ارائه سیستم تشخیص حملات سایبری در اینترنت اشیاء صنعتی با استفاده از شبکه عصبی در ترکیب با الگوریتم فرا-ابتکاری گرگ خاکستری و درخت تبرید

سجادعلی محمدی[۱]، محمدقاسمی تادوانی[۱]

[۱]پژوهشگاه جنگ، دانشگاه فرماندهی و ستاد آجا، تهران
{s.alimohamamadi33,m.ghasemi}@casu.ac.ir

چکیده

در این مقاله سعی شده است تا به‌منظور ارائه یک مدل جهت افزایش دقت سیستم تشخیص نفوذ اینترنت اشیای صنعتی در برابر حملات سایبری، روش ترکیبی مبتنی بر الگوریتم فرا-ابتکاری گرگ خاکستری و شبیه‌سازی تبرید با الگوریتم شبکه عصبی مورد مطالعه قرار گیرد. ابتدا داده‌ها پیش‌پردازش و نرمال‌سازی شده، سپس در مرحله بعد با استفاده از الگوریتم‌های فرا-ابتکاری گرگ خاکستری و شبیه‌سازی درخت تبرید و ترکیب آن‌ها با الگوریتم شبکه عصبی، داده‌ها مورد آزمون و ارزیابی قرارگرفته‌اند. بر مبنای نتایج به‌دست‌آمده، استفاده از الگوریتم ترکیبی SA-ANN دارای دقت ۸۸٫۴۰۹۱ درصد و GWO-KNN دارای دقت ۹۳٫۲۲۷۳ در تشخیص حملات سایبری می‌باشند. مشخص گردید که استفاده از الگوریتم ترکیبی GWO-KNN با دقت ۹۳٫۲۲۷۳ درصد ازنظر دقت در انتخاب ویژگی و همچنین میزان تشخیص حملات عملکرد بهتری دارد.

کلمات کلیدی: الگوریتم فرا-ابتکاری گرگ خاکستری، الگوریتم شبیه‌سازی درخت تبرید، الگوریتم شبکه عصبی، سایبر، احراز هویت، سیستم تشخیص نفوذ، اینترنت اشیاء صنعتی.

۱ مقدمه

در بحث شبکه‌هایی که بر پایه اینترنت اشیا (IoT) بنا نهاده شده‌اند، حفظ امنیت و مراقبت از حریم خصوصی از اهمیت زیادی برخوردار است. قطعاً مدیران و طراحان این شبکه‌ها نمی‌خواهند شبکه آن‌ها توسط سایر افراد و سازمان‌های غیرمجاز مورد پایش قرار گرفته و کنترل شبکه را در دست بگیرند؛ زیرا اگر چنین شرایطی ایجاد گردد کاربران غیرمجاز خواهند توانست به اطلاعات کاربران دسترسی پیدا نموده و باعث ایجاد اختلال و یا ازکارافتادن شبکه و دستگاه‌های آن شوند. در شبکه اینترنت اشیاء صنعتی با توجه به محدودیت‌هایی از قبیل تأمین انرژی، اغلب از دستگاه‌هایی استفاده می‌شود که دارای قدرت پردازشی کم، جهت مدیریت منابع تأمین انرژی هستند؛ بنابراین یک شرایط مناسب جهت اجرای حملاتی از قبیل کانال جانبی را در اختیار

افراد غیرمجاز قرار می‌دهد. ازاین‌رو برقراری یک محیط امن به‌منظور جلوگیری از فاش شدن اطلاعات، امری ضروری است.

امنیت و احراز هویت یک هدف اصلی در طراحی دستگاه‌های محاسبات فعلی، ازجمله دستگاه‌های تعبیه‌شده، دستگاه‌های سایبر فیزیکی و دستگاه‌های اینترنت اشیاء صنعتی .است در طول سال‌ها، فن‌های مختلف طراحی، تجزیه–تحلیل، احراز هویت و آزمون امنیت برای دستگاه‌ها ارائه‌شده است. احراز هویت تضمین می‌کند که یک سیستم در حضور خطا هم آیا می‌تواند بدون وقفه کار کند. در اینترنت اشیاء به‌منظور افزایش امنیت و قابلیت اطمینان، احراز هویت مطمئن داده‌ها موردتوجه قرارگرفته است. بر این اساس ایجاد مکانیسم احراز هویت قابل‌اطمینان از الزامات امنیتی این فناوری محسوب می‌شود که با توجه به اهمیت زمان در صنعت بایستی در کمترین زمان ممکن صورت پذیرد. ازاین‌رو، با توجه به پیشرفت‌های روزافزون مبتنی بر حملات مخرب و کاهش امنیت فن‌های تحمل‌پذیر خطا در اینترنت اشیاء، به‌کارگیری و ارائه روش‌هایی که بتوانند احراز هویت را تضمین کرده و یا خطا را به حداقل برسانند، لازم و ضروری بوده که نشان‌دهنده اهمیت بالای پژوهش حاضر است.

