فصلنامه پژوهش در علوم، مهندسی و فناوری
دوره ۸، شماره ۱، بهار ۱۴۰۱
صفحات ۹۱-۱۰۶
ISSN: 2476-5082
www.uctjournals.com

ارائه یک سیستم هوشمند جهت تشخیص نفوذ در شبکه های کامپیوتری مبتنی بر فازی

جواد رفیعی۱(نویسنده مسئول)، دکتر عادل جهانبانی۲، دکتر جعفر پرتابیان
۱.دانشجوگروه مهندسی کامپیوتر ، واحد لامرد، دانشگاه آزاد اسلامی، لامرد، ایران javadrafei@gmail.com
۲. استادیار گروه مهندسی کامپیوتر ، واحد لامرد، دانشگاه آزاد اسلامی، لامرد، ایران jahanbani.adel@iau.ac.ir
۳. استادیار گروه مهندسی کامپیوتر ، واحد لامرد، دانشگاه آزاد اسلامی، لامرد، ایرانjavadrafei@gmail.com

چکیده

فرآیند نظارت و تحلیل وقایع روی یک سیستم کامپیوتری یا شبکه به منظور شناسایی حملات و کشف اشکالات امنیتی آن را تشخیص نفوذ گویند با توجه به مطالعاتی که در زمینه‌ی سیستم‌های تشخیص نفوذ صورت گرفته و همچنین با توجه به تحقیقاتی که در زمینه الگوریتم‌های فراابتکاری بر روی سیستم‌های تشخیص نفوذ انجام گرفته است، علی‌رغم بهبود کارایی سیستم‌های تشخیص نفوذ با استفاده از انتخاب بهترین و مؤثرترین ویژگی‌ها از بین مجموعه ویژگی‌ها، سرعت اجرای الگوریتم‌ها را کاهش دهیم. با توجه به مطالعات صورت گرفته، الگوریتم فاخته در ترکیب با فازی تاکنون در زمینه سیستم-های تشخیص نفوذ به کار گرفته نشده است. در این مقاله سعی بر آن داریم که تکنیک های تشخیص نفوذ را شناسایی، متدولوژی فنی مورد نیاز برای پاسخگویی گسترده به بیان مساله را (روش پیشنهادی) بپردازیم آن را تجزیه و تحلیل و نتایج یافته شده را ارائه دهیم.

کلیدواژه: شبکه های کامپیوتری،تشخیص نفوذ،الگوریتم فازی،سیستم هوشمند

۱-مقدمه

تشخیص نفوذ، ساز و کارهای پیشگیرانه را برای بهبود امنیت سیستم تکمیل می‌کند. علاوه بر این، حتی اگر ساز و کارهای امنیتی پیشگیرانه بتوانند سیستم‌های اطلاعات را با موفقیت حفاظت کنند، این روند کماکان برای شناخت اینکه چه نفوذهایی روی داده یا در حال وقوع است مطلوب خواهد بود، به گونه‌ای که می‌توانیم خطرات و تهدیدهای امنیتی را دریافته و بنابراین برای خطرات آتی آمادگی بهتری داشته باشیم.به رغم اهمیت آنها، IDS ها جایگزین‌هایی برای ساز و کارهای امنیتی پیشگیرانه مانند کنترل دسترسی و تایید هویت محسوب نمی‌شوند. در واقع، خود IDS ها نمی‌توانند حفاظت کافی برای سیستم‌های اطلاعات ارائه کنند. برای نمونه، چنانچه یک مهاجم تمام داده‌ها را در یک سیستم اطلاعاتی پاک کند، شناسایی حمله‌ها نمی‌تواند خسارت را به طور کلی کاهش دهد. بنابراین، IDS ها باید همراه با ساز و کارهای امنیتی پیشگیرانه به عنوان بخشی از سیستم دفاعی جامع گسترش یابد.تکنیک‌های تشخیص نفوذ به طور معمول در یکی از این دو دسته جای می‌گیرد: تشخیص ناهنجاری و تشخیص سوء استفاده. تشخیص ناهنجاری بر رفتار عادی یک موضوع هر اقدامی که به طور بارز از رفتار عادی منحرف شده باشد مزاحم تلقی می‌گردد. تشخیص سوء استفاده نفوذها را برحسب ویژگی‌های حملات شناخته شده یا آسیب‌پذیری‌های سیستم به دست می‌آورد؛ هر اقدامی که با الگوی یک حمله شناخته شده یا آسیب پذیری مطابقت داشته باشد مزاحم تلقی می‌شود.

همچنین، IDS ها می‌توانند براساس منابع اطلاعات حسابری استفاده شده از سوی هریک از IDS ها به IDS های مبتنی بر میزبان، IDS های توزیع شده و IDS های توزیعی طبقه بندی شوند. IDS های مبتنی بر میزبان داده‌های حسابرسی را از مسیرهای حسابرسی میزبان گرفته و به طور معمول هدف آنها شناسایی حملات در برابر یک میزبان واحد می‌باشد؛ IDS های توزیعی داده‌های حسابرسی را از میزبان‌های متعدد و احتمالا شبکه که به میزبان‌ها متصل می‌شود جمع‌آوری کرده و حملات مشتمل بر میزبانهای چندگانه را مورد شناسایی قرار می‌دهد. IDS های مبتنی بر شبکه ترافیک شبکه را به عنوان منبع داده حسابرسی با کاستن بار حاصل بر میزبان‌ها مورد استفاده قرار دهند که به طور معمول خدمات محاسباتی معمولی را ارائه می‌کنند.

امروزه با رشد و گسترش اینترنت در فعالیت‌های تجاری و سرویس‌های زیرساختی، کاربردهای شبکه‌های رایانه‌ای در زمینه‌های مختلف همچون انتقال داده‌ها، سرویس‌های وب، معاملات الکترونیک، پخش صدا و تصویر و غیره گسترده‌تر شده است. سازمان‌ها عموماً دارای شبکه‌های اطلاعاتی پیچیده‌ای می‌باشند و این شبکه‌ها را به‌منظور اشتراک گذاری اطلاعات با شرکا و مشتریانشان باز گذاشته‌اند. از این رو شبکه‌های رایانه‌ای ناگزیر در معرض حملات سایبری رو به گسترش هستند و این حملات می‌توانند باعث بروز میلیون‌ها دلار خسارت مالی به سازمان‌ها شوند. درنتیجه امروزه ضرورت حفاظت و امنیت اطلاعات و مقابله با تهدیدات امنیتی بیش از پیش مورد توجه قرار گرفته است. در این زمینه، تشخیص نفوذ در شبکه‌های رایانه‌ای توجه پژوهشگران را به خود جلب کرده و ساخت سیستم تشخیص نفوذ کارآمد و قدرتمند به عنوان امری با بالاترین اولویت در محیط‌های دانشگاهی، ارگان‌های دولتی، مؤسسات تحقیقاتی و شرکت‌های صنعتی درنظر گرفته شده است.

فرآیند نظارت و تحلیل وقایع روی یک سیستم کامپیوتری یا شبکه به منظور شناسایی حملات و کشف اشکالات امنیتی آن را تشخیص نفوذ گویند. درمباحث امنیت سیستم‌ها و شبکه‌های کامپیوتری منظور ازتشخیص نفوذ تشخیص آن دسته ازحملات و نفوذهایی است که با استفاده ازمکانیزم‌های معمول پیش‌گیرانه ازجمله روش‌های هویت شناسی و اعتبارسنجی کنترل دسترسی حفاظ و رمزنگاری امکان پیش‌گیری ازبروز آن‌ها وجود ندارد. از جمله روش‌های موجود درتشخیص نفوذ، استفاده از تکنیک‌های داده‌کاوی است. داده‌کاوی فرایند کشف مدل‌های مختلف، خلاصه‌ها و مقادیر کسب شده از مجموعه داده می‌باشد. در این تحقیق به ارائه یک الگوریتم بر مبنای سیستم فازی و الگوریتم فاخته پرداخته‌ایم.

