Journal of Information and Communication Technology

Volume 16, Issue 61-62, Autumn and Winter 2024, pp. 218-228

Providing an Intrusion Detection System in the Industrial Internet of Things Using the Gray Wolf Algorithm

Sajjad Alimohammadi[1*], Mohammad Fathi[2]

[1] Faculty Member of Electrical Engineering, University of Kurdistan, Sanandaj, Iran
[2] Faculty Member of Electrical Engineering, University of Kurdistan, Sanandaj, Iran

Received: 19 April 2024, Revised: 26 June 2024, Accepted: 07 October 2024
Paper type: Research

Abstract

Security is a main goal in the design of the Industrial Internet of Things network. Due to the ever-increasing developments in the Internet of Things, it is necessary to use new methods in detecting active network attacks. In this article, an intrusion detection system for industrial Internet of Things is proposed. This system uses a combination of gray wolf meta-heuristic (GWO) and decision tree (DT), nearest neighbor (KNN) and artificial neural network (ANN) classification algorithms. First, the data is pre-processed and then normalized, in the next step, data feature extraction is performed using the gray wolf algorithm to extract its independent and effective features. Then it is trained and finally evaluated using classification algorithms. The obtained results show that the use of the combined GWO-ANN algorithm with 93.22% accuracy has a better performance in detecting attacks. Also, the ANN algorithm is more accurate than the DT and KNN algorithms in combination with the GWO algorithm.

Keywords: Intrusion Detection System, Gray Wolf Algorithm, Industrial Internet of Things.

[*] Corresponding Author's email: sajad.alimohamadi@uok.ac.ir

دوفصلنامه
فناوری اطلاعات و ارتباطات ایران

سال شانزدهم، شماره‌های ۶۱ و ۶۲، پاییز و زمستان ۱۴۰۳، صفحه ۲۱۸ الی ۲۲۸

ارائه سیستم تشخیص نفوذ در اینترنت اشیاء صنعتی با استفاده از الگوریتم گرگ خاکستری

سجاد علی محمدی[1*]، محمد فتحی[2]،

[1] دانشکده مهندسی برق، دانشگاه کردستان، سنندج، ایران

[2] دانشکده مهندسی برق، دانشگاه کردستان، سنندج، ایران

تاریخ دریافت: ۱۴۰۳/۰۱/۳۱ تاریخ بازبینی: ۱۴۰۳/۰۴/۰۶ تاریخ پذیرش: ۱۴۰۳/۰۷/۱۶

نوع مقاله: پژوهشی

چکیده

امنیت یک هدف اصلی در طراحی شبکه اینترنت اشیاء صنعتی است. با توجه به پیشرفت‌های روزافزون در اینترنت اشیاء لازم است از روش‌های جدید در تشخیص حملات فعال شبکه استفاده شود. در این مقاله یک سیستم تشخیص نفوذ برای اینترنت اشیاء صنعتی پیشنهاد شده است. این سیستم از ترکیب الگوریتم‌های فراابتکاری گرگ خاکستری (GWO) و الگوریتم‌های طبقه‌بندی درخت تصمیم (DT)، نزدیک‌ترین همسایه (KNN) و شبکه عصبی مصنوعی (ANN) استفاده می‌کند. ابتدا داده‌ها پیش‌پردازش و سپس نرمال‌سازی شده، در مرحله بعد استخراج ویژگی داده‌ها با استفاده از الگوریتم گرگ خاکستری برای استخراج ویژگی‌های مستقل و مؤثر آن انجام می‌شود. سپس با استفاده از الگوریتم‌های طبقه‌بندی، آموزش و درنهایت ارزیابی می‌شود. نتایج به‌دست‌آمده نشان می‌دهد که استفاده از الگوریتم ترکیبی GWO-ANN با دقت ۹۳/۲۲ درصد در میزان تشخیص حملات عملکرد بهتری دارد. همچنین الگوریتم ANN نسبت به الگوریتم‌های DT و KNN در تلفیق با الگوریتم GWO دارای دقت بالاتری است.

کلیدواژگان: سیستم تشخیص نفوذ، الگوریتم گرگ خاکستری، اینترنت اشیاء صنعتی.

* رایانامهٔ نویسنده مسؤول: sajad.alimohamadi@uok.ac.ir

۱- مقدمه

با گسترش روزافزون شبکه‌های مبتنی بر اینترنت اشیاء چالش‌های این حوزه روزبه‌روز در حال افزایش است. اینترنت اشیاء در زمینه‌های مختلف مانند صنعت، پزشکی و غیره کاربرد دارد. حفظ امنیت یکی از چالش‌های مهم شبکه‌های اینترنت اشیاء است. شبکه‌های مبتنی بر اینترنت اشیاء همواره با محدودیت‌هایی از قبیل تأمین انرژی، قدرت پردازش کم و غیره روبرو می‌باشند. ازاین‌رو، تأمین امنیت شبکه‌های اینترنت اشیاء از اهمیت بالایی برخوردار است زیرا مهاجم بانفوذ به شبکه قادر به انجام اعمال مخربی مانند سرقت اطلاعات یا اختلال در شبکه است. مهاجم با کشف یک آسیب‌پذیری و بهره‌برداری از آن به اهداف مخرب خود دست می‌یابد. با توجه به حساسیت شبکه اینترنت اشیاء صنعتی[۱](IIoT) در محیط‌های مختلف مانند صنایع پتروشیمی، خودروسازی و غیره نیاز به ایجاد یک شبکه امن لازم و ضروری است.