ضرورت این تحقیق در ۱) تشخیص حملات سایبری با استفاده از روش‌های مبتنی بر هوش مصنوعی به‌منظور افزایش امنیت و احراز هویت در اینترنت اشیای صنعتی؛ ۲) افزایش امنیت محاسبات مربوط به احراز هویت در اینترنت اشیای صنعتی؛ ۳) افزایش دقت سیستم تشخیص نفوذ با روش ترکیبی مبتنی بر گرگ خاکستری باهدف کاهش ابعاد ویژگی و کم کردن زمان تشخیص حملات؛ ۴) بررسی و طبقه‌بندی آسیب‌پذیری‌ها و چالش‌های اینترنت اشیای صنعتی و ۵) ارائه یک روش جدید مبتنی بر هوش مصنوعی به‌منظور افزایش امنیت و احراز هویت در اینترنت اشیای صنعتی است.

با توجه به بررسی‌های انجام‌شده در حوزه تحقیق و علی‌رغم تحقیقات انجام‌شده درزمینه امنیت و احراز هویت اینترنت اشیای صنعتی، این پژوهش با توجه به موضوع انتخاب‌شده دارای نوآوری و جنبه جدید است.

در این مقاله، به‌منظور افزایش امنیت و احراز هویت در اینترنت اشیای صنعتی باهدف تشخیص حملات سایبری از روش تلفیق الگوریتم‌های فرا ابتکاری گرگ خاکستری و درخت تصمیم با الگوریتم شبکه عصبی برای تعیین طبقه‌بندی ویژگی‌های انتخابی تولیدشده مبتنی بر گرگ خاکستری و شبیه‌سازی تبرید استفاده گردیده است و با توجه به این‌که موضوع زمان در صنعت از اهمیت ویژه‌ای برخوردار است، سعی شده تا با انتخاب کمترین ویژگی از میان ویژگی‌های تعریف‌شده در زمان کمتر و با دقت بیشتری وجود حمله در شبکه شناسایی گردد.

این پژوهش در قالب ۵ بخش نگارش شده است. در بخش اول مقدمه‌ای از پژوهش و الزامات آن اشاره‌شده است. در بخش دوم به بررسی مبانی نظری پژوهش، تعریف مفاهیم موردنیاز در اینترنت اشیای صنعتی، ویژگی‌های این شبکه‌ها و مسئله امنیت در شبکه اینترنت اشیا، بررسی حملات سایبری و ویژگی‌های آن‌ها، تعریف الگوریتم‌های موردنیاز جهت پژوهش و پژوهش‌های مرتبطی که در این زمینه انجام‌شده است بیان می‌شود. در بخش سوم، ما به بیان روش تحقیق مورداستفاده و مراحل انجام روش پیشنهادی که برای دستیابی به اهداف پژوهش شامل بالا بردن ضریب امنیت و احراز هویت که به کمک هوش مصنوعی صورت

می‌گیرد می‌پردازیم. در بخش چهارم، با تجزیه‌وتحلیل و بررسی داده‌های جمع‌آوری‌شده از اعمال روش‌های پیشنهادی، به ارائه نتایج به‌دست‌آمده خواهیم پرداخت و درنهایت، در بخش پنجم، نتیجه‌گیری و پیشنهادهای خود را که از نتایج به‌دست‌آمده در بخش چهار به آن رسیده‌ایم را ارائه می‌کنیم.

همچنین در این پژوهش به دنبال ارزیابی و ارائه یک سیستم تشخصی نفوذ با استفاده از الگوریتم‌های فرا ابتکاری مبتنی بر گرگ خاکستری و شبیه‌سازی درخت تبرید به‌منظور افزایش امنیت و احراز هویت در اینترنت اشیاء صنعتی (IIoT) هستیم.

۲ مبانی نظری و پیشینه های پژوهش

اینترنت اشیاء (IoT) به‌عنوان یک پلتفرم ذخیره و انتقال داده که به کاربران اجازه دسترسی به منابع محاسباتی از راه دور را می‌دهد ظهور کرده است. با توجه به کاربرد اینترنت اشیاء در حوزه های مختلفی همچون صنعت (دشپاند و همکاران، ۲۰۱۶)، آموزش (آجاز محرکان و همکاران، ۲۰۱۷)، سلامت (اسکارپاتو و همکاران، ۲۰۱۷)، کشاورزی (گاندچاوار و کویت کار، ۲۰۱۶) و (شنوی و پینگل، ۲۰۱۸)، تجارت (سینگ، ۲۰۱۵) و (سوهایب و همکاران، ۲۰۱۷) و غیره ظهور نموده است، داده های بسیار زیاد، حجیم و مهمی از این کاربردها تولید میگردد که بایستی در یک شبکه بیسیم منتقل شوند. وجود این داده های حجیم، منجر به بروز چالش هایی مانند پردازش حجم بالایی از داده و ذخیره سازی آن، برقراری امنیت در پردازش، احراز هویت و انتقال شده است. از طرفی، ارتباطات پیچیده شبکه و وجود میزان اطلاعات حجیم در این شبکه‌ها، مدیریت داده‌ها را بسیار دشوار کرده است. ازآنجا که موردحمله قرار گرفتن شبکه ارتباطی اجتناب‌ناپذیر است، درنتیجه شناسایی سریع در کمترین زمان ممکن، احراز هویت و بازیابی اطلاعات در شبکه از قابل‌اطمینان‌ترین مواردی هستند که برای یک عملیات با سرعت‌بالا ضروری می‌باشند [۳].