۱-۱بیان مسأله

اخیرا و با افزایش نیاز به استفاده از اینترنت در برنامه‌های کاربردی و حوزه‌ها تعداد بسته‌های متحرک و بار بر روی شبکه افزایش یافته است. بنابراین، مهمترین اطلاعات علی‌رغم وجود چندین سیستم محافظت از شبکه مانند سیستم فایروال که یک

سیستم مؤثر حفاظت و پیشگیری است، در معرض خطر است. از آنجا که امروزه تکنولوژی‌های مبتنی بر شبکه‌ها همه جا رایج شده‌اند. یکی از روش‌های مؤثر برای به دست آوردن امنیت بالاتر، استفاده از سیستم‌های تشخیص نفوذ می‌باشد که ابزارهای نرم‌افزاری برای تشخیص فعالیت‌های غیرنرمال در کامپیوتر و شبکه هستند. از جمله تهدیدهایی که برای یک سیستم کامپیوتری وجود دارند ویروس‌ها و نفوذها هستند. ویروس‌ها می‌توانند به طور گسترده با نصب نرم‌افزار آنتی ویروس و بروز رسانی به طور منظم کنترل شوند .

از مسائل مهم در تعیین صحت عملکرد سیستم‌های تشخیص نفوذ، مسئله هشدارهای غلط است. این هشدارها در دو دسته هشدارهای غلط-مثبت و هشدارهای غلط-منفی قرار می‌گیرند. هشدار غلط-مثبت زمانی رخ می‌دهد که سیستم به اشتباه یک فعالیت مجاز را فعالیتی نفوذی تشخیص دهد و در مورد دیگر، هشدار غلط-منفی، سیستم قادر به تشخیص رفتار نفوذی نخواهد بود. با توجه به نوع محیطی که سیستم به حفاظت از آن می‌پردازد، بهبود عملکرد سیستم با استفاده از کاهش یکی از این دسته هشدارهای غلط صورت می‌پذیرد. سیستم تشخیص نفوذ دارای سه عملکرد مشاهده، تشخیص و پاسخ دهی (هشدار) به فعالیت‌های بدون اجازه است و هدف آن، انجام عمل تشخیص نفوذ با سرعت مناسب و با کمترین تعداد هشدار نادرست است. سیستم تشخیص نفوذ فرآیندی از مشاهده و تحلیل رویدادهای پدید آمده در شبکه است. پایین بودن نرخ دقت تشخیص نفوذ و بالا بودن نرخ تشخیص اشتباه مثبت از چالش‌های اساسی این سیستم به شمار می‌آید. هدف این پژوهش بهبود چالش‌های اشاره شده در سیستم تشخیص نفوذ است، و در همین راستا روشی را با استفاده از تکنیک انتخاب ویژگی مطرح می‌کند. یکی از این الگوریتم‌ها الگوریتم فرااابتکاری فاخته می‌باشد. الگوریتم بهینه سازی فاخته[1]نیز یکی از پیاده‌سازی‌های موفق از فرآیندهای طبیعی است. این الگوریتم از شیوه زندگی پرنده‌ای به نام فاخته الهام گرفته شده است. شیوه زندگی خاص این پرنده، روش تخم‌گذاری و رشد منحصر به فرد و در نهایت تولید مثل این گونه، پایه و اساس این الگوریتم بهینه‌سازی را تشکیل می‌دهد. ویژگی شاخص این الگوریتم، شبیه‌سازی مفهوم بقا، مهاجرت برای یافتن منابع غذایی و انتخاب محیط بهینه برای زندگی است. جمعیت الگوریتم بهینه سازی فاخته را فاخته‌های بالغ و تخم‌های فاخته تشکیل می‌دهند. اثربخشی روش پیشنهادی بهینه‌ساز فاخته، از طریق مقایسه با سایر تکنیک‌های الهام گرفته شده از طبیعت، بر روی ۲۹ تابع، محک و چندین مسئله مهندسی دنیای واقعی بررسی شده است. نتایج آماری و مقایسه‌ها نشان می‌دهد که الگوریتم فاخته نتایج بسیار امیدوار کننده و گاها رقابتی را در مقایسه با دیگر تکنیک‌های فرااابتکاری شناخته شده دارد. در این پژوهش از تکنیک فازی در ترکیب با الگوریتم فاخته برای مجموعه داده KDDCUP99 استفاده می‌شود.

۱-۲اهمیت و ضرورت انجام تحقیق

اینترنت ساختارهای اجتماعی، سیاسی و اقتصادی را به طور مثبت تغییر داده و از بسیاری جهات مرزهای جغرافیایی را از بین برده است. سهم عظیم اینترنت در معاملات تجاری همراه با سهولت استفاده از آن منجر به افزایش تعداد کاربران اینترنت و در نتیجه افراد متجاوز شده است. حفاظت از منابع رایانه ای با کمک سیستم های تشخیص نفوذ[2] علاوه بر سیستم های جلوگیری از نفوذ بسیار مهم است. در زمان های اخیر، تجزیه و تحلیل ترافیک عظیم شبکه ای که در چند ثانیه در ترابایت ایجاد می شود با روش مبتنی بر قاعده سنتی دشوار است. از این رو، محققان باید تکنیک های الگوریتم های تکاملی را با تأکید بر دقت تشخیص نفوذ، در معرض تشخیص نفوذ قرار دهند. انتخاب ویژگی مربوط منجر به سریعتر و افزایش سرعت تشخیص دقیق می شود. بنابراین ، این تحقیق یک IDS را بر اساس ترکیب الگوریتم فاخته با فازی ارائه می دهد.

۱-۳اهداف مشخص تحقیق

توسعه سیستم تشخیص نفوذ در شبکه‌های کامپیوتری از طریق انتخاب بهترین و مؤثرترین ویژگی‌ها از میان مجموعه ویژگی‌ها با استفاده از تکنیک فازی در ترکیب با الگوریتم فاخته می‌باشد. در این روش چند هدف را دنبال می‌کنیم:

۱- افزایش نرخ تشخیص

[1] Cuckoo Optimization Algorithm
[2](IDS)

۲- کاهش نرخ مثبت کاذب
۳- افزایش دقت
۴- کاهش سرعت زمان اجرا

۱-۴سؤالات تحقیق:

تکنیک فازی در ترکیب با الگوریتم فاخته چگونه می تواند باعث بهبود و کارایی سیستم‌های تشخیص نفوذ شود؟

۱-۵فرضیه‌های تحقیق:

استفاده از تکنیک فازی در ترکیب با الگوریتم فاخته به منظور انتخاب ویژگی در سیستم‌های تشخیص نفوذ می‌تواند باعث بهبود دقت سیستم شود.

استفاده از تکنیک فازی در ترکیب با الگوریتم فاخته در انتخاب ویژگی می‌تواند با انتخاب ویژگی‌های بهتر، باعث بالا رفتن سرعت اجرای پردازش در سیستم شود.

۲-تشخیص ناهنجاری

۲-۱مدل‌های آماری

مدل‌سازی آماری نخستین روش‌های مورد استفاده برای تشخیص نفوذها در سیستم‌های اطلاعاتی الکترونیک محسوب می‌شود. مفروض است که رفتار یک نفوذکننده به طور قابل توجهی متفاوت از رفتار کاربر عادی است و مدل‌های آماری برای گردآوری رفتار کاربر و تشخیص یک مهاجم از یک کاربر عادی مورد استفاده قرار می‌گیرد. این تکنیک‌ها برای سایر موضوعات مانند گروههای کاربران و برنامه‌ها نیز اجرایی است. در اینجا ما دو الگوی آماری NIDES/STAT و های‌استاک[3]را که برای تشخیص ناهنجاری مطرح شده است بیان می‌کنیم.

