دانشنـــــــامــه علــوم فنــی و مهندســی

جلد۱،شماره۱،آذر۱۴۰۲

صفحه ۵۳-۶۵

ترکیب الگوریتم های سنجاقک و ملخ در انتخاب ویژگی ها برای تشخیص نفوذ در شبکه

۱- رضا شمسائی ۲- علی دلیری بیدختی

۱- استادیار دانشکده کامپیوتر و فناوری اطلاعات، دانشگاه صنعتی سجاد
۲- دانشجوی کارشناسی ارشد شبکه های کامپیوتری دانشگاه صنعتی سجاد

r_shamsaee@hotmail.com
Daliri@gmail.com

چکیده

یادگیری ماشین می‌تواند راهکار مناسب سیستم تشخیص نفوذ باشد بشرط آنکه بتواند در مواجه با ویژگی‌های متعدد داده‌های تشخیص نفوذ و همچنین الگو متفاوت آن‌ها، فاز انتخاب ویژگی و فاز طبقه بندی کارا داشته باشد. در مدل پیشنهادی پیدا کردن تاثیرگذارترین ویژگی‌ها توسط روش تکاملی ترکیبی انجام شده و سپس روش یادگیری چندتایی با رأی گیری، نتیجه تشخیص نفوذ را تعیین می‌کند. نتایج مدل پیشنهادی یعنی روش تکاملی ترکیبی الگوریتم سنجاقک و ملخ در فاز انتخاب ویژگی و طبقه بندی چندتایی در مقایسه با الگوریتم جستجوی ممنوع، سنجاقک و ملخ بصورت تکی در فاز انتخاب ویژگی و طبقه بندی جنگل تصادفی بروی مجموعه داده استاندارد UNSW-NB نشان داد که مدل پیشنهادی توانسته بطور میانگین تا ۱،۸ درصد تأثیر مثبت بروی نتایج تشخیص نفوذ داشته باشد.

کلمات کلیدی: تشخیص نفوذ، الگوریتم بهینه سازی سنجاقک، الگوریتم بهینه سازی ملخ، یادگیری چندتایی

۱. مقدمه

در عصر نوظهور شبکه های مختلف از جمله اینترنت اشیاء، اعتماد یک فاکتور حیاتی برای استقرار دستگاه‌های هوشمند بدون مداخله انسان همراه با کاهش ریسک محسوب می‌شود. درحالی که ادغام اشیاء فیزیکی، اجزای سایبری و انسان در زیرساخت شبکه اینترنت اشیاء نگرانی های جدیدی را برای دستیابی به اعتماد موجب شده است که میتوان به شبکه‌های حسگر و عملگرها در شهر هوشمند اشاره نمود.برای ایجاد امنیت کامل در یک شبکه در دنیای امروز، علاوه بر تجهیزات جلوگیری از حمله، سیستم‌هایی به نام سیستم‌های تشخیص حمله مورد نیاز است تا بتوانند در صورتی که نفوذگر از تجهیزات امنیتی عبور کرد و وارد شبکه شد، آن را تشخیص داده و چاره ای برای مقابله با آن بیاندیشند. در این پژوهش به ارائه راهکاری جدید برای بهبود تشخیص حمله در شبکه‌ها با استفاده از الگوریتم ترکیبی بهینه سازی سنجاقک و ملخ ارائه می‌شود.

۲. بیان مسئله

در سیستم‌های تشخیص نفوذ یکی از چالش‌ها وجود ویژگی‌های غیرمهم در مجموعه داده است و نیاز به انتخاب ویژگی دارد. کاهش ابعاد از طریق انتخاب و استخراج ویژگی یکی از مهم‌ترین مراحل در آماده سازی داده است. در انتخاب ویژگی، هدف انتخاب زیرمجموعه از ویژگی‌ها که قادر به توصیف هدف یا کلاس داده‌ها باشند، است. گاهی ویژگی‌های اضافی، تأثیر نامطلوبی بر روی دسته بندی داده‌ها داشته و این تأثیرات می‌تواند با انتخاب ویژگی مناسب کاهش یابد[۱،۲].

مسئله انتخاب بهترین مجموعه ویژگی در مجموعه داده‌های نفوذ به شبکه از مسائل سخت است که روش‌های قدیمی راهکار مناسب برای حل نیستند زیرا محاسباتی بالایی داشته و به بهترین زیر مجموعه ویژگی‌ها، همیشه نمی‌رسند. الگوریتم‌های فراابتکاری با تکاملی روش‌های مناسبی است اما نکته‌ای که هست اینکه الگوریتم‌های تکاملی همیشه در هر اجرا، به یک مجموعه ویژگی ثابت دست پیدا نمی‌کنند، بنابراین مسئله انتخاب بهترین مجموعه ویژگی یک مسئله با بهینه‌های متعدد بشمار می‌رود[۳].

در این تحقیق پیشنهاد یک راهکار در انتخاب ویژگی در مجموعه داده‌های نفوذ با استفاده از الگوریتم ترکیبی سنجاقک[۴] و ملخ [۵] مطرح شده است. در واقع الگوریتم سنجاقک دارای قدرت اکتشاف بالا و الگوریتم ملخ دارای قدرت استخراج بالا است و بنظر می‌رسد ترکیب این دو الگوریتم بتواند فضای جستجو را با دقت بیشتری جستجو نماید. بطور خلاصه نکات زیر در ابعاد بیان مسئله مطرح است:

- مسئله انتخاب ویژگی از مسائل سخت است که روش‌های قدیمی راهکار مناسبی برای حل آن نیست زیرا محاسباتی بالایی داشته و بهترین زیر مجموعه ویژگی‌ها را حاصل نمی‌کنند .
- الگوریتم‌های فراابتکای روش مناسبی برای مسئله انتخاب ویژگی بوده زیرا هزینه محاسبات مناسب داشته و به بهترین ویژگی‌ها دست پیدا می‌کنند.
- سیستم‌های تشخیص نفوذ با مجموعه داده‌هایی برخورد دارند که شامل ویژگی‌های غیرمهم است و بایستی انتخاب ویژگی در این حوزه انجام شود.
- الگوریتم سنجاقک یک الگوریتم کارا در بهینه سازی و مسئله انتخاب ویژگی است اما بصورت ترکیبی با الگوریتم بهینه سازی ملخ در حوزه سیستم‌های تشخیص نفوذ استفاده نشده است.

۳.ادبیات تحقیق

در این قسمت بطور ویژه در حوزه امنیت در شبکه ها در مقالات سال اخیر برخی از تحقیقات اشاره می شود. درواقع هدایت و کنترل کارآمد سیستم های مقیاس بزرگ در شبکه، وظیفه پیچیده و چالش برانگیزی است. سیستم عامل های محاسباتی باید توانایی پردازش و تجزیه و تحلیل ایمن و به هنگام را برای داده های بزرگ داشته باشند [۶و۷]. علاوه بر این، ظرفیت و توان عملیاتی سیستم باید زیاد باشد تا انتقال داده را با کمترین زمان تأخیر و درصد اطمینان بالا انجام دهند. الگوریتم ها و مدل های یادگیری ماشینی به طور قابل توجهی عملکرد این حوزه را از نظر درصد اطمینان و امنیت بهبود بخشیده اند. این الگوریتم ها پتانسیل زیادی برای بررسی چالش های امنیتی در شبکه ها دارند [۸و۹].

در تحقیق [۱۰] یک مدل انتخاب ویژگی رپر و طبقه بندی چندتایی برای شناسایی حمله به شبکه ارائه شده است. در این مدل از یادگیری با استفاده از قسمت رمزگذاری که بصورت تطبیقی عمل میکند، استفاده شده است. در این مقاله محققان از مجموعه داده KDD-CUP ۹۹ برای ارزیابی طرح پیشنهادی خود استفاده کردند و دقت تشخیص حمله ۹۴,۷۱٪ بود. نتایج تجربی آنها ثابت کرد که عملکرد مدل آنها بهتر از عملکرد شبکه باور عمیق است.