از طریق اینترنت اشیاء صنعتی (IIoT)، انواع مختلف تجهیزات صنعتی در یک مکان صنعتی هوشمند تعامل خوشه ای را با سایر دستگاه‌ها برقرار می‌کنند و در این مکان که تعامل خوشه‌ای برقرار گردیده داده‌های دستگاه‌ها دیگر مستقل نیستند. از طریق برخورد و تلفیق داده ها، می‌توان دگرگونی و به‌روزرسانی فرآیند تولید هوشمند و شبکه ای را ارتقاء داد. در حقیقت ، روندها و ابتکارات فعلی صنعت باهدف استفاده از اینترنت برای اتصال اشیای غیر مرتبط در صنعت و تحقق چهارمین انقلاب صنعتی است [۴].

باوجود شرایط توسعه و پیشرفت‌های عرصه سایبری، برقراری امنیت کافی برای زیرساخت‌های اینترنت اشیاء که در صنعت مورداستفاده قرار می‌گیرند، روزبه‌روز بر اهمیتشان افزوده می‌شود. برنامه‌هایی که درزمینه صنعت بر پایه اینترنت اشیاء استفاده می‌شوند در برابر هرگونه حمله، اختلال و سرقت اطلاعات، آسیب‌پذیر می‌باشند. اینترنت اشیاء نقش مهمی در اتصال تقریباً همه وسایل (دستگاه‌های موبایل، دوربین‌ها، لوازم‌خانگی، دستگاه‌های بهداشتی، تجهیزات نظامی (ازجمله بی‌سیم‌های نظامی و غیره) به اینترنت، از طریق فناوری‌های مختلف ارتباطی مانند Wi-Fi خواهد داشت. برخی از این اطلاعات که از طریق اینترنت اشیاء به هم مرتبط هستند. این وسایل ممکن است بسیار حساس باشند و حریم خصوصی و امنیت آن‌ها نباید به خطر افتاده باشد. یکی از چالش‌های مهم جامعه بشری با توجه به پیشرفت‌های فنّاوری امروز، حفظ حریم

خصوصی و احراز هویت در تبادل اطلاعات است که سرقت رفتن آن‌ها می‌تواند آسیب‌های جبران‌ناپذیری را به بار آورد [۵].

ازاین‌رو در این پژوهش تلاش می‌گردد که به مسئله امنیت و احراز هویت در اینترنت اشیای صنعتی که از اهمیت بالایی به دلیل ماهیت صنعتی بودن آن برخوردار است، پرداخته شود و با استفاده از فن‌های هوش مصنوعی و یادگیری ماشین در این حوزه، به این مسئله پرداخته شود.

الگوریتم تبرید شبیه‌سازی‌شده یکی از انواع کلی و مؤثر بهینه‌سازی است و معمولاً برای پیدا کردن بهینه‌ی سراسری در حضور تعداد زیادی بهینه‌ی محلی مؤثر است.

اصطلاح تبرید ریشه در مبحث ترمودینامیک دارد که علی‌الخصوص به پروسه گداخته و سرد شدن دوباره‌ی آهن اشاره می‌کند. در مباحث معاملاتی تبرید شبیه‌سازی‌شده، به‌جای انرژی یک ماده، تابع هدف را مورداستفاده قرار می‌دهیم. استفاده از تبرید شبیه‌سازی‌شده بسیار آسان است.

به بیان ساده این الگوریتم یک الگوریتم از خانواده‌ی "Hill Climbing" محسوب می‌شود با این تفاوت که به‌جای انتخاب بهترین حرکت، ما شیوه‌ی انتخاب تصادفی را اختیار می‌کنیم. درصورتی‌که حرکت انتخاب‌شده باعث بهبود معیار موردنظر شود آن را می‌پذیریم در غیر این صورت الگوریتم به هر شکل حرکت بعدی را با احتمالی کوچک‌تر از ۱ انجام می‌دهد. احتمال به‌صورت نمایی به تبعیت از میزان نامطلوب بودن حرکت کاهش می‌یابد.

در یک مسئله‌ی تصمیم‌گیری، نمونه‌های دسته‌بندی به‌صورت مثبت یا منفی برچسب می‌خورد. در این راستا، به‌منظور ارزیابی عملکرد الگوریتم پیشنهادی و بررسی میزان دقت آن از معیارهای ارزیابی استفاده می‌کنیم که در زیر معرفی‌شده‌اند [۶].