۲-۲رویکرد ایمنی شناسی رایانه

رویکرد ایمنی‌شناسی رایانه مبتنی بر مقایسه قابلیت سیستم ایمنی تعیین خود از دیگری است. این رویکرد خود را در قالب مجموعه‌ای از زنجیره l طولی نشان می‌دهد که l پارامتر یک سیستم گسترده است. زنجیره l طولی به صورت دیگری در نظر گرفته می‌شود مشروط بر آنکه هیچ گونه زنجیره‌ای متعلق به خود در آن هماهنگ نباشد. برای ایجاد تعیین کننده‌هایی که بتوانند دیگری را از خود تشخیص دهند یک رویکرد ساده برای تولید اتفاقی زنجیره l طولی و کنترل هماهنگی آن در زنجیره خودی وجود دارد. در صورتی که پاسخ مثبت است زنجیره تولید شده حذف می‌شود در غیراینصورت در قالب یک تعیین کننده مورد استفاده قرار می‌گیرد. با این حال، این رویکرد ساده در تعدادی از زنجیره‌های خودی وقت‌گیر است.

۲-۳شیوه‌های مبتنی بر ویژگی

کو[4] راشیتزکا[5]و کی لویت[6](۱۹۹۷) شیو مبتنی بر ویژگی را برای تشخیص نفوذ پینشهاد کردند. عقیده آنها استفاده از مسیرهای سفارش شده پیامدهای وقایع اجرایی بود تا بتوانند رفتارهای مورد انتظار برنامه‌های جاری در سیستم توزیع را مشخص کنند. یک ویژگی پیامدهای عملیات مورد تایید اجرای یک یا چند برنامه موسوم به را تشریح می‌کند. زنجیره عملیات اجرا شده توسط این موضوع که با ویژگی مورد نظر همخوانی ندارد در قالب تخطی از امنیت در نظر گرفته می‌شود. هر ویژگی به نام یک خط مشی مسیریابی نامیده می‌شود. یک گرامر موسوم به گرامرهای محیط موازی برای تعیین سیاست‌های مسیر توسعه یافتند.

مزیت این رویکرد از لحاظ تئوری آن است که می‌تواند بعضی از انواع جدید حملاتی را مشخص کند که نفوذکننده‌ها در آینده ابداع می‌کنند. نقطه ضعف آن سیستم آن است که عملکرد اصلی آن برای تعیین دقیق رفتار بسیاری از برنامه‌های سیستم مورد نظر ضروری بوده و این ویژگی‌ها، ویژگی سیستم عملیاتی می‌باشند.

[3] Haystack
[4] Ko
[5] Ruschitzka
[6] K. Levitt

۲-۴محدودیت تعیین ناهنجاری

اگرچه تعیین ناهنجاری می‌تواند الگوهای ناشناخته حملات را در برگیرد ممکن است از بعضی از نقاط ضعف نیز رنج ببرد. مشکل متداول تمام رویکردهای تعیین ناهنجاری به استثنای رویکرد مبتنی بر ویژگی‌ها آن است که رفتار نرمال شخص براساس داده‌های گردآوری شده در خلال عملکرد طبیعی مدل‌بندی می‌شود.

۲-۵تعیین سوء استفاده

تعیین سوء استفاده در تعیین ناهنجاری به عنوان امری مقدماتی در نظر گرفته می‌شود. منطق آن این است که الگوهای حملات مشخص را می‌توان به شیوه‌ای موثرتر و کارآمدتر با استفاده از دانش ضمنی آنها تشخیص داد. لذا سیستم‌های تعیین سوء استفاده به دنبال الگوهای مشخص حملات ناشناخته یا نقاط ضعف آنها می‌باشند. آنها می‌توانند فعالیت تشخیصی را مخفی کنند حتی اگر به اندازه‌ای ناچیز باشندکه رویکردهای تعیین ناهنجاری آنها را نادیده بگیرد.

مسئله اصلی در تعیین سوء استفاده چگونگی نشان دادن الگوهای مشخص حملات است. الگوریتم‌های تشخیص معمولاً مستقیماً از مکانیسم‌های نمایش پیروی می‌کند. در این بخش، در مورد شیوه‌های متداول نمایش حملات بحث می‌کنیم.

۲-۶ زبان‌های قانون محور

سیستم متخصص زبان محور، گسترده‌ترین رویکرد مورد استفاده در تعیین سوء استفاده است. الگوی حملات مشخص در قالب مجموعه قوانینی مشخص شده و سیستم متخصص زنجیره رو به جلو معمولاً برای بررسی علائم نفوذ مورد استفاده قرار می‌گیرد. در اینجا در مورد دو زبان قانون محور یعنی زبان ارزیابی زنجیره قانون محور[۷] و مجموعه ابزار سیستم تخصصی تولید محور[۸] بحث می‌کنیم. زبان‌های قانون محور دیگری نیز وجود دارند اما همگی آنها دارای این احساس مشترک هستند که الگوهای حملات مشخص را در قالب الگوهای رخداد تعیین می‌کنند.

۲-۷جعبه ابزار تحلیل انتقال حالت

اگرچه زبان‌های قانون محور در تشریح الگوهای حمله برای تعیین سوء استفاده انعطاف پذیر و مشخص می‌باشند در عمل معمولاً به سختی مورد استفاده قرار می‌گیرند. همانگونه که در مشاهده شد به طور کلی مبانی قانونی تخصصی مشهود نبوده و به مهارت برنامه‌ریزان مجرب قانون‌محور برای به روز رسانی آنها نیاز است. STAT برای تشریح این مسئله ایجاد شد.

۲-۸ماشین‌های خودکار رنگی پتری

کومار و اسپافورد (۱۹۹۴) و کومار (۱۹۹۵) تعیین سوء استفاده را در قالب فرآیند هماهنگی الگو مورد بررسی قرار دادند. آنها یک سلسله مراتب خلاصه برای طبقه‌بندی امضاهای نفوذ (یعنی الگوهای حمله) را براساس روابط ساختاری میان رخدادهایی پیشنهاد کردند که شامل امضا بودند. رخدادهای چنین سلسله مراتبی رخدادهای سطح بالایی هستند که می‌توانند برحسب رخدادهای آزمایشی بازبینی سطح پایین تعریف شده و برای تعیین سلسله مراتب خلاصه در یک حالت مشخص مورد استفاده قرار گیرند. مزیت این طرح طبقه‌بندی آن است که پیچیدگی تعیین امضاها در هر سطح سلسله مراتب را تشریح می‌کند.

۲-۹تعیین نفوذ انتزاعی محور

اجرای بسیاری از رویکردهای تعیین سوء استفاده با یک مشکل مشترک مواجه است: هر سیستم برای یک محیط خاص نوشته شده و ثابت کرده است که استفاده از سایر محیط‌ها که ممکن است سیاست‌ها و مسائل مشابهی داشته باشند مشکل است. هدف اصلی تعیین نفوذ انتزاعی محور بیان این مشکل است.