در تحقیق [۱۱]، رمزگذار غیر متقارن برای افزایش امنیت ارائه شدهاست که ویژگی ها را به روشی بدون نظارت یاد می گیرد.سازندگان، مدل پیشنهادی خود را در واحد پردازش گرافیکی دارای تانسور پیاده سازی کرده و مدل را با استفاده از مجموعه داده NSL-KDD ارزیابی کردند. نشان داده شده است که مدل مقاله دارای زمان آموزش کمتری است اما نتوانسته دقت بسیار بالایی داشته باشد و دقت تشخیص حمله ۸۹/۲۲ درصد گزارش شده است.

در پژوهش [۱۲] یک شبکه یادگیری سریع را با ترکیب روش بهینه سازی گروه ذرات، ارائه کردند. در مدل مقاله از الگوریتم گروه ذرات برای تنظیم پارامترهای شبکه استفاده شده است و بدین شکل سرعت یادگیری در این روش افزایش یافته است. نویسنگان، طرح پیشنهادی خود را با استفاده از مجموعه داده KDD ۹۹ اجرا کردند. دقت پیش بینی حمله مدل پیشنهادی آنها ۹۸,۹۲٪ بود. اگرچه مدل آنها عملکرد رضایت بخشی را ارائه می دهد، اما پیچیدگی مدل آنها زیاد است و برای دستگاههای منابع محدود مناسب نیست.

در تحقیق [۱۳] یک الگوریتم ژنتیکی ترکیبی جدید و ماشین بردار پشتیبانی با طرح مبتنی بر بهینه سازی گروه ذرات برای تشخیص حمله DoS ارائه دادند. در واقع در مدل پیشنهادی مقاله از الگوریتم ترکیبی ژنتیک و گروه ذرات استفاده شده است و این الگوریتم ترکیبی برای تنظیم دو پارامتر اصلی در ماشین بردار پشتیبان استفاده شده است. محققان طرح پیشنهادی خود را با استفاده از مجموعه داده KDD ۹۹ اجرا کردند و به دقت ۹۶,۳۸٪ دست یافتند.

در تحقیق [۱۴] با پیشنهاد یک مدل مبتنی بر ماشین بردار پشتیبانی، دقت طبقه بندی را بهبود بخشید. در این مدل پیشنهادی سعی شده است با تنظیم هسته مناسب برای ماشین بردار پشتیبان و همچنین انجام انتخاب ویژگی در مجموعه داده، به افزایش دقت در تشخیص برسند. نویسندگان مقاله در آزمایشات خود را با استفاده از مجموعه داده ADFA-LD که دارای ویژگی های بیشتری نسبت به مجموعه داده KDD است، آزمایشات را انجام دادند و به دقت ۹۴,۵۱٪ دست یافتند.

در [۱۵] با استفاده از انتخاب ویژگی مبتنی بر آنتروپی و درخت تصمیم c۴,۵، در مدلی با یادگیری، حملات شبکه را با موفقیت شناسایی و طبقه بندی کردند. آنها مدل خود را با استفاده از حالتهای مختلف ازدرخت تصمیم آزمایش کردند و بهترین معماری از هر روش را برای تشخیص نفوذ استفاده کردند. نویسندگان مدل پیشنهادی را مجموعه داده KDD CUP ۹۹ پیاده سازی کردند و به دقت ۹۱,۵۰٪ دست یافتند.

در مقاله [۱۶] یک مدل مبتنی بر انتخاب ویژگی غیرنظارتی و مدل ترکیبی ماشین بردار پشتیبان، درخوت c ۴,۵ و روش بیزین، برای تشخیص حملات تزریق داده های کاذب ارائه دادند. در این مدل از انتخاب ویژگی برای کاهش ابعاد داده ها و در مدل یادگیری چندتایی بهره برده شده است. محققان طرح پیشنهادی خود را با استفاده از مجموعه داده های IEEE ۱۱۸ اجرا کردند. دقت تشخیص حمله مدل پیشنهادی آنها ۹۱,۸۰٪ بود.

در مقاله [۱۷] طرح تشخیص نفوذ مبتنی انتخاب ویژگی الگوریتم میمون عنکبوتی پیشنهاد شده است. این رویکرد در کنار طبقه بند شبکه عصبی روبه جلو، با موفقیت توانست حملات DoS و حملات شخص میانی را در برنامه های شبکه با مجموعه داده های متفاوت شناسایی کند. محققان طرح پیشنهادی خود را با استفاده از مجموعه داده های NSL-KDD ارزیابی کردند و دقت تشخیص حمله مدل آنها ۹۱,۶۵٪ بود.

در تحقیق [۱۸] یک چارچوب یادگیری مبتنی بر انتخاب ویژگی با ضریب همبستگی مبتنی بر شبکه، برای حملات فیشینگ و بات نت ارائه دادند.در این مدل سیستم تشخیص نفوذ در ابر قرار داشته و عملیات تشخیص در این سطح و با استفاده از شبکه عصبی روبه جلو انجام شده است. در واقع یادگیری الگو های فیشینگ و نفوذ در شبکه و دسترسی به اطلاعات را یادگیری نموده و در هر درخواست برای استفاده از منابع ابری، وجود تشخیص نفوذ را بررسی میکند.در این مقاله برای حملات فیشینگ و بات نت، نتایج تجربی به ترتیب مقادیر دقت ۹۴,۳۰٪ و ۹۴,۸۰٪ را نشان میدهد.

در مقاله [۱۹] یک روش یادگیری مبتنی بر انتخاب ویژگی با الگوریتم ژنتیک و طبقه بندی ترکیبی نزدیکتریم همسایه، ماشین بردار پشتیبان، درخت تصمیم j۴۸ را برای تشخیص نفوذ شبکه ارائه دادند. محققان دقت طرح پیشنهادی را با استفاده از مجموعه داده های NSL-KDD ارزیابی کردند. دقت رویکرد آنها ۹۲/۳۵ درصد بود. در تحقیق دیگری [۲۰] یک تحلیل مقایسه ای روی تکنیک های موجود مبتنی بر یادگیری ماشینی برای تشخیص حمله شبکه ارائه دادند و در آن الگوریتم ژنتیک برای انتخاب ویژگی و مدل یادگیری چندتایی شامل شبکه عصبی، ماشین بردار پشتیبان و درخت تصمیم j۴۸ استفاده شده است و نتایج خوبی تا ۹۱ درصد دقت داشته است.

در تحقیق [۲۱] یک طرح تشخیص نفوذ در حوزه شبکه حوزه پزشکی با انتخاب ویژگی الگوریتم پروانه و طبقه بندی شبکه عصبی ارائه شد. محققان طرح خود را با استفاده از مجموعه داده های NSL-KDD ارزیابی کردند. نتایج تجربی آنها کارایی را تا ۹۲ درصد نشان داده اند.در تحقیق دیگری [۲۲] الگوریتم گروه ذرات در انتخاب ویژگی و طبقه بند درخت های تصمیم از جمله CART| و C۴,۵ و J۴۸ برای تشخیص حمله سایبری پیشنهاد کردند. محققان عملکرد طرح پیشنهادی را با الگوریتم های درخت تصمیم و مدل رگرسیون لاجستیک مقایسه کردند. در مقاله [۲۳] یک طرح شناسایی نفوذ مبتنی بر انتخاب ویژگی با اطلاعات متقابل ارائه دادند. آنها درباره استقرار طرح در گیت های شبکه مباحثی ارائه کردند. آنها با استفاده از طرح پیشنهادی خود با موفقیت، ترافیک مخرب، اسکن درگاه و حمله جستجوی فراگیر را با انتخاب ویژگی ضریب همبستگی و طبقه بندی جنگل تصادفی و شبکه عصبی شناسایی کردند. در مقاله [۲۴] یک مدل ترکیبی در انتخاب ویژگی با الگوریتم گرگ خاکستری و گروه ذرات و طبقه بند جنگل تصادفی و شبکه عصبی در سیستم تشخیص نفوذ برای حملات از راه دور پیشنهاد دادند. آنها با استفاده از مجموعه داده NSL-KDD هر دو حمله را با موفقیت شناسایی کردند.