TP: تعداد نمونه‌های مثبت که درست تشخیص داده شده‌اند.

TN: تعداد نمونه‌های منفی که درست تشخیص داده شده‌اند.

FP: تعداد نمونه‌های مثبت که به‌اشتباه منفی تشخیص داده شده‌اند.

FN: تعداد نمونه‌های منفی که به‌اشتباه مثبت تشخیص داده شده‌اند.

جدول ۱: ماتریس درهم‌ریختگی

		مقادیر واقعی	
مقادیر پیش‌بینی‌شده	طبقه‌بندی به‌عنوان حمله	حملات عادی	داده‌های عادی
	طبقه‌بندی به‌عنوان حمله	True Positive (TP)	False Positive (FP)
		False Negative (FN)	True Negative (TN)

دقت[1]:

نسبت تعداد نمونه‌های درست طبقه‌بندی‌شده به تمام نمونه‌ها که از رابطه‌ی زیر محاسبه می‌شود:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \times ۱۰۰ \qquad (۱)$$

[1] Accuracy

حساسیت[۲]:

عبارت است از نسبت نمونه‌های مثبت درست طبقه‌بندی‌شده به کل نمونه‌های مثبت موجود:

$$\text{Percision} = \frac{TP}{TP + FP} \times ۱۰۰ \quad (۲)$$

فراخوانی[۳]:

عبارت است از نسبت نمونه‌های مثبت درست طبقه‌بندی‌شده به کل نمونه‌هایی که مثبت هستند. دقت شود که برخی از نمونه‌هایی که مثبت تشخیص داده شده‌اند اشتباه هستند و در مجموعه‌ی FN قرار می‌گیرند.

$$\text{Recall} = \frac{TP}{TP + FN} \times ۱۰۰ \quad (۳)$$

پیشینه تحقیق

علی الحکیم و نیکوقدم [۱] در مقاله خود به ارائه طرحی برای احراز هویت متقابل گمنام و توافق کلید بین دستگاه‌ها در اینترنت اشیاء پرداخته‌اند. در این مقاله در ابتدا به بررسی طرح احراز هویت رستم پور و همکاران پرداخته و اثبات کرده‌اند که طرح آن‌ها به‌درستی کار نمی‌کند و نمی‌تواند طرح مناسبی برای حوزه اینترنت اشیاء باشد و سپس پروتکلی بهبودیافته برای احراز هویت و توافق کلید برای دستگاه‌های اینترنت اشیاء را ارائه داده‌اند. در ادامه اثبات کرده‌اند که طرح پیشنهادی در برابر تمامی حملات مقاوم بوده و نیازهای امنیتی مختلف را تأمین می‌کند. درنهایت امنیت طرح بهبودیافته با ابزار رسمی scyther ارزیابی‌شده است.

نیکویی و چاله‌چاله [۲] در مقاله خود به بررسی و ارزیابی روش‌های احراز هویت برای دسترسی به وسایل خانگی هوشمند در بستر اینترنت اشیاء پرداختند. هدف اصلی این پژوهش معرفی و ارزیابی پروتکل‌های FIDO و OAuth است. این امر به‌منظور احراز هویت کاربری است که درخواست مجوز دسترسی به دستگاه‌های خانگی هوشمند به کمک پروتکل موردنیاز در بستر ابری را دارد. نتایج حاصل از پژوهش نشان داد که استفاده از رمزنگاری دستگاه‌ها به‌جای رمز عبور امنیت ارتباطی بهتری را بین کاربر و سیستم تأمین می‌نماید. این امر به‌منظور جلوگیری از جعل هویت کاربر و دسترسی غیرمجاز به سیستم هوشمند به کمک پروتکل‌های نام‌برده صورت می‌پذیرد.

بخشی و همکاران [۷] در مقاله خود به بررسی مسئله امنیت در اینترنت اشیای صنعتی با در نظرگیری مدل‌های Cisco و Microsoft پرداختند. این مقاله به‌منظور ارائه راهنمایی برای کسانی که می‌خواهند امنیت اینترنت اشیاء را بررسی کرده و در بهبود آن مشارکت داشته باشند، فهرستی از تهدیدها و مسائل امنیتی در لایه ابری اینترنت اشیا که شامل تجمع داده‌ها و سطوح مختلف احراز هویتی است، ارائه می‌دهد. به همین دلیل، از معماری‌های Cisco و Microsoft Azure در اینترنت اشیاء به‌عنوان مدل‌های مرجع استفاده‌شده است. سپس، دو لایه از مدل معماری مرجع سیسکو برای بررسی مسائل امنیتی آن‌ها انتخاب‌شده است.

[2] Percision
[3] Recall

،درنهایت ملاحظات مربوط به مسائل امنیتی و احراز هویت صاحبان داده مورد بررسی و ارزیابی قرارگرفته است.