۳-تعیین نفوذ در سیستم‌های توزیع شده

سرعت رشد اینترنت نه تنها ابزاری را برای اشتراک منابع و اطلاعات فراهم می‌کند بلکه چالش‌های جدیدی را پیش روی انجمن تعیین نفوذ قرار می‌دهد. با توجه به پیچیدگی و میزان داده‌های بازبینی تولید شده در سیستم‌های بزرگ مقیاس، IDS

[7] RUSSEL
[8] P-BEST

های سنتی که برای هاست‌های شخصی و سیستم‌های شبکه بندی شده کوچک مقیاس طراحی شدند نمی‌توانند مستقیماً در سیستم‌های بزرگ مقیاس مورد استفاده قرار گیرند.

بررسی تعیین نفوذ در سیستم‌های توزیع شده اخیرا بر دو مبحث اصلی متمرکز می‌شود: مقیاس‌پذیری و ناهمگنی. IDS ها در سیستم‌های بزرگ توزیع شده باید مقیاس‌پذیر باشند تا حجم بزرگی از داده‌های بازبینی در چنین سیستم‌هایی را در برداشته باشند. همچنین، این IDS ها باید بتوانند اطلاعات ناهمگن انواع مختلف سیستم‌های مولفه را بررسی کرده و سیستم‌های توزیعی بزرگ را به وجود آورند که بتوانند با سایر انواع IDS ها هماهنگ باشند.

بررسی تعیین نفوذ توزیعی در سه محدوده اصلی صورت می‌گیرد. ابتدا مردم IDS های مقیاس‌پذیر و توزیعی را به وجود آورده یا IDS های موجود را برای مقیاس پذیر کردن آنها در سیستم‌های بزرگ توسعه می‌دهند. دوم، IDS های شبکه محور برای برخورداری از مزیت پروتکل‌های شبکه استاندارد در جهت اجتناب از داده‌های بازبینی ناهمگن در پلتفورم‌های مختلف توسعه می‌یابند. سوم، استانداردها و تکنیک‌ها برای تسهیل اشتراک اطلاعات در میان IDS های مختلف و احتمالا ناهمگن در حال توسعه می‌باشند.

۴-اشتراک اطلاعات در سیستم‌های تعیین نفوذ

با توجه به IDS های تجاری متعدد، این IDS ها باید بتوانند اطلاعات را به اشتراک گذاشته و با یکدیگر تعامل برقرار کرده و در نتیجه عملکرد بهتری داشته باشند. فعالیت‌های تحقیق و توسعه اخیرا به شیوه ای برای توانمندسازی IDSهای مختلف و احتمالاً ناهمگن برای اشتراک اطلاعات تبدیل شدند.

چارچوب تشخیص نفوذ معمول[9]برای مهیا نمودن تشخیص نفوذهای مختلف و مولفه‌های پاسخ (IDR) به اطلاعات مشترک و همکاری و منابع ایجاد شده است. این روند به عنوان بخشی از برنامه قابلیت باززیستی اطلاعات عامل پروژه پژوهشی پیشرفته دفاعی[1]با تمرکز بر امکان فعالیت مشترک پروژه‌های DARPA آغاز می‌شود. CIDF سیستم‌های IDR را که متشکل از چهار نوع مولفه هستند در بر می‌گیرد که از طریق ارسال پیام تعامل دارند: تولیدکننده‌های رویداد (E باکس‌ها)، تحلیل‌گرهای رویداد (A باکس‌ها)، پایگاه داده‌های رویداد (D باکس‌ها) و واحدهای پاسخ (R باکس‌ها) می‌باشد. چارچوب ارتباطی و زبان خاص نفوذ عادی برای همکاری مشترک میان مولفه‌های CIDF ارائه شده است.

۵-مجموعه داده

در این تحقیق، مجموعه داده KDD CUP 99 برای اموزش و تست سیستم تشخیص نفوذ مورد استفاده قرار گرفته است. این مجموعه داده، شامل اطلاعاتی است که از سرورهای شبکه‌های نظامی استخراج شده است. در جدول (۱)، ویژگی‌های مربوط به مجموعه داده مورد استفاده (۴۱ ویژگی)، نمایش داده شده‌اند:

[9] Common Intrusion Detection Framework
[1] Defense Advanced Research Project Agency

جدول ۱. لیست ویژگی‌های مجموعه داده KDD

#	Name	#	Name	#	Name
1	Duration	15	Su-attempted	29	Same-srv-rate
2	Protocol-type	16	Num-root	30	Diff-srv-rate
3	Service	17	Num-file-creations	31	Srv-diff-host-rate
4	Flag	18	Num-shells	32	Dst-host-count
5	Src-bytes	19	Num-access-files	33	Dst-host-srv-count
6	Dst-bytes	20	Num-outbound-cmds	34	Dst-host-same-srv-rate
7	Land	21	Is-hot-login	35	Dst-host-diff-srv-rate
8	Wrong-fragment	22	Is-guest-login	36	Dst-host-same-src-port-rate
9	Urgent	23	Count	37	Dst-host-srv-diff-host-rate
10	Hot	24	Srv-count	38	Dst-host-serror-rate
11	Num-failed-logins	25	Serror-rate	39	Dst-host-srv-serror-rate
12	Logged-in	26	Srv-serror-rate	40	Dst-host-rerror-rate
13	Num-compromised	27	Rerror-rate	41	Dst-host-srv-rerror-rate
14	Root-shell	28	Srv-rerror-rate		

۱-۵متغیرهای تحقیق

در این تحقیق، مجموعه داده KDD CUP 99 برای اموزش و تست سیستم تشخیص نفوذ مورد استفاده قرار گرفته است. این مجموعه داده، شامل اطلاعاتی است که از سرورهای شبکه‌های نظامی استخراج شده است.

چهار دسته موجود در مجموعه داده KDD CUP 99 و زیرمجموعه‌های آن‌ها در جدول (۲) نمایش داده شده‌اند.

جدول ۲. دسته‌بندی انواع حملات موجود در مجموعه داده KDD CUP 99 بر اساس چهار گروه اصلی نفوذ در این مجموعه

DOS (Denial Of Service)	**R2L** (User-to-Root)	**U2R** (Remote-to-Local)	**Probing**
apache2	buffer_overflow	ftp_write	ipsweep
back	httptunnel	guess_password	mscan
land	ps	imap	nmap portsweep
mailbomb	loadmodule	named	saint
Neptune	Multihop	phf	satan
pod	Perl	sendmail	
proccesstable	rootkit	snmpgetattack	
smurf	sqlattack	snmpguess	
teardrop	xterm	spy	
udpstorm		warezclient	
		warezmaster	
		worm	
		xlock	
		xsnoop	

با احتساب داده‌های نرمال، مجموعه داده KDD CUP 99 در ۵ دسته normal، probing، Dos، U2R و R2L قرار می‌گیرند. در پیاده‌سازی‌های صورت گرفته در این تحقیق، ۵ کلاس در نظر گرفته‌ایم:

۱.کلاس normal ۲.کلاس Dos ۳.R2L ۴. U2R ۵.probing

در این تحقیق، کلاس‌های normal، Dos، R2L، U2R و probing را به ترتیب با اعداد ۱ تا ۵ نمایش داده‌ایم.

۳-۵روش تحقیق

در روش پیشنهادی در این تحقیق، ابتدا به پیش‌پردازش داده‌ها پرداخته، داده‌های پیش‌پردازش شده را به الگوریتم، وارد می-نماییم.

الگوریتم پیشنهادی در این تحقیق به صورت زیر است:

۱. پیش‌پردازش داده‌ها: در مرحله پیش‌پردازش، در ابتدا داده‌ها به صورتی در می‌آیند که قابل استفاده توسط الگوریتم باشند، به این صورت که مقادیر ستون‌ها را به صورت عددی تبدیل می‌نماییم.