در مقاله [۲۵] یک چارچوب یادگیری عمیق دو سطحی برای تشخیص بات نت ارائه دادند. محققان با استفاده از الگوریتم تولید دامنه، حملات و ترافیک عادی را با موفقیت دسته بندی کردند. نتایج تجربی آنها بهبود کارایی طرح پیشنهادی را از نظر دقت، نمره F۱ و سرعت تشخیص اثبات کرد.

جدول ۱: مقایسه روشهای مرور ادبیات در سال های اخیر

مقاله	نقطه قوت	نقطه ضعف	روش انتخاب ویژگی / روش طبقه بندی
[۱۵]	دقت دسته بندی بالا تا ۹۴,۷۱ درصد	زمان اجرای بالا در تجزیه	مبتنی بر آنتروپی / درخت تصمیم C۴,۵
[۱۶]	دقت دسته بندی در کلاسهای زیاد تا ۸۹,۲۲ درصد	زمان آموزش بالا	انتخاب ویژگی غیرنظارتی / مدل چندتایی ماشین بردار پشتیبان، درخت C۴,۵ و بیزین
[۱۷]	دقت دسته بندی در مجموعه داده های معمول تا ۹۸,۹۲ درصد	زمان اجرای بالا در مجموعه داده های نا متوازن	الگوریتم میمون عنکبوتی / شبکه عصبی رو به جلو
[۱۸]	تعمیم پذیری روش و دقت دسته بندی تا ۹۶,۳۸ درصد	مقیاس پذیری روش بررسی نشده است	مبتنی بر ضریب همبستگی / شبکه عصبی رو به جلو
[۱۹]	دقت دسته بندی در مجموعه داده هایی با نمونه کم تا ۹۴,۵۱ درصد	تعمیم پذیری محدود در سایر مجموعه داده ها	الگوریتم ژنتیک / مدل چندتایی ماشین بردار پشتیبان، درخت J۴۸ و نزدیکترین همسایه
[۲۰]	سرعت در اجرا و تعمیم پذیری با دقت ۹۱,۸۰ درصد	مقایس پذیری روش پیشنهادی نشان داده نشده است	الگوریتم ژنتیک / مدل چندتایی ماشین بردار پشتیبان، درخت C۴,۵و شبکه عصبی
[۲۱]	تعمیم پذیری در مجموعه داده های مختلف با دقت ۹۲,۴۹ درصد	زمان آموزش بالا است	الگوریتم پروانه / شبکه عصبی رو به جلو
[۲۲]	دقت دسته بندی را تا ۹۱,۶۵ درصد بهمراه داشته باشد	مقایس پذیری روش پیشنهادی نشان داده نشده است	الگوریتم گروه ذرات / درخت تصمیم Cart ,C۴,۵,J۴۸
[۲۳]	دقت تشخیص برای کشف حملات DOS تا ۹۴,۸۰ درصد	تعمیم پذیری بروی حملات دیگر ندارد	مبتنی بر اطلاعات متقابل / شبکه عصبی و جنکل تصادفی
[۲۴]	دقت تشخیص برای کشف حملات DOS تا ۹۲,۳۵ درصد	تعمیم پذیری بروی حملات دیگر در شبکه های دیگر از جمله vanet	ضریب همبستگی و اطلاعات متقابل / جنگل تصادفی و نزدیکترین همسایه
[۲۵]	سرعت تشخیص حملات DoS با دقت ۸۴,۸۶ درصد	مقایس پذیری روش پیشنهادی نشان داده نشده است	ترکیب الگوریتم گرگ خاکستری و گروه ذرات / شبکه عصبی و جنگل تصادفی

به طور خلاصه، اکثر محققان با هدف قرار دادن برخی از برنامه های خاص در شبکه، طرح های تشخیص حمله خود را ارائه دادند. آنها بیشتر مدلهای خود را با استفاده از مجموعه داده های موجود KDD Cup ۹۹ و NSL-KDD ارزیابی کردند. این مجموعه های داده مدت زمان طولانی مورد استفاده قرار گرفته و برنامه های خاص شبکه را هدف قرار داده است. روشهای استفاده شده در مرور ادبیات دارای معماری کم عمق شامل فاز انتخاب ویژگی و طبقه بند هستند و همانطور که مشخص است میزان دقت تشخیص در مقالات با توجه به اینکه مجموعه داده های یکسانی داشته اند با تغییر روش ها، متغییر بوده است.

۴.روش پیشنهادی

در این پیشنهاد یک روش یادگیری ماشین مبتنی بر انتخاب ویژگی برای اولین بار مبتنی بر الگوریتم جدید ترکیبی سنجاقک [۴] و الگوریتم ملخ [۵] ارائه شده است. در مقاله [۱] از مجموعه داده متناسب برای انواع کاربرد شبکه ها استفاده شده است و نیاز به ارائه سیستم تشخیص نفوذ بررسی شده است، در این مقاله از روش جستجوی ممنوع برای انتخاب ویژگی و روش جنگل تصادفی برای طبقه بندی استفاده شده است. اما در مدل پیشنهادی برای انتخاب ویژگی از روش ترکیبی سنجاقک و ملخ و برای طبقه بندی از مدل یادگیری چندتایی شامل جنگل تصادفی، روش نزدیکترین همسایه و ماشین بردار پشتیبان استفاده شده است.

با توجه به اینکه روش انتخاب ویژگی با جستجوی ممنوع و طبقه بندی با جنگل تصادفی [۱] از الگوریتم های فراابتکاری برای انتخاب ویژگی استفاده نکرده و همچنین از یادگیری چندتایی بجای یادگیری تکی بهره نگرفته است، در روش پیشنهادی از الگوریتم فراابتکاری هوش جمعی ملخ و سنجاقک بشکل ترکیبی و یادگیری چندتایی استفاده شده است. در مدل پیشنهادی نیز از مجموعه داده استفاده شده مربوط به مجموعه داده به نام UNSW-NB۱۵ استفاده می شود[۱] تا نتایج با روش این مقاله از نظر صحت تشخیص حمله مقایسه شود.

در روش پیشنهادی از الگوریتم هوش جمعی سنجاقک و ملخ برای انتخاب ویژگی ها استفاده می‌شود و از روش نزدیک‌ترین همسایه برای مشخص کردن میزان برازندگی این مجموعه ویژگی ها استفاده می شود. در روش پیشنهادی هر خانه از سنجاقک یا ملخ (جواب ممکن) که مقدار ۱ داشته باشد بدین معنی است که آن ویژگی در مجموعه ویژگی انتخاب قرار دارد. در روش پیشنهادی برای ارزیابی هر جواب ممکن (مجموعه ویژگی) باید میزان برازندگی یا همان تابع برازندگی تخصیص یابد که برای اینکار از روش نزدیک‌ترین همسایه برای ارزیابی مجموعه ویژگی‌های انتخابی با استفاده از معادله زیر محاسبه می‌شود:

$$\text{fitness}_i(\text{selectedFeatures}) = \propto \left(\text{classificatnAccuracy}\ (\text{selectedFeatures})\right) + (1-\propto)\left(\frac{N_t - N_s}{N_t}\right) \quad (1)$$

که در آن N_t و N_s به ترتیب نماینده تعداد کل ویژگی‌ها و تعداد ویژگی‌های انتخاب‌شده و α یک ضریب بین ۰ و ۱می‌باشند.میزان classificatnAccuracy آمده در تابع برازندگی از رابطه زیر بدست می‌آید:

$$\text{classificatnAccuracy} = \frac{TN+TP}{TN+FN+TP+FP} \quad (2)$$

در معادله ۲ صورت کسر شامل نمونه‌هایی است که درست طبقه بندی شده و مخرج کسر شامل کل نمونه‌ها است.