لین و همکاران [۸] در مقاله خود به بررسی تأیید اعتبار متقابل مبتنی بر بلاک‌چین با سیستم کنترل دسترسی ریز (Bseln) برای صنعت 4.0 پرداخته‌اند. در این مقاله با بهره‌گیری از چارچوب مفهومی به‌طور کارآمد یک کارخانه هوشمند، انعطاف‌پذیر و قابل تنظیم را موردبررسی قرار داده‌اند. بااین‌حال، امنیت ذاتی دستگاه‌ها و شبکه‌های موجود (مستقل) و همچنین شبکه‌هایی را که ممکن است در چنین ادغام‌هایی ایجاد شود، طراحی کارخانه‌های هوشمند را تحت تأثیر قرار داده است؛ بنابراین، یک سیستم مبتنی بر بلاک‌چین برای احراز هویت متقابل امن، BSeIn، برای اجرای سیاست‌های کنترل دسترسی ریز ارائه‌شده است. سیستم پیشنهادی (با امضای ویژگی یکپارچه، رمزگذاری چند گیرنده و کد احراز هویت پیام) برای ارائه تضمین‌های حریم خصوصی و امنیتی مانند احراز هویت بدون نام، قابلیت کنترل و محرمانه بودن طراحی‌شده است. این رویکرد به دلیل استفاده از قرارداد هوشمند مقیاس خوبی دارد و امنیت و عملکرد کارخانه را بهبود داده است.

آلوِز و همکاران [۹] در مقاله خود به ارائه یک مدل ایمن مبتنی رمزگذاری و دستگاه‌های پیشگیری از نفوذ یادگیری ماشین در کنترل‌کننده‌های منطقی قابل‌برنامه‌ریزی پرداخته‌اند. این مقاله، یک طرح جایگزین را با استفاده از یک PLC منبع باز توصیف می‌کند که برای رمزگذاری تمام داده‌هایی که از طریق شبکه ارسال می‌کند، مستقل از پروتکل مورداستفاده، اصلاح‌کرده است. علاوه بر این، یک IPS مبتنی بر یادگیری ماشینی به پشته شبکه PLC اضافه‌شده که مکانیسمی امن در برابر حملات سیل شبکه مانند انکار سرویس (DoS) ارائه می‌کند. نتایج تجربی نشان داد که لایه رمزگذاری و IPS امنیت پیوند بین PLC و نرم‌افزار نظارتی را افزایش داده و از ره‌گیری، تزریق و حملات DoS جلوگیری می‌کند.

تاوالبه و همکاران [۹] در مقاله خود به ارائه یک مدل احراز هویت در سیستم سلامت از راه دور مبتنی بر اینترنت اشیاء پرداختند. در این مقاله، مدل‌های مختلف امنیتی ارائه‌شده برای مسائل امنیتی و احراز هویت معماری اینترنت اشیاء در بستر ابر موردبررسی قرارگرفته و یک مدل امنیتی با استفاده از پلت فرم ارائه‌دهنده خدمات وب آمازون (AWS) با کنترل دسترسی مبتنی بر ویژگی (ABAC) در سطوح مختلف اینترنت اشیاء، بستر ابر و لبه ارائه‌شده است. مدل پیشنهادی تا حد مطلوبی قادر است امنیت و احراز هویت در شبکه اینترنت اشیاء را فراهم نماید.

زو و همکاران [۱۱] در مقاله خود به بررسی استفاده از بلاک‌چین مبتنی بر صنعت 4.0 جهت هوشمند سازی کارخانه‌ها و محیط‌های صنعتی پرداخته‌اند. در این مقاله یک بررسی جامع در مورد بلاک‌چین در صنعت 4.0، کاربردها، معماری، فن‌ها و چالش‌های این حوزه ارائه‌شده است. درنهایت یک معماری مرجع بلاک‌چین برای تولید هوشمند پیشنهاد می‌شود که راهنمای مورداستفاده از فناوری بلاک‌چین در برنامه‌های مختلف کارخانه هوشمند و زنجیره تأمین هوشمند است.

وو و انصاری [۱۲] در مقاله خود یک سیستم امن و قابلیت اعتماد مبتنی بر بلاک‌چین در اینترنت اشیای صنعتی ارائه دادند. در این مقاله یک اینترنت اشیای صنعتی مبتنی بر بلاک‌چین با ارزیابی اعتماد برای هر گره به‌منظور ارزیابی وزن در رأی‌گیری تصمیمات نهایی مربوط به هویت داده‌ها و گره‌ها پیشنهادشده است.

با طراحی مکانیسم رأی‌گیری جدید با ارزیابی مطمئن برای کنترل دسترسی در شبکه اینترنت اشیای صنعتی مبتنی بر بلاک‌چین، سیستم پیشنهادی، احتمال بیشتری برای صدور تأییدیه صحیح حتی باوجود حملات مخرب دارد. نتایج شبیه‌سازی نشان داده است که مکانیسم ارزیابی اعتماد می‌تواند با احتمال اطمینان بالایی، احراز هویت و امنیت در شبکه اینترنت اشیای صنعتی مبتنی بر بلاک‌چین را ارائه نماید.