۲. نرمالسازی داده‌ها: به منظور کسب نتایج بهتر، مقادیر هر ویژگی را بین ۰ تا ۱ نرمالیزه نموده، سپس سطرهای ماتریس کلی داده را به صورت تصادفی جابه‌جا می‌نمائیم تا ترتیب داده‌ها از حالت اولیه جمع‌آوری شده، خارج شود. نرمالیزه نمودن به دلیل دستیابی به دقت بالاتر است. اگر A مقدار ویژگی در یک ستون، A_{max} ماکزیمم مقدار ویژگی و A_{min} مینیمم مقدار ویژگی باشد و A_1 را مقدار نرمال شده در نظر بگیریم، از رابطه (۱) برای نرمال‌سازی داده‌ها استفاده می‌گردد:

$$A_1 = \frac{A - A_{min}}{A_{max} - A_{min}} \quad (۱)$$

۳. مقداردهی جمعیت تصادفی اولیه

۴. اعمال الگوریتم فاخته جهت انتخاب ویژگی

۵. طبقه بندی با شبکه عصبی فازی انفیس

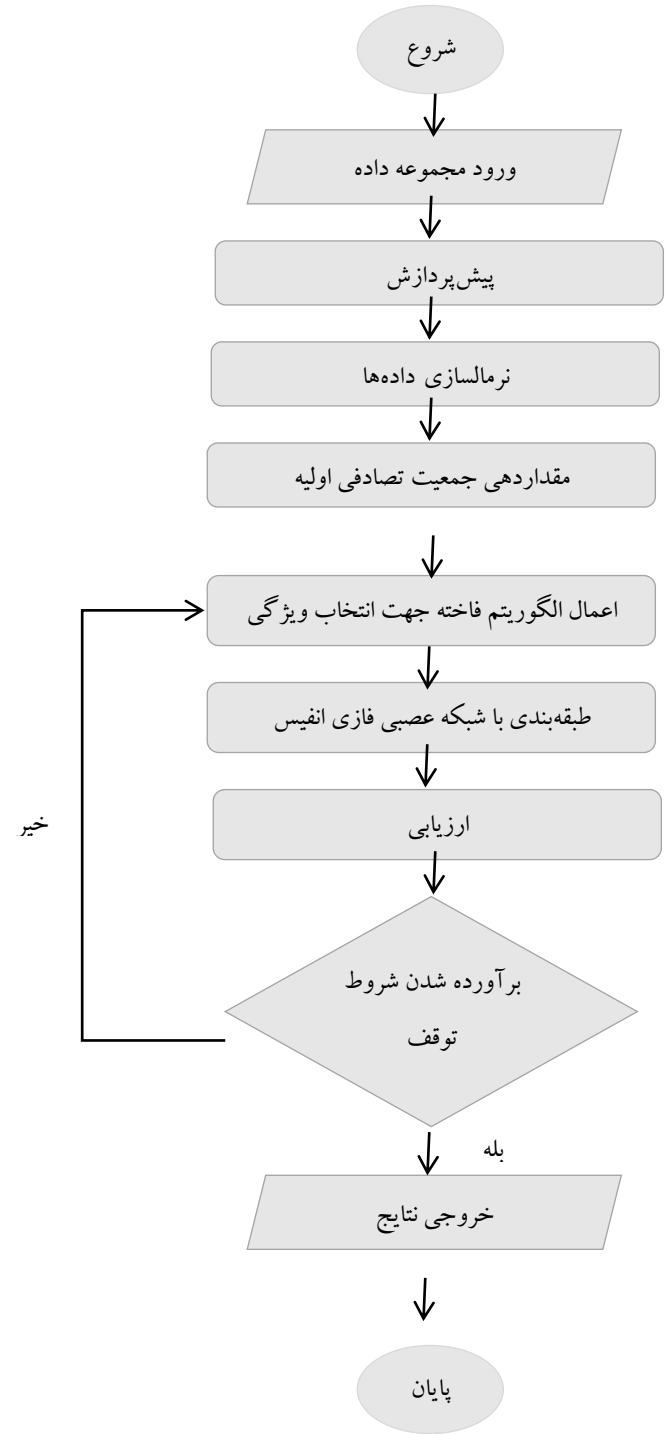
۶. ارزیابی

۷. درصورت برآورده شدن شرط یا شروط توقف (میزان خطا یا تعداد تکرار باشد) رفتن به مرحله بعد در غیر این صورت بازگشت به مرحله ۴

۸. خروجی نتایج

۹. پایان

فلوچارت روش پیشنهادی در این تحقیق به صورت شکل (۱)، نمایش داده شده است.



شکل ۱. فلوچارت روش پیشنهادی در تحقیق حاضر

پیاده‌سازی شبیه‌سازی‌ها در این تحقیق با نرم افزار Matlab انجام می شود. به این دلیل که در اکثر مقالات پیشین که در زمینه تشخیص نفوذ مجموعه داده مورد استفاده در این تحقیق را به کارگرفته‌اند، از نظر پارامتر accuracy مورد بررسی قرار گرفته‌اند، در این تحقیق نیز کارایی روش پیشنهادی را با پارامتر accuracy بررسی می‌کنیم.

(۲)

$$\text{Accuracy} = (TP + TN)/(TP + TN + FP + FN)$$

که در این معادله:

- مثبت واقعی (TP[1]): تعداد نمونه‌های حمله که به درستی دسته‌بندی شدند.
- منفی حقیقی (TN[2]): تعداد نمونه‌های عادی که به درستی دسته‌بندی شدند.
- مثبت اشتباه (FP[3]): تعداد نمونه‌های عادی که به اشتباه به عنوان حمله دسته‌بندی شدند.
- منفی اشتباه (FN[4]): تعداد نمونه‌های حمله‌ای که به اشتباه به صورت عادی دسته‌بندی شدند.

۶-تجزیه و تحلیل

در این بخش ابتدا در آزمایش ۱ به روش پیشنهادی پرداخته، دو آزمایش دیگر نیز با دو روش شبکه عصبی و سیستم فازی انجام می‌دهیم و نتایج آزمایشات را با یکدیگر مقایسه می‌نماییم.

➢ آزمایش ۱: روش پیشنهادی در تحقیق حاضر

پیش‌پردازش داده‌ها: جهت پیش‌پردازش داده‌ها ابتدا مقادیر ستون‌ها را به صورت عددی تبدیل نموده، همچنین، ورودی‌های ستون کلاس به کلاس‌های دودویی ۰ یا ۱ تبدیل می‌شوند که در آن ۰ رکورد عادی و ۱ رکورد حمله را نشان می‌دهد. سپس سطرهای ماتریس کلی داده را به صورت تصادفی جابه‌جا نموده‌ایم تا ترتیب داده‌ها از حالت اولیه جمع‌آوری شده، خارج شود.

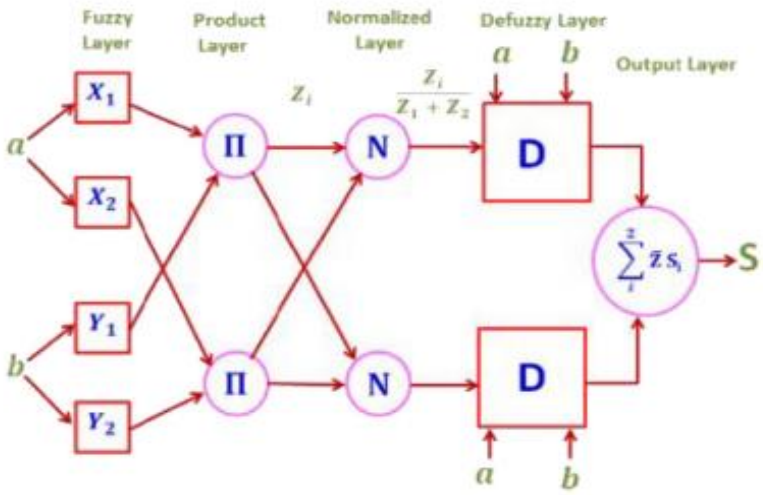
نرمال‌سازی داده‌ها: در مرحله بعد، نرمال‌سازی داده‌ها مطابق توضیحات فصل ۳، انجام شده است. نرمال‌سازی داده‌ها فرآیند تبدیل یا مقیاس‌دهی مقادیر داده‌های هر ویژگی به یک محدوده متناسب است. نرمال‌سازی داده‌ها یک مرحله مهم برای حذف سوگیری با ویژگی‌هایی با مقادیر بزرگ‌تر از مجموعه داده‌ها است.