مراحل الگوریتم بهینه سازی پیشنهادی به شرح زیر است:

ورودی: اندازه جمعیت N شامل مقادیر مختلف ۰ و ۱ به تعداد ویژگی‌های مجموعه داده

خروجی: بهترین سنجاقک یا ملخ که شامل بهترین ویژگی‌های انتخابی است

۱-تولید جمعیت اولیه جواب‌ها بصورت تصادفی $X_i=(i=1,2,...,N)$

۲-تا وقتیکه به تعداد تکرار نهایی الگوریتم بهینه سازی نرسیده‌ایم مراحل زیر انجام شود:

۲-۱- محاسبه برازندگی همه جوابها تولید شده (سنجاقک) با معادله ۱

۲-۲- جواب با بهترین برازندگی غذا (X^+) و بدترین برازندگی دشمن (X^-) است.

۲-۳- برای هر سنجاقک/ ملخ مراحل زیر انجام شود:

۲-۳-۱- مقدار پارامترهای w, s, a, c, f ,l و e شعاع همسایگی را بروزرسانی کن.

۲-۳-۲- مقادیر S و A و C و F و E را بروزرسانی شود و مقدار نهایی هر سنجاقک (X_{t+1}) با استفاده از معادلات ۳ الی ۹ .

حرکت تفکیک در سنجاقک ها:

$$S_i = -\sum_{j=1}^{N} X - X_j \quad (3)$$

X موقعیت سنجاقک جاری و X_j موقعیت jامین سنجاقک در همسایگی و N تعداد افراد در همسایگی است.

حرکت هم ترازی در سنجاقک ها:

$$A_i = \frac{\sum_{j=1}^{N} v_j}{N} \quad (4)$$

v_j سرعت jامین سنجاقک که در همسایگی است و Nتعداد همسایه ها می باشد.

حرکت انسجام در سنجاقک ها:

$$C_i = \frac{\sum_{j=1}^{N} X_j}{N} - X \quad (5)$$

X موقعیت سنجاقک جاری و Nتعداد همسایه ها را نشان میدهد و X_j موقعیت jامین سنجاقک در همسایگی است.

جذب شدن به طرف منبع غذایی:

$$F_i = X^+ + X \quad (6)$$

Xموقعیت سنجاقک جاری است و X^+ موقعیت منبع غذا است.

فرار از دشمن :

$$E_i = X^- - X \quad (7)$$

Xموقعیت سنجاقک جاری است و X^- موقعیت دشمن است. به منظور بروزرسانی سنجاقک ها در فضای جستجو ، دو بردار طول گام و بردار موقعیت X در نظر گرفته می شود. بردار طول گام :

$$\Delta X_{t+1} = (sS_i + aA_i + cC_i + fF_i + eE_i) + w\Delta X_t \quad (۸)$$

s ضریب تفکیک است و Si میزان تفکیک سنجاقک iام است و a ضریب هم ترازی است و Ai هم ترازی سنجاقک iام است ، c ضریب انسجام است و Ci انسجام iامین سنجاقک است و fفاکتور تغذیه و Fi منبع غذای iامین سنجاقک و eفاکتور دشمنی است و Eiموقعیت دشمن وw وزن اینرسی و t شمارنده تکرار است.

$$\Delta X_t = X_t + w\Delta X_{t+1} \quad (۹)$$

که در آن tشمارنده تکرار الگوریتم می باشد.

۲-۳-۳- درصورتیکه در همسایگی یک سنجاقک همسایه ای نبود، مقدرا هر سنجاقک با معادلات ۱۰ الی ۱۳ محاسبه خواهد شد.

به منظور بهبود رفتارهای اتفاقی دراکتشاف سنجاقک ها، نیاز است تا در زمانی که راه حلی در همسایگی وجود ندارد در فضای جستجو با یک طول گام تصادفی پرواز کنند. در این موقعیت ، سنجاقک ها با رابطه زیر بروزرسانی میشوند.

$$X_{t+1} = X_t + Levy(d) \times X_t \quad (۱۰)$$

tشمارنده تکرار فعلی است و dبعد های بردار موقعیت است و مقدارlevy از رابطه:

$$Levy(x) = ۰.۰۱ \times \frac{r_1 \times \sigma}{|r_1|^{\frac{1}{\beta}}} \quad (۱۱)$$

r1و r2دو عدد تصادفی در بازه صفر و یک است و بتا یک ثابت و الفا از رابطه:

$$\sigma = \left(\frac{\mathcal{L}(1+\beta) \times \sin\left(\frac{\pi\beta}{2}\right)}{\mathcal{L}\left(\frac{1+\beta}{2}\right) \times \beta \times 2^{\left(\frac{\beta-1}{2}\right)}} \right)^{\frac{1}{\beta}} \quad (۱۲)$$

$$\mathcal{L}(x) = (x-1)! \quad (۱۳)$$

۲-۳-۴- بروزرسانی منطقه آسایش برای ملخ ها با استفاده از رابطه ۱۴

ضریب C منطقه آسایش متناسب با تعداد تکرارها کاهش می یابد و بصورت زیر محاسبه می‌شود:

$$C = Cmax - i\frac{Cmax - Cmin}{l} \quad (۱۴)$$

که در آن cmax بیشترین مقدار، و cmin کمترین مقدار می باشد و I شماره تکرار فعلی را نشان میدهد . و L حداکثر تعداد دفعات تکرار الگوریتم می باشد . در شبیه سازی ها مقدار cmax را ۱ و مقدار cmin را ۰,۰۰۰۰۱ در نظر گرفته شده است .

۲-۳-۵- نرمال سازی فاصله ملخها در بازه [۴،۱] و بروزرسانی مکان هر ملخ با استفاده از رابطه ۱۵

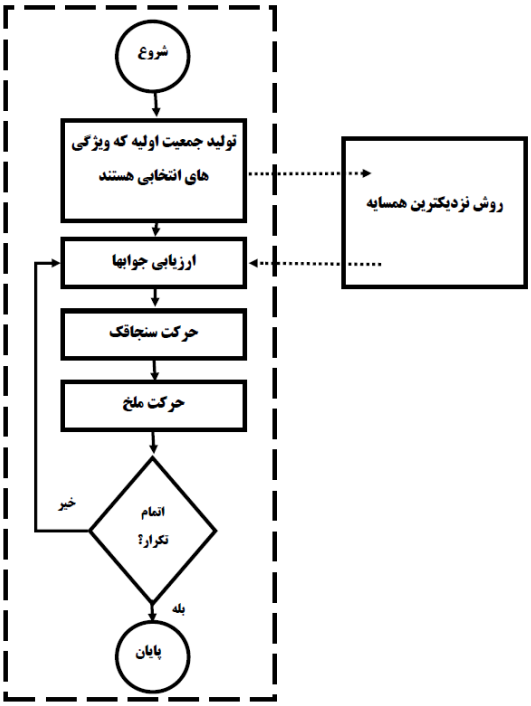
$$x_i^d = c\left(\sum_{\substack{j=1 \\ j \neq i}}^{N} c\frac{Ub_d - lb_d}{2} s\left(\left|x_j^d - x_i^d\right|\right)\frac{x_j - x_i}{d_{ij}} \right) + \widehat{T_d} \quad (۱۵)$$

که در آن Ub_d حد بالا در بعد d ام می باشد و lb_d حد پایین در بعد d ام می باشد و$\widehat{T_d}$ مقدار بعد d ام در هدف (بهترین جوابی که تاکنون دیده شده است) می باشد و c منطقه آسایش است.