۳ روش‌شناسی پژوهش

در ابتدا مجموعه داده‌های مربوط به حملات سایبری که هرکدام خود دارای چه ویژگی‌هایی هستند، می‌بایست شناسایی و جمع‌آوری گردد که برای این منظور از مجموعه داده‌های KDDcup99 [۹] استفاده‌شده است؛ بنابراین به‌منظور پیاده‌سازی، ما از این مجموعه داده‌ها استفاده می‌کنیم که دارای ۴۱ ویژگی است که این ویژگی‌ها به سه گروه زیر تقسیم می‌شوند:

۱. ویژگی‌های اساسی
۲. ویژگی‌های محتوایی
۳. ویژگی‌های ترافیک

ویژگی‌های اساسی (مانند protocol-type، duration، service و ...) شامل خصوصیاتی می‌باشند که قادرند از اتصال TCP/IP استخراج گردند. ویژگی‌های محتوایی (مانند Num_failed_logins، logged_in، num_compromised و ...) امکان تشخیص رفتارهای مشکوک در شبکه را فراهم می‌نمایند. ویژگی‌های ترافیکی که خود به دو ویژگی، میزبان یکسان و سرویس یکسان تقسیم می‌شوند نیز بر مبنای تجزیه‌وتحلیل اتصالات برقرارشده در دو رکورد قبل محاسبه می‌شوند. نام این ۴۱ ویژگی و همچنین طبقه هرکدام از ویژگی‌ها در جدول زیر آورده شده است.

همچنین در این پایگاه داده از 22 حمله که در 4 دسته DOS، RU2، LR2 و Probe استفاده‌شده است. همچنین دسته‌های از تارگت‌ها با نام normal نام‌گذاری شده‌اند که به معنای عدم وقوع حمله در شبکه است. ویژگی‌های جدول فوق در پایگاه داده مورداستفاده در قالب a1 تا a41 برچسب‌گذاری گردیده‌اند. از طرفی با توجه به اینکه ویژگی‌های a2 تا a4 به‌صورت رشته یا کلمه می‌باشند و a7 تا a22 نیز مقدار صفر را داشتند و در شبکه تأثیرگذار نمی‌باشند از پایگاه داده حذف گردیدند تا درنهایت 22 ویژگی در پایگاه داده باقی بمانند تا بتوان با استفاده از روش پیشنهادی، بهینه‌ترین جواب را به دست آوریم. نکته‌ای که می‌بایست در روش پیشنهادی پس از به دست آوردن جواب بهینه و انتخاب ویژگی‌ها موردتوجه قرار گیرد، این است که آیا نمونه به‌دست‌آمده جدید حمله است یا خیر. به عبارتی، در این روش هدف به‌صورت یک عدد باینری در نظر گرفته می‌شوند که صفر یعنی حالت نرمال و بدون حمله برقرار است و 1 به معنی رخ دادن حمله است.

دسته‌بندی ویژگی‌ها در این تحقیق بر اساس الگوریتم یادگیری ماشین تحت نظارت یعنی شبکه عصبی است. همچنین از الگوریتم فرا ابتکاری گرگ خاکستری و شبیه‌سازی درخت تبرید برای جواب بهینه استفاده می‌شود.

جدول ۲: مقایسه همبستگی (رگرسیون) توابع آموزش مختلف شبکه عصبی

شماره	نوع تابع	تعریف تابع	ضریب همبستگی
1	trainlm	Levenberg-Marquardt	۰/۹۰۶۵
2	trainscg	Scaled conjugate gradient	۰/۸۵۶۳
3	trainbr	Bayesian regularization	۰/۷۸۵۲
4	trainbfg	BFGS quasi-Newton	۰/۸۳۴۶
5	traincgb	Conjugate gradient backpropagation with Powell-Beale restarts	۰/۹۰۴۱
6	traincgp	Conjugate gradient backpropagation with Polak-Ribiére updates	۰/۸۴۲۶
7	traingda	Gradient descent with adaptive learning rate	۰/۸۴۷۴
8	traingdm	Gradient descent with momentum	۰/۸۹۷۹
9	traingdx	Gradient descent with momentum and adaptive learning rate	۰/۸۹۳۴
10	trainoss	One-step secant	۰/۸۲۱۹

۴ تجزیه و تحلیل داده‌ها

در این مدل شبکه عصبی، الگوریتمی برای یادگیری نظارتی شبکه عصبی با استفاده از گرادیان کاهشی که در آن گرادیان تابع خطا نسبت به وزن‌های شبکه عصبی محاسبه می‌شود، تعیین می‌گردد. موضوعی که در شبکه عصبی حائز اهمیت است، تابع آموزش‌دهنده است. در جدول ۲ مقایسه‌ای بین رگرسیون توابع آموزش گرادیان کاهشی مختلف ارائه‌شده تا مشخص شود کدام تابع کارایی بالاتری دارد.