انتخاب ویژگی: KDDCUP 99 شامل ۴۱ ویژگی است، که البته همه آن‌ها برای ساختن IDS اهمیت ندارند. یک زیرمجموعه از این ویژگی‌ها باید برای دستیابی به نرخ تشخیص بالاتر و هشدار اشتباه پایین‌تر انتخاب شوند. علاوه بر این، فرآیند انتخاب ویژگی برای حذف تعداد ویژگی‌هایی که برای ساختن IDS ضروری هستند مهم می‌باشد. در این تحقیق، از الگوریتم فاخته جهت انتخاب ویزگی استفاده گردیده است.

طبقه‌بندی با سیستم فازی عصبی انفیس: وقتی فرآیند انتخاب ویژگی با استفاده از الگوریتم فاخته صورت گرفت، مجموعه ویژگی‌ها با استفاده از سیستم استنتاج فازی- عصبی تطبیقی (انفیس) آموزش می‌بینند تا بین کلاس‌های عادی و تهاجمی تمایز صورت گیرد. سپس، مدل آموزشی با استفاده از مجموعه آزمایشی ارزیابی می‌شود.

روش سیستم استنتاج فازی- عصبی تطبیقی، روشی است که با استفاده از یک شبکه بازگشتی چند لایه و الگوریتم‌های یادگیری شبکه عصبی و منطق فازی به طراحی نگاشت غیرخطی بین فضای ورودی و خروجی می‌پردازد. شکل (۳-۱) ساختار یک شبکه انفیس ساده را با دو متغیر ورودی x و yنمایش می‌دهد. این شبکه از ۵ لایه تشکیل گردیده است که هر متغیر ورودی دارای دو زیر مجموعه فازی می‌باشد. همان‌طور که شکل (۱-۳) مشاهده می‌نمایید A_1 و A_2 زیرمجموعه x و B_1 و B_2 زیر مجموعه y هستند.

[1] True Positive
[2] True Negative
[3] False Positive
[4] False Negative



شکل ۲- ساختار شبکه انطباقی مبتنی بر سیستم‌های استنتاج فازی (Rajesh et al,. 2022)

سیستم استنتاج فازی این شبکه، از نوع تاکاگی-سوگونو بوده و ساختار آن دارای ۵ لایه می‌باشد. اولین لایه‌ی پنهان، متغیرهای ورودی را به طور نسبی به توابع عضویت نظیر می‌کند. گره‌های این لایه انطباقی بوده و خروجی حاصل از این لایه بر اساس فرمول (۳) محاسبه می‌گردد.

$$O_i^1 = \mu_{A_i}(x) \qquad (۳)$$

x ورودی گره‌ی i، Ai زیر مجموعه فازی مربوط به متغیر ورودی x است O_i^1 خروجی حاصل از iامین گره از اولین لایه و $\mu_{A_i}(x)$ تابع عضویت مربوط به زیر مجموعه فازی متغیر ورودی x است با ماکسیمم مقدار ۱ و مینیمم مقدار ۰.

گره‌های ثابت لایه‌ی دوم، میزان آستانه هر یک از قوانین را از طریق ضرب مقادیر ورودی محاسبه کرده و به عنوان خروجی در نظر می‌گیرند. کارکرد هر گره در این لایه، یافتن وزن (w) برای قوانین فازی داده شده با استفاده از توابع عضویت می‌باشد. دستیابی به این مقدار از طریق فرمول (۲-۳) ممکن می‌گردد.

$$w_i = \mu_{A_i}(x) \times \mu_{B_i}(y) \quad , i = 1,2, \ldots \qquad (۴)$$

وزن‌های بدست آمده از لایه‌ی دوم، توسط گره‌های سومین لایه و از طریق فرمول (۵)، نرمال می‌گردند.

$$\bar{w}_i = \frac{w_i}{w_1 + w_2} \qquad (۵)$$

گره‌ی iام در لایه چهارم، سهم iامین قانون را در خروجی نهایی از طریق تابع گره‌ی (۴) محاسبه می‌نماید. wi خروجی لایه‌ی سوم، {pi, qi, ri}مجموعه پارامترهای موسوم به پارامترهای پیامد می‌باشند. این لایه بخش نتیجه‌ی قانون را در مدل ANFIS نمایش می‌دهد.

$$O_i^4 = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \qquad (۶)$$

تنها گره‌ی لایه‌ی پنجم، خروجی نهایی را از طریق محاسبه‌ی برآیند سیگنال‌های ورودی و با استفاده از فرمول (۵-۳)، به دست می‌آورد.

$$O_i^4 = \text{overall output} = \sum_i \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i} \qquad (۷)$$

➢ آزمایش۲

در آزمایش دوم، طبقه‌بندی را با شبکه عصبی انجام می‌دهیم. در واقع مراحل این آزمایش، به جز مرحله ۵، مشابه روش پیشنهادی هستند. در مرحله ۵، طبقه‌بندی با شبکه عصبی صورت می‌گیرد.مراحل آزمایش ۲ به صورت زیر هستند:

۱. پیش‌پردازش داده‌ها
۲. نرمال‌سازی داده‌ها
۳. مقداردهی جمعیت تصادفی اولیه
۴. اعمال الگوریتم فاخته جهت انتخاب ویژگی
۵. طبقه بندی با شبکه عصبی
۶. ارزیابی
۷. درصورت برآورده شدن شرط یا شروط توقف (میزان خطا یا تعداد تکرار باشد) رفتن به مرحله بعد در غیر این صورت بازگشت به مرحله ۴
۸. خروجی نتایج
۹. پایان

کلیه مراحل به جز مرحله ۵ مطابق با توضیحات صورت گرفته برای روش پیشنهادی صورت می‌گیرند، در مرحله ۵، از شبکه عصبی پرسپترون سه لایه استفاده شده است که دارای یک لایه ورودی، یک لایه خروجی و یک لایه پنهانی است.

➢ آزمایش۳

در آزمایش سوم، طبقه‌بندی را با سیستم فازی انجام می‌دهیم. مراحل آزمایش ۳ نیز مانند آزمایش ۲، به جز مرحله ۵، مشابه روش پیشنهادی هستند. در مرحله ۵، طبقه‌بندی با سیستم فازی صورت می‌گیرد. مراحل آزمایش ۳ به صورت زیر هستند:

۱. پیش‌پردازش داده‌ها
۲. نرمال‌سازی داده‌ها
۳. مقداردهی جمعیت تصادفی اولیه
۴. اعمال الگوریتم فاخته جهت انتخاب ویژگی
۵. طبقه بندی با سیستم فازی
۶. ارزیابی
۷. درصورت برآورده شدن شرط یا شروط توقف (میزان خطا یا تعداد تکرار باشد) رفتن به مرحله بعد در غیر این صورت بازگشت به مرحله ۴
۸. خروجی نتایج
۹. پایان

کلیه مراحل به جز مرحله ۵ مطابق با توضیحات صورت گرفته برای روش پیشنهادی صورت می‌گیرند، در مرحله ۵، جهت طبقه‌بندی از مکانیزم فازی سوگنو استفاده شده است.