۲-۳-۶- بروزرسانی برازندگی همه جواب‌های تولید شده با بهره برداری از میزان صحت دسته بندی رابطه ۱

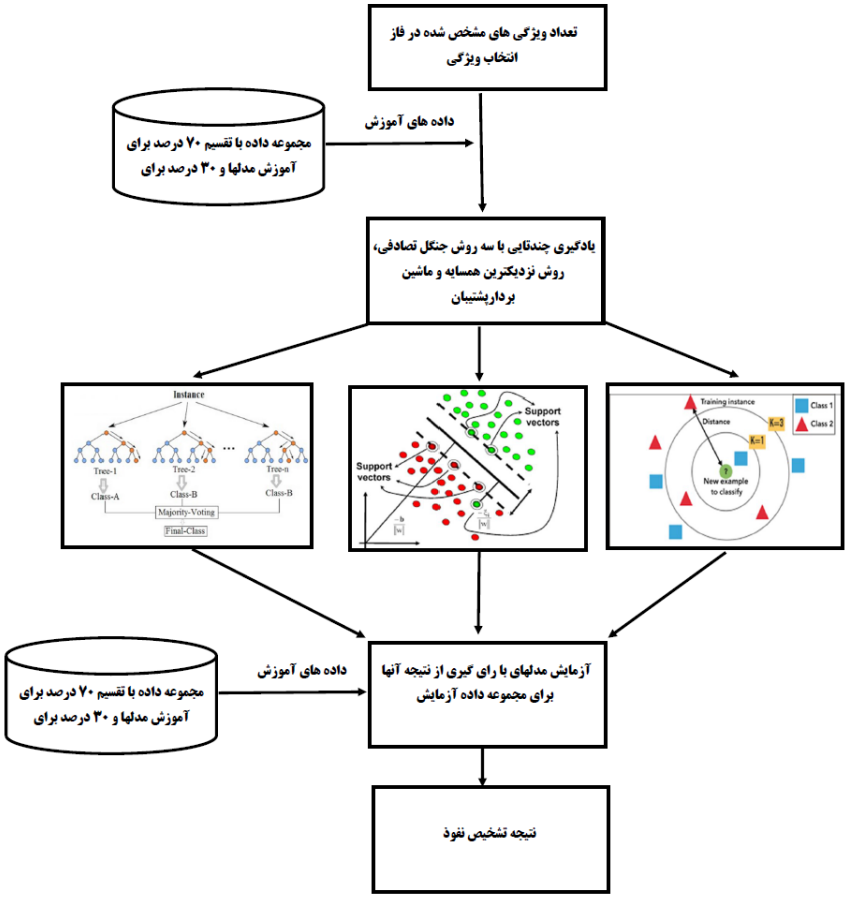
۳- در صورت اتمام تکرارهای الگوریتم بهینه سازی پایان و ارائه موقعیت بهترین جواب در غیر این صورت رفتن به مرحله ۲-۳

نمایی از مراحل در روش پیشنهادی و نحوه انجام کار در شکل ۱ آمده است.



شکل ۱: مراحل الگوریتم بهینه سازی پیشنهادی

مدل طبقه بندی با یادگیری چندتایی شامل جگل تصادفی، روش نزدکیترین همسایه و ماشین بردار پشتیبان انجام می‌شود که در شکل ۲ آمده است.



شکل ۲ : مدل طبقه بندی با یادگیری چندتایی

۵.نتایج

مجموعه داده UNSW-NB۱۵ شامل ۴۱ ویژگی است. آزمایشات، طبقه بندی برای مسئله دوکلاسی و همچنین برای مسئله ده کلاسی انجام شده است. تعداد داده آموزش ۱۷۵۳۴۱ نمونه و داده‌های آزمایش ۸۲۳۳۲ نمونه است که در مجموعه این مجموعه داده دارای ۲۵۷۶۷۳ نمونه است. در داده‌های آزمایش ۴۵۳۳۳ نمونه دارای کلاس حمله و تعداد ۳۷۰۰۰ نمونه کلاس غیرحمله داشته و در داده‌های آموزش ۵۶۰۰۰ نمونه دارای کلاس بدون حمله و ۱۱۹۳۴۲ نمونه دارای کلاس حمله بوده‌اند. این مجموعه داده دارای کلاس ۰ و ۱ برای کلاس غیرحمله و حمله است و همچنین دارای کلاس نوع حمله نیز با ۹ حمله مختلف است که شامل حملات Backdoor، Analysis، Fuzzers، Exploits، DOS، Generic، Reconnaissance، Shellcode و Worms است که در کنار کلاس بدون حمله یعنی normal در مجموعه ۱۰ کلاس دارد.
جهت ارزیابی مدل‌ها جهت تشخیص حمله از سناریوهای مختلف استفاده شده است که شامل:
سناریو اول: انتخاب ویژگی با الگوریتم بهینه سازی ملخ در کنار روش نزدیک‌ترین همسایه در مدل رپر برای دو کلاس (حمله و غیرحمله) و مقایسه سه روش ماشین بردار پشتیبان، جنگل تصادفی و یادگیری چندتایی شامل روش نزدیک‌ترین همسایه، ماشین بردارپشتیبان و جنگل تصادفی است.
سناریو دوم: انتخاب ویژگی با الگوریتم بهینه سازی ملخ در کنار روش نزدیک‌ترین همسایه در مدل رپر برای ده کلاس(چهارحمله و غیرحمله) و مقایسه سه روش ماشین بردار پشتیبان، جنگل تصادفی و یادگیری چندتایی شامل روش نزدیک‌ترین همسایه، ماشین بردارپشتیبان و جنگل تصادفی است.
سناریو سوم: انتخاب ویژگی الگوریتم سنجاقک در کنار روش نزدیک‌ترین همسایه در مدل رپر برای دو کلاس(حمله و غیرحمله) و مقایسه سه روش ماشین بردار پشتیبان، جنگل تصادفی و یادگیری چندتایی شامل روش نزدیک‌ترین همسایه، ماشین بردارپشتیبان و جنگل تصادفی است.
سناریو چهارم: انتخاب ویژگی با الگوریتم سنجاقک در کنار روش نزدیک‌ترین همسایه در مدل رپر برای ده کلاس(چهارحمله و غیرحمله) و مقایسه سه روش ماشین بردار پشتیبان، جنگل تصادفی و یادگیری چندتایی شامل روش نزدیک‌ترین همسایه، ماشین بردارپشتیبان و جنگل تصادفی است.