همان‌طور که در جدول ۲ مشاهده می‌شود، از میان توابع فوق، تابع Levenberg-Marquardtکه در کتابخانه متلب تحت عنوان trainlm شناخته می‌شود دارای بیشترین مقدار همبستگی است و همچنین این تابع سریع‌ترین الگوریتم پس انتشار در جعبه‌ابزار متلب است و به‌عنوان یک الگوریتم نظارت‌شده انتخاب اول، توجیه می‌شود. ازاین‌رو از این تابع به‌منظور آموزش پس انتشار گرادیان کاهشی برای آموزش داده‌ها استفاده می‌شود. قبل از مدل‌سازی طبقه‌بندی حملات سایبری با استفاده از شبکه عصبی پس انتشار، ابتدا نیاز است که تعداد لایه‌ها و تعداد نرون‌های مطلوب جهت مدل‌سازی تعیین گردد. این موضوع در جدول ۳ ارائه‌شده است.

جدول ۳: مقایسه معماری‌های مختلف شبکه عصبی MLP

تعداد لایه	تعداد نرون	Performance	ضریب همبستگی
۱۰	۱۰	۰/۰۰۱۸	۰/۹۶۷۴
۱۰	۲۰	۰/۰۰۵۲	۰/۹۸۶۰
۲۰	۱۰	۰/۰۰۲۱	۰/۹۷۴۶
۱۵	۱۵	۰/۰۰۳۷	۰/۴۸۹۰
۵	۵	۰/۰۰۳۳	۰/۹۶۱۲
۸	۸	۰/۰۰۷	۰/۸۸۲۰
۱۰	۵	۰/۰۰۱۳	۰/۹۷۸۶

بر مبنای معماری‌های اشاره‌شده در جدول ۳، می‌توان به این نتیجه رسید که بهترین معماری برای شبکه

جدول ۴: نتایج حاصل از اعمال الگوریتم ANN از نوع MLP به الگوریتم گرگ خاکستری

پارامتر	مقدار
TP	۲۹۵۱
TN	۱۱۵۱
FP	۱۵۳
FN	۱۴۵
Precision	۹۵/۰۷۰۹ درصد
Recall	95/3165 درصد
Accuracy	۹۳/۲۲۷۳ درصد
F-Measure	۹۵/۱۹۳۵ درصد

جدول ۵: ماتریس درهم‌ریختگی حاصل از اعمال الگوریتم ANN از نوع MLP به الگوریتم گرگ خاکستری

		مقادیر واقعی	
		۰	۱
مقادیر پیش‌بینی‌شده	۰	۶۷/۰۷ درصد	۳/۳ درصد
	۱	۳/۴۸ درصد	۲۶/۱۶ درصد

عصبی MLP به‌منظور تعیین عملکرد طبقه‌بندی حملات سایبری برابر با 10 لایه و 20 نرون است. نتیجه این بررسی در جدول ۴ ارائه‌شده است.

همان‌طور که در جدول ۴ مشاهده می‌شود، اعمال الگوریتم ANN از نوع MLP به الگوریتم گرگ خاکستری در راستای کاهش ویژگی و شناسایی حمله به میزان دقت ۹۳/۲۲۷۳ درصد است.

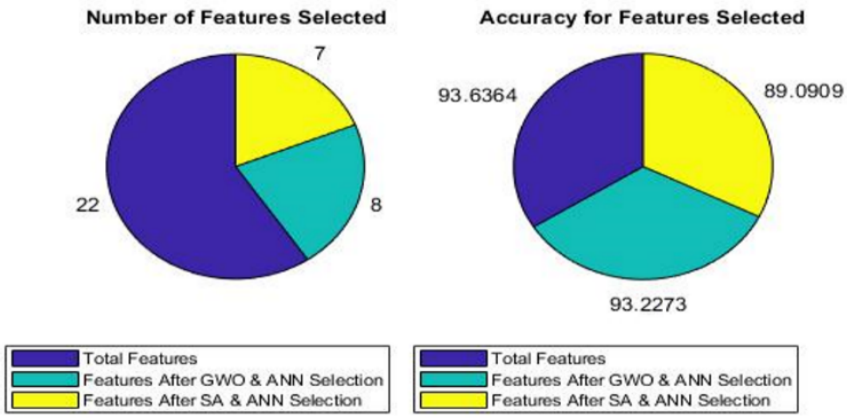
جدول درهم‌ریختگی ۵ نیز نشان می‌دهد که در تلفیق الگوریتم گرگ خاکستری با الگوریتم ANN از نوع MLP از مجموع ۴۴۰۰ داده جهت آزمایش، ۹۳/۲۳ درصد از داده‌ها به نحو صحیح تشخیص داده‌شده‌اند.