معیارهای عملکردی: چندین معیار برای ارزیابی الگوریتم‌های انتخاب ویژگی وجود دارند. معیار منتخب به ماهیت کاربرد بستگی دارد. در این تحقیق معیاری که برای برازش کارایی الگوریتم پیشنهادی در نظر گرفته شده، عبارت است از: accuracy و فرمول آن به صورت معادله (۶-۴) می‌باشد.

$$\text{Accuracy} = (TP + TN)/(TP + TN + FP + FN) \quad (۸)$$

که در این معادله:

- مثبت واقعی (TP): تعداد نمونه‌های حمله که به درستی دسته‌بندی شدند.

- منفی حقیقی (TN): تعداد نمونه‌های عادی که به درستی دسته‌بندی شدند.
- مثبت اشتباه (FP): تعداد نمونه‌های عادی که به اشتباه به عنوان حمله دسته‌بندی شدند.
- منفی اشتباه (FN): تعداد نمونه‌های حمله‌ای که به اشتباه به صورت عادی دسته‌بندی شدند.

در این تحقیق همچنین در الگوریتم پیشنهادی مطرح شده در بخش (۵-۳)، در مرحله بعد انتخاب ویژگی با الگوریتم بهینه‌سازی فاخته، طبقه‌بندی را با الگوریتم‌های ANN و سیستم فازی نیز انجام داده و نتایج را با روش پیشنهادی مقایسه می‌نماییم. جدول (۳) مقایسه بین الگوریتم پیشنهادی تحقیق حاضر و این روش‌ها را نمایش می‌دهد.

جدول۳. مقایسه الگوریتم پیشنهادی تحقیق حاضر و الگوریتم‌های COA-ANN و COA-Fuzzy

	Normal	**Dos**	**R2L**	**Probing**	**U2R**
COA-ANN	**95.18**	**95.24**	**89.63**	**91.57**	**84.26**
COA-Fuzzy	**96.37**	**96.12**	**93.2**	**94.93**	**86.53**
COA-ANFIS (Proposed method)	97.53	96.71	93.26	95.68	91.20

ارزیابی دقت در سه آزمایش انجام شده، نشان می‌دهد که دقت الگوریتم پیشنهادی نسبت به دو آزمایش دیگر، بالاتر بوده است.

۷-نتیجه گیری

در این پایان‌نامه با بهره‌گیری از الگوریتم فاخته و سیستم فازی با هدف افزایش دقت تشخیص، به ارائه روشی ترکیبی پرداختیم. از الگوریتم فاخته که یک الگوریتم تکاملی است برای انتخاب ویژگی های متمایزتر، استفاده شده است..دلیل برتری الگوریتم فاخته نسبت به الگوریتم های تکاملی دیگر، در کارکرد چندگانه عملگرهای الگویتم فاخته از قبیل تخم‌گذاری و مهاجرت است که باعث تحقق همزمان چندین هدف می شود. همچنین برای طبقه بندی نیز از سیستم فازی استفاده گردیده است.

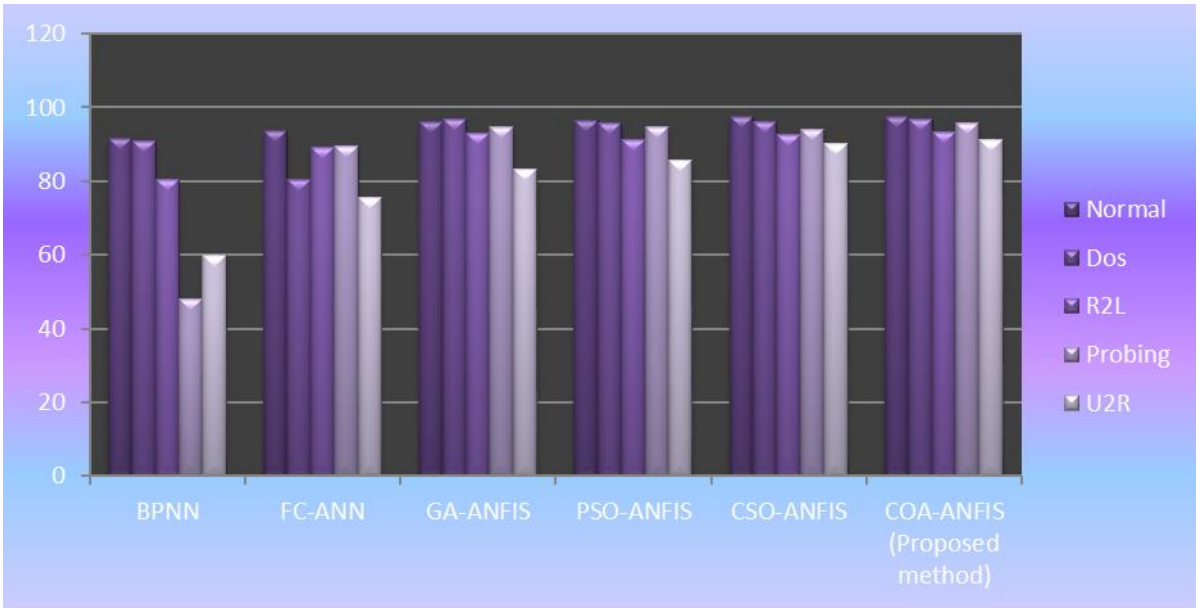
در این تحقیق علاوه بر روش پیشنهادی، دو آزمایش دیگر نیز صورت گرفت، به این صورت که مرحله طبقه‌بندی در روش پیشنهادی با شبکه عصبی و سیستم فازی انجام شد. نتایج، نشان دهنده دقت بالاتر روش پیشنهادی نسبت به دو آزمایش دیگر بود.

در این قسمت همچنین قصد داریم روش پیشنهادی را با روش‌های مورد بررسی در تحقیق صورت گرفته توسط Manimurugan و همکاران در سال ۲۰۲۰ مقایسه نماییم. جدول (۴) مقایسه بین الگوریتم ترکیبی مورد استفاده در این پایان‌نامه و این روش‌ها را نمایش می‌دهد.

جدول ۴ مقایسه الگوریتم ترکیبی مورد استفاده در تحقیق حاضر و الگوریتم‌های مورد بررسی در تحقیق Manimurugan و همکاران (۲۰۲۰)

	Normal	**Dos**	**R2L**	**Probing**	**U2R**
BPNN	91.50	90.94	80.53	48.13	60.00
FC_ANN	93.87	80.35	89.12	89.57	75.58
GA-ANFIS	96.22	96.70	93.18	94.89	83.33
PSO-ANFIS	96.46	95.90	91.35	94.72	85.62
CSO-ANFIS	97.41	96.25	92.51	94.15	90.26
COA-ANFIS (Proposed method)	97.53	96.71	93.26	95.68	91.20

همچنین در شکل (۳)، نمودار مقایسه دقت الگوریتم پیشنهادی در تحقیق حاضر و الگوریتم‌های مورد استفاده در تحقیق انجام شده توسط Manimurugan و همکاران (۲۰۲۰) نمایش داده شده است.



شکل ۳. نمودار مقایسه دقت الگوریتم‌های مورد بررسی

کسب نتایج بهتر روش به کار گرفته شده در این تحقیق را نسبت به دیگر الگوریتم‌های مقایسه‌شده نشان می‌دهند. این الگوریتم در شرایط مشابه دقت بالاتری داشته است.

در کل ارزیابی رویکرد پیشنهادی از نقطه‌نظر دقت نشان می‌دهد که دقت الگوریتم طبقه‌بندی مورد بررسی نسبت به دیگر رویکردهای مشابه بهتر بوده است.