سناریو پنجم: انتخاب ویژگی با مدل الگوریتم سنجاقک - ملخ در کنار روش نزدیک‌ترین همسایه در مدل رپر برای دو کلاس(حمله و غیرحمله) و مقایسه سه روش ماشین بردار پشتیبان، جنگل تصادفی و ماشین بردار پشتیبان بهبود یافته با مدل پیشنهادی که الگوریتم سنجاقک است.
سناریو ششم: انتخاب ویژگی با مدل الگوریتم سنجاقک - ملخ در کنار روش نزدیک‌ترین همسایه در مدل رپر برای ده کلاس(چهارحمله و غیرحمله) و مقایسه سه روش ماشین بردار پشتیبان، جنگل تصادفی و یادگیری چندتایی شامل روش نزدیک‌ترین همسایه، ماشین بردارپشتیبان و جنگل تصادفی است.
سناریو هفتم: انتخاب ویژگی با مدل الگوریتم جستجوی ممنوع در کنار روش نزدیک‌ترین همسایه در مدل رپر برای دو کلاس(حمله و غیرحمله) و مقایسه سه روش ماشین بردار پشتیبان، جنگل تصادفی و ماشین بردار پشتیبان بهبود یافته با مدل پیشنهادی که الگوریتم سنجاقک است.
سناریو هشتم: انتخاب ویژگی با مدل الگوریتم جستجوی ممنوع در کنار روش نزدیک‌ترین همسایه در مدل رپر برای ده کلاس(چهارحمله و غیرحمله) و مقایسه سه روش ماشین بردار پشتیبان، جنگل تصادفی و یادگیری چندتایی شامل روش نزدیک‌ترین همسایه، ماشین بردارپشتیبان و جنگل تصادفی است.
نتایج مقایسه در سناریو اول برای سه روش ماشین بردار پشتیبان، جنگل تصادفی و یادگیری چندتایی بروی مجموعه داده تشخیص نفوذ در ۲ کلاس با انتخاب ویژگی مدل الگوریتم سنجاقک در کنار روش نزدیکترین همسایه در مدل رپر با ۷ ویژگی انجام شده است. در روش ماشین بردار پشتیبان میانگین نتایج برای دقت دسته بندی عدد ۸۷است و انحراف معیار این نتایج عدد ۰,۱۱ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰,۸۷ و انحراف معیار ۰,۰۰۱ بوده است. در روش جنگل تصادفی میانگین نتایج برای دقت دسته بندی عدد ۹۰,۷ است و انحراف معیار این نتایج عدد ۰,۲۱ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰,۹۰ و انحراف معیار ۰,۰۰۲ است. در روش یادگیری چندتایی (روش نزدیکترین همسایه، جنگل تصادفی، ماشین بردار پشتیبان) میانگین نتایج برای دقت دسته بندی عدد ۹۱,۹۸ است و انحراف معیار این نتایج عدد ۰,۱۷ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰,۹۱ و انحراف معیار ۰,۰۰۲ است.
نتایج مقایسه در سناریو دوم برای سه روش ماشین بردار پشتیبان، جنگل تصادفی و یادگیری چندتایی بروی مجموعه داده تشخیص نفوذ در ۱۰ کلاس با انتخاب ویژگی مدل الگوریتم سنجاقک در کنار روش نزدیکترین همسایه در مدل رپر با ۱۰ ویژگی انجام شده است. در روش ماشین بردار پشتیبان میانگین نتایج برای دقت دسته بندی عدد ۸۶,۴۸ است و انحراف معیار این نتایج عدد ۰,۰۹ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰,۸۶ و انحراف معیار ۰,۰۰۱ است. در روش جنگل تصادفی میانگین نتایج برای دقت دسته بندی عدد ۹۰,۱۱ است و انحراف معیار این نتایج عدد ۰,۳۵ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰,۹۰ و انحراف معیار ۰,۰۰۴ است. در روش یادگیری چندتایی (روش نزدیکترین همسایه، جنگل تصادفی، ماشین بردار پشتیبان) میانگین نتایج برای دقت دسته بندی عدد ۹۱,۳۰ است و انحراف معیار این نتایج عدد ۰,۰۸ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰,۹۱ و انحراف معیار ۰,۰۰۱ است.
نتایج مقایسه در سناریو سوم برای سه روش ماشین بردار پشتیبان، جنگل تصادفی و یادگیری چندتایی بروی مجموعه داده تشخیص نفوذ در ۲ کلاس با انتخاب ویژگی مدل الگوریتم ملخ در کنار روش نزدیک‌ترین همسایه در مدل رپر با ۵ ویژگی انجام شده است. در روش ماشین بردار پشتیبان میانگین نتایج برای دقت دسته بندی عدد ۸۸,۵۹ است و انحراف معیار این نتایج عدد ۰,۲۶ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰,۸۸ و انحراف معیار ۰,۰۰۳ است. در روش جنگل تصادفی میانگین نتایج برای دقت دسته بندی عدد ۹۲,۳۰ است و انحراف معیار این نتایج عدد ۰,۱ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰,۹۲و انحراف معیار ۰,۰۰۲ است. در روش یادگیری چندتایی (روش نزدیکترین همسایه، جنگل تصادفی، ماشین بردار پشتیبان) میانگین نتایج برای دقت دسته بندی عدد ۹۴,۷۹ است و انحراف معیار این نتایج عدد ۰,۱۳ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰,۹۴۷و انحراف معیار ۰,۰۰۳ است.
نتایج مقایسه در سناریو چهارم برای سه روش ماشین بردار پشتیبان، جنگل تصادفی و یادگیری چندتایی شامل روش نزدیکترین همسایه، ماشین بردار پشتیبان و جنگل تصادفی بروی مجموعه داده تشخیص نفوذ در ۱۰ کلاس با انتخاب ویژگی مدل الگوریتم ملخ در کنار روش نزدیک‌ترین همسایه در مدل رپر با ۷ ویژگی انجام شده است. در روش ماشین بردار پشتیبان میانگین نتایج برای دقت دسته بندی عدد ۸۸,۰۵ است و انحراف معیار این نتایج عدد ۰,۱۶ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰,۸۸و انحراف معیار ۰,۰۰۳ است. در روش جنگل تصادفی میانگین نتایج برای دقت دسته بندی عدد ۹۲,۱۵ است و انحراف معیار این نتایج عدد ۰,۰۵ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰,۹۲و انحراف معیار ۰,۰۰۱ است. در روش یادگیری چندتایی (روش نزدیکترین همسایه، جنگل تصادفی، ماشین بردار

پشتیبان) میانگین نتایج برای دقت دسته بندی عدد ۹۴،۵۴ است و انحراف معیار این نتایج عدد ۰،۱ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰،۹۴و انحراف معیار ۰،۰۰۲ است.

همانطور که از نتایج مقایسه در سناریو پنجم برای سه روش ماشین بردار پشتیبان، جنگل تصادفی و یادگیری چندتایی بروی مجموعه داده تشخیص نفوذ در ۲ کلاس با انتخاب ویژگی الگوریتم بهینه سازی ترکیبی سنجاقک و ملخ در کنار روش نزدیکترین همسایه در مدل رپر با ۶ ویژگی انجام شده است. در روش ماشین بردار پشتیبان میانگین نتایج برای دقت دسته بندی عدد ۹۱،۱۱ است و انحراف معیار این نتایج عدد ۰،۱۳ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰،۹۱ و انحراف معیار ۰،۰۰۲ است. در روش جنگل تصادفی میانگین نتایج برای دقت دسته بندی عدد ۹۵،۵۲ است و انحراف معیار این نتایج عدد ۰،۱۱ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰،۹۵۴و انحراف معیار ۰،۰۰۲ است. در روش یادگیری چندتایی (روش نزدیکترین همسایه، جنگل تصادفی، ماشین بردار پشتیبان) میانگین نتایج برای دقت دسته بندی عدد ۹۷،۵۴است و انحراف معیار این نتایج عدد ۰،۱ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰،۹۷ و انحراف معیار ۰،۰۰۱ است.