همچنین میزان خطای به‌دست‌آمده از این رویکرد برابر با ۰/۰۶۷۷ است. با اعمال الگوریتم ANN از نوع MLP به الگوریتم SA نیز نتایجی به‌دست آمده که مقایسه آن با الگوریتم GWO-ANN در شکل ۱ ارائه شده است.

مقایسه الگوریتم‌های بکار رفته در جدول ۶ ارائه شده است.

جدول ۶: مقایسه دقت الگوریتم پیشنهادی با الگوریتم‌های استفاده‌شده

الگوریتم	دقت	تعداد ویژگی انتخاب شده
GWO-ANN	۹۳/۲۲۷۳	۸
SA-ANN	۸۹/۰۹۰۹	۷



شکل ۱: مقایسه عملکرد الگوریتم GWO-ANN با الگوریتم SA-ANN

۵ نتیجه‌گیری و پیشنهادها

در این پژوهش پایگاه داده مورداستفاده از مرجع KDDCup99 استفاده گردید. این داده‌ها به‌منظور کشف و شناسایی حملات احتمالی صورت گرفته و مخصوص دستگاه‌های اینترنت اشیای صنعتی است. با پیاده‌سازی الگوریتم انتخابی، بررسی چهار معیار مهم ارزیابی یعنی دقت، نرخ تشخیص، فراخوانی و حساسیت انجام پذیرفت. بر مبنای نتایج به‌دست‌آمده مشخص گردید که استفاده از الگوریتم ترکیبی GWO-ANN با دقت ۹۳٫۲۲۷۳ درصد ازنظر دقت در انتخاب ویژگی و همچنین میزان تشخیص حملات عملکرد بهتری دارد.

درمجموع الگوریتم گرگ خاکستری در انتخاب ویژگی بسیار موفق است که با مقایسه آن با الگوریتم شبیه‌سازی تبرید در راستای انتخاب ویژگی، ثابت شد که با انتخاب تعداد ویژگی‌های بهینه می‌تواند نرخ تشخیص حملات را بهبود بخشد.

مراجع

[۱] حیدر محمد علی الحکیم، ش.و م. نیکوقدم، ارائه طرحی برای احراز هویت متقابل گمنام و توافق کلید بین دستگاه ها در اینترنت اشیا، در پنجمین کنگره بین‌المللی مهندسی برق، کامپیوتر و مکانیک. ۱۳۹۹.

[۲] نیکویی، م و چاله‌چاله، ع. بررسی و ارزیابی روش‌های احراز هویت برای دسترسی به وسایل خانگی هوشمند در بستر اینترنت اشیا، هفتمین کنگره ملی تازه یافته‌های مهندسی برق ایران، تهران. ۱۳۹۹.

[3] Fei, X. and G. Tian, Optimization of communication network fault identification based on NB-IoT. Microprocessors and Microsystems, 2021. 80: p. 103531.

[4] Wan, J. et al. A blockchain-based solution for enhancing security and privacy in smart factory. IEEE Transactions on Industrial Informatics, 2019. 15(6): p. 3652-3660.

[5] Sengupta, J. S. Ruj, and S.D. Bit, A comprehensive survey on attacks, security issues and blockchain solutions for IoT and IIoT. Journal of Network and Computer Applications, 2020. 149: p. 102481.

[6] Kamel, S.R. R. YaghoubZadeh, and M. Kheirabadi, Improving the performance of support-vector machine by selecting the best features by Gray Wolf algorithm to increase the accuracy of diagnosis of breast cancer. Journal of Big Data, 2019. 6(1): p. 1-15.

[7] Bakhshi, Z. Balador, A. Mustafa, J. Industrial IoT security threats and concerns by considering Cisco and Microsoft IoT reference models. In 2018 IEEE Wireless Communications and Networking Conference Workshops (WCNCW) (pp. 173-178). IEEE, 2018.

[8] Lin, C. He, D. Huang, X. Choo, K. K. R. Vasilakos, A. V. BSeIn: A blockchain-based secure mutual authentication with fine-grained access control system for industry 4.0. Journal of Network and Computer Applications, 116, 42-52, 2018.

[9] Tawalbeh, M. M. Quwaider, and A.T. Lo'ai. Authorization model for IoT healthcare systems: case study. in 2020 11th IEEE International Conference on Information and Communication Systems (ICICS). 2020.

[10] Zuo, Y. Making smart manufacturing smarter–a survey on blockchain technology in Industry 4.0. Enterprise Information Systems, 1-312, 2020. Wu, D. Ansari, N. A trust-evaluation-enhanced blockchain-secured industrial IoT system. IEEE Internet of Things Journal, 8(7), 5510-5517, 2020.

[11] KDD Cup 1999. Available on: http://kdd.ics.uci.edu/databases/kddcup 99/kddcup99.html, Ocotber 2007.

[12] Electronics and Applications (ICIEA) (pp. 419-423). IEEE.