۷-۱پیشنهادها

با توجه به نتایج این پایان‌نامه، مهم‌ترین پیشنهاداتی که می‌تواند به‌عنوان چارچوبی برای تحقیقات آتی مدنظر قرار گیرند عبارتند از:

- استفاده از الگوریتم‌های تکاملی دیگر جهت انتخاب ویژگی و ارزیابی و مقایسه نتایج
- استفاده از دیگر الگوریتم‌های طبقه‌بندی به جای الگوریتم‌های طبقه‌بندی مورد استفاده در تحقیق حاضر و مقایسه نتایج به‌دست آمده با نتایج این تحقیق
- در نظر گرفتن معیارهای دیگر به‌جز دقت و ارزیابی نتایج

۷-۲ منابع

- Almasoudy, F. H., Al-Yaseen, W. L., & Idrees, A. K. (2020). Differential Evolution Wrapper Feature Selection for Intrusion Detection System. Procedia Computer Science, 167, 1230-1239.
- Alazzam, H., Sharieh, A., & Sabri, K. E. (2020). A feature selection algorithm for intrusion detection system based on pigeon inspired optimizer. Expert Systems with Applications, 148, 113249.
- Acharya, Neha, and Shailendra Singh. (2018)“An IWD-based feature selection method for intrusion detection system.” Soft Comput. 22 (13): 4407– 4416.
- A Jahanbani, H Karimi .A new approach for detecting intrusions based on the PCA neural networks Journal of Basic and Applied Scientific Research, 2012.
- A Jahanbani, M Keshtgari, F Monadjemi -Intrusion Detection System Using New Synthetic Neural Networks, Journal of Basic and Applied Scientific Research, 2012

- Chakravarty, S. (2020, June). Feature Selection and Evaluation of Permission-based Android Malware Detection. In 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184) (pp. 795-799). IEEE.
- Dua, M. (2020). Attribute Selection and Ensemble Classifier based Novel Approach to Intrusion Detection System. Procedia Computer Science, 167, 2191-2199.
- Dwivedi, S., Vardhan, M., & Tripathi, S. (2020). An Effect of Chaos Grasshopper Optimization Algorithm for Protection of Network Infrastructure. Computer Networks, 107251.
- Enache, A. C., Sgarciu, V., & Petrescu-Niţă, A. (2015, May). Intelligent feature selection method rooted in Binary Bat Algorithm for intrusion detection. In 2015 IEEE 10th Jubilee International Symposium on Applied Computational Intelligence and Informatics (pp. 517-521). IEEE.
- Heidari, A. A., Mirjalili, S., Faris, H., Aljarah, I., Mafarja, M., & Chen, H. (2019). Harris hawks optimization: Algorithm and applications. Future Generation Computer Systems, 97, 849-872.
- Hajisalem,V.,Babaie,S., (2018). A hybrid intrusion detection system based on abc-afs algorithm for misuse and anomaly detection. Computer Networks 136,37–50.doi:10.1016/j.comnet.2018.02.028.
- Khammassi, C., & Krichen, S. (2020). A NSGA2-LR wrapper approach for feature selection in network intrusion detection. Computer Networks, 107183.
- Kumar, G. (2020). An improved ensemble approach for effective intrusion detection. The Journal of Supercomputing, 76(1), 275-291.
- Koay, A., Chen, A., Welch, I., & Seah, W. K. (2018, January). A new multi classifier system using entropy-based features in DDoS attack detection. In Information Networking (ICOIN), 2018 International Conference on (pp. 162-167). IEEE.
- Kamarudin, M. H., Maple, C., Watson, T., & Safa, N. S. (2017). A New Unified Intrusion Anomaly Detection in Identifying Unseen Web Attacks. Security and Communication Networks, 2017.
- KDDCup,1999.Availableon:http://kdd.ics.uci.edu/databases/kddcup99/ kddcup99.html, Ocotber 2007.
- Khan, J. A., & Jain, N. (2016). A survey on intrusion detection systems and classification techniques. Int. J. Sci. Res. Sci., Eng. Technol., 2(5), 202-208.
- Kosamkar, V. and Chaudhari, S.S. (2014), "Improved Intrusion Detection System using C4.5 Decision Tree and Support Vector Machine" International Journal of Computer Science and Information Technologies, Vol. 5 (2) , 14631467.
- Khalid, S., Khalil, T., & Nasreen, S. (2014, August). A survey of feature selection and feature extraction techniques in machine learning. In 2014 Science and Information Conference (pp. 372-378). IEEE.
- Li, X., Chen, W., Zhang, Q., & Wu, L. (2020). Building Auto-Encoder Intrusion Detection System Based on Random Forest Feature Selection. Computers & Security, 101851.
- Mafarja, M., Heidari, A. A., Habib, M., Faris, H., Thaher, T., & Aljarah, I. (2020). Augmented whale feature selection for IoT attacks: Structure, analysis and applications. Future Generation Computer Systems.
- Mishra, P., Pilli, E. S., Varadharajan, V., & Tupakula, U. (2017). Intrusion detection techniques in cloud environment: A survey. Journal of Network and Computer Applications, 77, 18-47.

- Mafarja, M. M., & Mirjalili, S. (2017). Hybrid whale optimization algorithm with simulated annealing for feature selection. Neurocomputing, 260, 302-312.
- Manimurugan, S., Majdi, A. Q., Mohmmed, M., Narmatha, C., & Varatharajan, R. (2020). Intrusion detection in networks using crow search optimization algorithm with adaptive neuro-fuzzy inference system. Microprocessors and Microsystems, 79, 103261.
- Nguyen, B. H., Xue, B., & Zhang, M. (2020). A survey on swarm intelligence approaches to feature selection in data mining. Swarm and Evolutionary Computation, 54, 100663.
- Phutane, T. and Pathan, A. (2015), "Intrusion Detection System using Decision Tree and Apriori Alghoritm" International Journal of Computer Engineering and Technology, Volume 6,Issue 7.
- Perin, A. and Gamback, B. (2013), "Ensembles of Decision Trees for Network Intrusion Detection Systems" International Journal on Advances in Security, vol 6 no 1 & 2.
- Ramin Rajabioun,Cuckoo Optimization Algorithm,Applied Soft Computing,Volume 11, Issue 8,2011, Pages 5508-5518,ISSN 1568-4946.
- Sugianela, Y., & Ahmad, T. (2020, February). Pearson Correlation Attribute Evaluation-based Feature Selection for Intrusion Detection System. In 2020 International Conference on Smart Technology and Applications (ICoSTA) (pp. 1-5). IEEE.
- Salo, F., Nassif, A. B., & Essex, A. (2019). Dimensionality reduction with IG-PCA and ensemble classifier for network intrusion detection. Computer Networks, 148, 164175.
- Singh, K., & Singh, K. (2018). Intrusion Detection and Recovery of MANET by Using ACO Algorithm and Genetic Algorithm. In Next-Generation Networks (pp. 97-109). Springer, Singapore.
- Tubishat, M., Idris, N., Shuib, L., Abushariah, M. A., & Mirjalili, S. (2020). Improved Salp Swarm Algorithm based on opposition based learning and novel local search algorithm for feature selection. Expert Systems with Applications, 145, 113122.
- Vargas-Muñoz, M. J., Martínez-Peláez, R., Velarde-Alvarado, P., Moreno-García, E., TorresRoman, D. L., & Ceballos-Mejía, J. J. (2018, February). Classification of network anomalies in flow level network traffic using Bayesian networks. In Electronics, Communications and Computers (CONIELECOMP), 2018 International Conference on (pp. 238-243). IEEE.