نتایج مقایسه در سناریو ششم برای سه روش ماشین بردار پشتیبان، جنگل تصادفی و یادگیری چندتایی بروی مجموعه داده تشخیص نفوذ در ۱۰ کلاس با انتخاب ویژگی الگوریتم بهینه سازی ترکیبی سنجاقک و ملخ در کنار روش نزدیک‌ترین همسایه در مدل رپر با ۸ ویژگی انجام شده است. در این آزمایشات با توجه به نتایج بهتر در هسته گوسی برای ماشین بردار پشتیبان، در کلیه آزمایشات از این هسته استفاده شده است. در ادامه نتایج بدست آمده در ۱۰ بار اعتبارسنجی با روش اعتبارسنجی fold-۱۰ آمده است. در روش ماشین بردار پشتیبان میانگین نتایج برای دقت دسته بندی عدد۹۰،۰۴ است و انحراف معیار این نتایج عدد ۰،۱۲ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰،۹ و انحراف معیار ۰،۰۰۲ است. در روش جنگل تصادفی میانگین نتایج برای دقت دسته بندی عدد ۹۵،۴ است و انحراف معیار این نتایج عدد ۰،۰۸ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰،۹۵۵ و انحراف معیار ۰،۰۰۲ است. در روش یادگیری چندتایی (روش نزدیکترین همسایه، جنگل تصادفی، ماشین بردار پشتیبان) میانگین نتایج برای دقت دسته بندی عدد ۹۷،۳ است و انحراف معیار این نتایج عدد ۰،۱۱ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰،۹۷ و انحراف معیار ۰،۰۰۲ است.

نتایج مقایسه در سناریو هفتم برای سه روش ماشین بردار پشتیبان، جنگل تصادفی و یادگیری چندتایی بروی مجموعه داده تشخیص نفوذ در ۲ کلاس با انتخاب ویژگی الگوریتم بهینه سازی جستجوی ممنوع در کنار روش نزدیکترین همسایه در مدل رپر با ۷ ویژگی انجام شده است. در روش ماشین بردار پشتیبان میانگین نتایج برای دقت دسته بندی عدد ۸۵،۸۵ است و انحراف معیار این نتایج عدد ۰،۱۱ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰،۸۵ و انحراف معیار ۰،۰۰۸ است. در روش جنگل تصادفی میانگین نتایج برای دقت دسته بندی عدد ۹۰،۰۵ است و انحراف معیار این نتایج عدد ۰،۱۶ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰،۹۰و انحراف معیار ۰،۰۱۲ است. در روش یادگیری چندتایی (روش نزدیکترین همسایه، جنگل تصادفی، ماشین بردار پشتیبان) میانگین نتایج برای دقت دسته بندی عدد ۹۱،۶۳ است و انحراف معیار این نتایج عدد ۰،۱ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰،۹۱ و انحراف معیار ۰،۰۰۱ است.

نتایج مقایسه در سناریو هشتم برای سه روش ماشین بردار پشتیبان، جنگل تصادفی و یادگیری چندتایی بروی مجموعه داده تشخیص نفوذ در ۱۰ کلاس با انتخاب ویژگی الگوریتم بهینه سازی جستجوی ممنوع در کنار روش نزدیک‌ترین همسایه در مدل رپر با ۹ ویژگی انجام شده است. در این آزمایشات با توجه به نتایج بهتر در هسته گوسی برای ماشین بردار پشتیبان، در کلیه آزمایشات از این هسته استفاده شده است. در ادامه نتایج بدست آمده در ۱۰ بار اعتبارسنجی با روش اعتبارسنجی fold-۱۰ آمده است. در روش ماشین بردار پشتیبان میانگین نتایج برای دقت دسته بندی عدد۸۵،۴۳ است و انحراف معیار این نتایج عدد ۰،۱۳ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰،۸۵ و انحراف معیار ۰،۰۱۲ است. در روش جنگل تصادفی میانگین نتایج برای دقت دسته بندی عدد ۸۹،۳۶ است و انحراف معیار این نتایج عدد ۰،۰۰۸ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰،۸۸ و انحراف معیار ۰،۰۲۲ است. در روش یادگیری چندتایی (روش نزدیک‌ترین همسایه، جنگل تصادفی، ماشین بردار پشتیبان) میانگین نتایج برای دقت دسته بندی عدد ۹۱،۰۵ است و انحراف معیار این نتایج عدد ۰،۲۱ است. برای سطح زیر منحنی roc این میزان بصورت میانگین عدد ۰،۹۱ و انحراف معیار ۰،۰۱۲ است.

۶.نتیجه گیری

در این مقاله، الگوریتم تکاملی بهینه سازی سنجاقک و الگوریتم ملخ بصورت ترکیبی برای انتخاب ویژگی در تشخیص حمله در شبکه‌های اینترنت اشیاء با مجموعه داده‌ای مربوط به این حوزه ارائه شده است، در این خصوص نشان داده شد که تشخیص حمله در تحقیقات سال‌های اخیر در دست تحقیق بوده و محققان تلاش داشته‌اند تا سیستم تشخیص مناسبی برای شناسایی و جلوگیری از حملات ارائه نمایند. بنابراین ارائه سیستم تشخیص نفوذ یا حمله مناسب، ارتباط مستقیم با کارایی یادگیری ماشین و روش دسته بندی دارد که نیازمند فاز انتخاب ویژگی و طبقه بندی است. نتایج در هشت سناریو در جدول ۲ نشان داده شده است.

جدول ۲ : مقایسه نتایج در حالات مختلف در ۸ سناریو مطرح شده

مدل	آزمایشات	روش آماده سازی داده‌ها	نتیجه
مدل طبقه بندی با یادگیری چندتایی (جنگل تصادفی- روش نزدیکترین همسایه – ماشین بردار پشتیبان)	سناریو ۱	انتخاب ویژگی با الگوریتم ملخ با ۲ کلاس و ۷ ویژگی انتخاب شده	۹۱,۹
	سناریو ۲	انتخاب ویژگی با الگوریتم ملخ با ۱۰ کلاس و ۱۰ ویژگی انتخاب شده	۹۱,۳
	سناریو ۳	انتخاب ویژگی با الگوریتم سنجاقک با ۲ کلاس و ۵ ویژگی انتخاب شده	۹۴,۷
	سناریو۴	انتخاب ویژگی با الگوریتم سنجاقک با ۱۰ کلاس و ۷ ویژگی انتخاب شده	۹۴,۵
	سناریو ۵	انتخاب ویژگی با الگوریتم ترکیبی سنجاقک و ملخ با ۲ کلاس و ۶ ویژگی انتخاب شده	۹۷,۵
	سناریو ۶	انتخاب ویژگی با الگوریتم ترکیبی سنجاقک و ملخ با ۱۰ کلاس و ۸ ویژگی انتخاب شده	۹۷,۳
	سناریو ۷	انتخاب ویژگی با الگوریتم جستجوی ممنوع با ۲ کلاس و ۷ ویژگی انتخاب شده	۹۰,۶
	سناریو ۸	انتخاب ویژگی با الگوریتم جستجوی ممنوع با ۱۰ کلاس و ۹ ویژگی انتخاب شده	۹۱
مدل طبقه بندی با جنگل تصادفی	سناریو ۱	انتخاب ویژگی با الگوریتم ملخ با ۲ کلاس و ۷ ویژگی انتخاب شده	۹۰,۷
	سناریو ۲	انتخاب ویژگی با الگوریتم ملخ با ۱۰ کلاس و ۱۰ ویژگی انتخاب شده	۹۰,۱
	سناریو ۳	انتخاب ویژگی با الگوریتم سنجاقک با ۲ کلاس و ۵ ویژگی انتخاب شده	۹۲,۳
	سناریو۴	انتخاب ویژگی با الگوریتم سنجاقک با ۱۰ کلاس و ۷ ویژگی انتخاب شده	۹۲,۱
	سناریو ۵	انتخاب ویژگی با الگوریتم ترکیبی سنجاقک و ملخ با ۲ کلاس و ۶ ویژگی انتخاب شده	۹۵,۵
	سناریو ۶	انتخاب ویژگی با الگوریتم ترکیبی سنجاقک و ملخ با ۱۰ کلاس و ۸ ویژگی انتخاب شده	۹۵,۴
	سناریو ۷	انتخاب ویژگی با الگوریتم جستجوی ممنوع با ۲ کلاس و ۷ ویژگی انتخاب شده	۹۰
	سناریو ۸	انتخاب ویژگی با الگوریتم جستجوی ممنوع با ۱۰ کلاس و ۹ ویژگی انتخاب شده	۸۹,۳

در جدول ۲ مقایسه نتایج در حالات طبقه بندی جنگل تصادفی و یادگیری چندتایی در ۸ سناریو مطرح شده مقایسه شده است. در واقع مقایسه بین جنگل تصادفی در تحقیق [۱] و یادگیری چندتایی نشان می‌دهد که بطور میانگین ۱,۸ درصد بهبود نتایج در مدل پیشنهادی بوده است و همچنین نتیجه می‌شود که مدل ترکیبی سنجاقک – ملخ قوی‌تر از مدل‌های تکی سنجاقک یا ملخ در فاز انتخاب ویژگی است.

