره‌ای به منظور تشخیص حملات فیشینگ در بانکداری الکترونیکی ارائه شد. این معیارها در قالب ۲۷ پارامتر و در شش بخش مختلف شامل مشخصه‌های دامنه وب سایت، قابلیت‌های امنیت و رمزنگاری مورد استفاده، کدهای اسکریپتی موجود در صفحه، ظاهر صفحه و محتوای آن، آدرس صفحه وب و ویژگیهای رفتاری، تعیین شده‌اند. با ترکیب مقادیر مختلف قابل استخراج برای هریک از این خصوصیات، پایگاه دانایی سیستم خبره تشکیل گردید. به منظور کاهش تعداد قواعد و همچنین افزایش سرعت استنتاج، بخشی از قوانین موجود در پایگاه دانایی از طریق شبکه‌های عصبی مصنوعی ارائه گردید. جهت پیاده‌سازی سیستم خبره پیشنهادی نیز از پوسته VP-Expert استفاده شد. ارزیابی سیستم خبره پیشنهادی از طریق اطلاعات جمع‌آوری شده از سایت PhishTank، نشان دهنده بهبود نسبی این سیستم در تشخیص و شناسایی حملات فیشینگ با میزان خطای ۱۰٪ است.

به منظور توسعه سیستم پیشنهادی، می‌توان از سایر روشهای یادگیری ماشین جهت تعیین وزن کلیه پارامترهای ورودی استفاده نمود که با توجه به وجود قابلیت یادگیری برای سیستم خبره در این روشها، کارایی و سرعت استنتاج، بسیار افزایش خواهد یافت.

مراجع

[1] Vrîncianu, & Popa, , 2010. Considerations Regarding the Security and Protection of E-Banking Services Consumers' Interests. Amfiteatru Economic, 12(28), pp.388-403.

properties. In In NYS Cyber Security Conference., 2006.

[17] Chen, J. & Ch., G., 2006. Online Detection and Prevention of Phishing Attacks. In First International Conference on Communications and Networking. China, 2006.

[18] Aburrous, , Hossain, M.A. & Dahal, , 2010. Experimental Case Studies for Investigating E-Banking Phishing Techniques and Attack Strategies. Cogn Comput, Springer Science, pp.242-53.

[19] Damodaram, R. & Valarmathi, M.L., 2011. Phishing Website Detection and Optimization Using Particle Swarm Optimization Technique. International Journal of Computer Science and Security (IJCSS), 5(5), pp.477-90.

[20] Sumathi, R. & Vidhya Prakash, R., 2012. Prediction of Phishing Websites Using Optimization Techniques. International Journal of Modern Engineering Research (IJMER), 2(1), pp.341-48.

[21] APWG, 2011. Phishing Activity Trends Report 2nd Half 2011. Anti-Phishing Working Group (APWG).

[22] Waterman, D.A., 1986. A guide to expert systems, illustrated, reprinted. California, USA: Addison-Wesley.

زیرنویس‌ها

[1] E-Commerce
[2] Electronic Banking (E-Banking)
[3] Phishing Attacks
[4] Anti-Phishing Working Group (APWG)
[5] Plugin
[6] Internet Explorer
[7] AntiPhish Tool
[8] Supervised Machine Learning
[9] Document Object Model (DOM)
[10] Simulated Annealing (SA)
[11] Support Vector Machines (SVM)
[12] Knowledge Base
[13] Inference Engine
[14] User Interface
[15] Straddling Attack
[16] Pharming Attack
[17] Submit
[18] Pop-up window
[19] Hexadecimal
[20] Genuine, Doubtful, Fraud
[21] Artificial Neural Network (ANN)
[22] Expert System Shell
[23] IF….THEN…. Rules
[24] Classification based on association rules

[2] Chou, N. et al., 2004. Client–side defense against web–based identity theft. In In Proceedings of 11th Annual Network and Distributed System Security Symposium., 2004.

[3] Pan, Y. & Ding, X., 2006. Anomaly based web phishing page detection. In 23th annual computer security applications conference., 2006.

[4] Al-Fairuz, & Renaud, , 2010. Multi-channel, Multi-level Authentication for More Secure eBanking. ISSA

[5] Weir, C.S., Douglas, , Richardson, & Jack, , 2010. Usable security: User preferences for authentication methods in eBanking and the effects of experience. Interacting with Computers, (22), pp.153-64.

[6] Yang, Y.-J., 1997. The Security of Electronic Banking. In International Systems Security Conference. National Computer Security Center., 1997., pp. 41-52.

[7] Wu, M., Miller, R. & Little, G., 2006. Web wallet: preventing phishing attacks by revealing user intentions. MIT Computer Science and Artificial Intelligence Lab.

[8] Jain, A. & Richariya, V., 2011. Implementing a Web Browser with Phishing Detection Techniques. World of Computer Science and Information Technology Journal (WCSIT), 1(7), pp.289-91.

[9] Abu-Nimeh, S., Nappa, D., Wang, X. & Nair, S., 2007. A Comparison of Machine Learning Techniques for Phishing Detection. In In Proceedings of the Anti-Phishing Working Group eCrime Researchers Summit., 2007.

[10] Alkhozae, M.G. & Batarfi, O.A., 2011. Phishing Websites Detection based on Phishing Characteristics in the Webpage Source Code. International Journal of Information and Communication Technology Research, 1(9), pp.238-91.

[11] Sharif, T., 2006. Phishing filter in IE7. [Online] Available at: http://www.blogs.msdn.com/ie/archive/2005/09/09/463204.aspx.

[12] Kirda, E. & Kruegel, C., 2005. Protecting users against phishing attacks. The Computer Journal.

[13] Aburrous, M., Hossain, M.A., Dahal, K. & Thabatah, F., 2010. Intelligent phishing detection system for e-banking using fuzzy data mining. Expert Systems with Applications, (37), p.7913–7921.

[14] Fotiou, N., Marias, G.F. & Polyzos, , 2012. Fighting Phishing the Information-Centric Way. In 5th International Conference on New Technologies, Mobility and Security (NTMS)., 2012., pp. 1-5.

[15] Maurer et al., "Using Visual Website Similarity for Phishing Detection and Reporting," in Proceedings of the 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts, 2012.

[16] Chandrasekaran, M., Narayanan, K. & Upadhyaya, S., 2006. Phishing email detection based on structural