۷. منابع

[۱] Nazir, A., & Khan, R. A. (۲۰۲۱). A novel combinatorial optimizatn based feature selectn method for network intrusn detectn. Computers & Security, ۱۰۲, ۱۰۲۱۶۴.

[۲] Li, Y., Ghoreishi, S., & Issakhov, A. (۲۰۲۱). Improving the Accuracy of Network Intrusn Detectn System in Medical T Systems through Butterfly Optimizatn Algorithm. Wireless Personal Communicatns.

[۳] Shima Kamyab, Mahdi Eftekhari. (۲۰۱۵). Feature Selectn using Multimodal Optimizatn Techniques. Neurocomputing,۱-۱۲.

[۴] Seyedali Mirjalili. (۲۰۱۶). Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. Neural Computing and Applications۲(۳):۱–۲۹.

[۵] Seyedali Mirjalili. (۲۰۱۷).Grasshopper Optimisation Algorithm: Theory and application. Advances in Engineering Software۳۰-۴۷.

[۶] Jianing Chen,Jun Wu, Haoran Liang,Shahid Mumtaz,Jianhua Li,Kostromitin Konstantin, Ali Kashif Bashir, and Raheel Nawaz. (۲۰۱۹). Collaborative trust blockchain based unbiased control transfer mechanism for industrial automation. IEEE Transactions on Industry Applications.۱۸-۲۵.

[۷] William Grant Hatcher and Wei Yu. (۲۰۱۸). A survey of deep learning: platforms, applications and emergingres earchtrends. IEEEAccess,۶:۲۴۴۱۱–۲۴۴۳۲.

[۸] Qingchen Zhang, Laurence T Yang, Zhikui Chen, and Peng Li. (۲۰۱۸). A tensortrain deep computation model for industry informatics big data feature learning. IEEE Transactions on Industrial Informatics, ۱۴(۷):۳۱۹۷–۳۲۰۴.

[۹] Fahimeh Farahnakian and Jukka Heikkonen. (۲۰۱۸). A deep auto-encoder based approach for intrusion detection system. In ۲۰۱۸ ۲۰th International Conference on Advanced Communication Technology (ICACT), pages ۱۷۸–۱۸۳.

[۱۰] Nathan Shone, Tran Nguyen Ngoc, Vu Dinh Phai, and Qi Shi. (۲۰۱۸). A deep learning approach to network intrusion detection. IEEE Transactions on Emerging Topics in Computational Intelligence, ۲(۱):۴۱–۵۰.

[۱۱] Mohammed Hasan Ali, Bahaa Abbas Dawood Al Mohammed, Alyani Ismail, and Mohamad Fadli Zolkipli. (۲۰۱۸). A new intrusion detection system based on fast learning network and particle swarm optimization. IEEE Access, ۶:۲۰۲۵۵–۲۰۲۶۱.

[۱۲] Mehdi Moukhafi, Khalid El Yassini, and Seddik Bri. (۲۰۱۸). A novel hybrid ga and svm with pso feature selection for intrusion detection system. Int. J. Adv. Sci. Res. Eng., ۴:۱۲۹–۱۳۴.

[۱۳] R Vijayanand, D Devaraj, and B Kannapiran. (۲۰۱۸). A novel intrusion detection system for wireless mesh network with hybrid feature selection technique based on ga and mi. Journal of Intelligent & Fuzzy Systems, ۳۴(۳):۱۲۴۳– ۱۲۵۰.

[۱۴] L Khalvati, M Keshtgary, and N Rikhtegar. (۲۰۱۸). Intrusion detection based on a novel hybrid learning approach. Journal of AI and data mining, ۶(۱):۱۵۷– ۱۶۲.

[۱۵] Nimbalkar, P., & Kshirsagar, D. (۲۰۲۱). Feature selection for intrusion detection system in Internet-of-Things (IoT). ICT Express, ۷(۲), ۱۷۷–۱۸۱.

[۱۶] Rahman, M. A., Asyhari, A. T., Wen, O. W., Ajra, H., Ahmed, Y., & Anwar, F. (۲۰۲۱). Effective combining of feature selection techniques for machine learning-enabled IoT intrusion detection. Multimedia Tools and Applications.

[۱۷] Parimala, G., & Kayalvizhi, R. (۲۰۲۱). An Effective Intrusion Detection System for Securing IoT Using Feature Selection and Deep Learning. ۲۰۲۱ International Conference on Computer Communication and Informatics (ICCCI).

[۱۸] Sumaiya Thaseen, I., Saira Banu, J., Lavanya, K., Rukunuddin Ghalib, M., & Abhishek, K. (۲۰۲۰). An integrated intrusion detection system using correlation-based attribute selection and artificial neural network. Transactions on Emerging Telecommunications Technologies.

[۱۹] Halim, Z., Yousaf, M. N., Waqas, M., Sulaiman, M., Abbas, G., Hussain, M., … Hanif, M. (۲۰۲۱). An effective genetic algorithm-based feature selection method for intrusion detection systems. Computers & Security, ۱۱۰, ۱۰۲۴۴۸.

[۲۰] Saba, T., Sadad, T., Rehman, A., Mehmood, Z., & Javaid, Q. (۲۰۲۱). Intrusion Detection System Through Advance Machine Learning for the Internet of Things Networks. IT Professional, ۲۳(۲), ۵۸–۶۴.

[۲۱] Li, Y., Ghoreishi, S., & Issakhov, A. (۲۰۲۱). Improving the Accuracy of Network Intrusion Detection System in Medical IoT Systems through Butterfly Optimization Algorithm. Wireless Personal Communications.

[۲۲] Kumar, V., Das, A. K., & Sinha, D. (۲۰۱۹). UIDS: a unified intrusion detection system for IoT environment. Evolutionary Intelligence.

[۲۳] Wu, C., & Li, W. (۲۰۲۱). Enhancing intrusion detection with feature selection and neural network. International Journal of Intelligent Systems, ۳۶(۷), ۳۰۸۷–۳۱۰۵.

[۲۴] Kumar, P., Gupta, G. P., & Tripathi, R. (۲۰۲۱). Toward Design of an Intelligent Cyber Attack Detection System using Hybrid Feature Reduced Approach for IoT Networks. Arabian Journal for Science and Engineering, ۴۶(۴), ۳۷۴۹–۳۷۷۸.

[۲۵] Keserwani, P. K., Govil, M. C., Pilli, E. S., & Govil, P. (۲۰۲۱). A smart anomaly-based intrusion detection system for the Internet of Things (IoT) network using GWO–PSO–RF model. Journal of Reliable Intelligent Environments, ۷(۱), ۳–۲۱.