اینترنت اشیاء امکان جمع‌آوری، انتقال و پردازش داده‌ها را در یک شبکه صنعتی فراهم می‌کند. از طرفی، پیچیدگی تجهیزات یک شبکه صنعتی و نیاز به هماهنگی تمام اجزاء شبکه مبتنی بر اینترنت اشیاء دشوار است. ازاین‌رو، وجود آسیب‌پذیری در یک شبکه IIoT اجتناب‌ناپذیر است. شناسایی سریع حملات فعال خواهد توانست از دسترسی مهاجم به منابع و اختلال در عملکرد شبکه جلوگیری کند [۱].

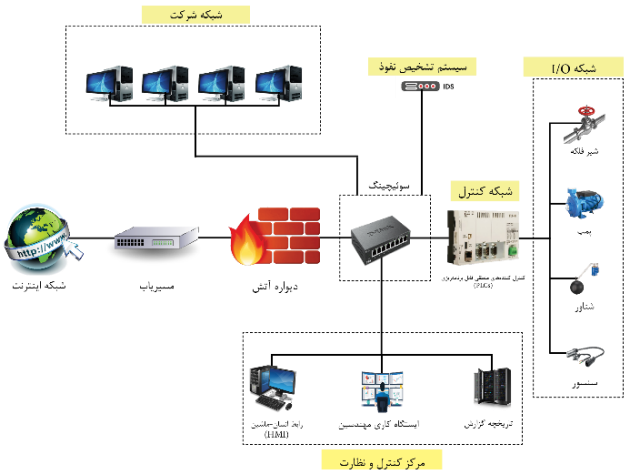
با گسترش فناوری‌های مختلف و ظهور مفاهیم جدید مانند کلان داده، محاسبات ابری، سیستم فیزیکی سایبری[۲] و غیره، صنایع به سطح بالایی از توسعه ارتقا یافته‌اند. درنتیجه، با استفاده از IIoT دستگاه‌های صنعتی هوشمند به وجود آمدند که به دنبال تحقق تولید هوشمند هستند. IIoT ارتباط انواع مختلف تجهیزات صنعتی موجود در یک محیط صنعتی هوشمند را به‌صورت تعاملی با سایر دستگاه‌ها برقرار می‌کنند. در یک محیط تعاملی داده‌ها دیگر مستقل از یکدیگر نیستند. با ارتباط تجهیزات صنعتی به یکدیگر می‌توان دگرگونی در فرآیند تولید هوشمند ایجاد کرد. در حقیقت، روند فعلی صنعت به دنبال استفاده از اینترنت برای اتصال تجهیزات صنعتی و تحقق چهارمین انقلاب صنعتی است [۲]. در دهه اخیر، مهندسی تولید، هوشمندسازی و روش‌های محاسباتی به دنبال تکامل IIoT هستند. اجزای محاسباتی و ارتباطی ادغام‌شده در دستگاه‌های کنترل صنعتی و دستگاه‌های تولید در کارخانه‌ها، به‌طور پیوسته در حال افزایش است. کنترل‌کننده‌های منطقی قابل‌برنامه‌ریزی با دستگاه‌های فیزیکی سایبری پیشرفته‌تر جایگزین شده‌اند که سبب شده تا دستگاه‌های نهفته[۳] به‌صورت آزادانه فرآیندهای فیزیکی را برنامه‌ریزی و کنترل کنند. در دستگاه‌های کنترل صنعتی، مفهوم امنیت به‌طور عمومی تقریباً همان معنای ایمنی را دارد، یعنی حفاظت از انسان، محیط‌زیست و ماشین‌ها در برابر پیامدهای ناشی از خرابی دستگاه‌ها است [۳].

امروزه برقراری امنیت سایبری در زیرساخت‌های اینترنت اشیاء که در صنعت مورد استفاده قرار می‌گیرند دارای اهمیت ویژه‌ای است. برنامه‌هایی که در صنعت بر پایه اینترنت اشیاء استفاده می‌شوند در برابر هرگونه حمله، اختلال و سرقت اطلاعات، آسیب‌پذیر می‌باشند. این برنامه‌ها ممکن است بسیار حساس و حفظ حریم خصوصی و امنیت آن‌ها دارای اهمیت بالای باشد [۴].

ازاین‌رو در این مقاله تلاش می‌گردد که به مسئله امنیت در IIoT که از اهمیت بالایی برخوردار است، پرداخته شود و با استفاده از روش‌های مبتنی بر یادگیری ماشین به بررسی این مسئله پرداخته شود. در این مقاله اثر ترکیب الگوریتم فرا ابتکاری گرگ خاکستری با مدل‌های یادگیری ماشین مرسوم ازجمله درخت تصمیم، شبکه عصبی و k- نزدیک‌ترین همسایه برای سیستم تشخیص نفوذ در اینترنت اشیا مورد ارزیابی قرار می‌گیرد.

دستگاه‌های کنترل صنعتی بخشی جدایی‌ناپذیر از زیرساخت‌های حیاتی هستند و مدت‌زمان طولانی در نظارت بر ماشین‌آلات صنعتی و فرآیندها مورد استفاده قرار گرفته‌اند. این دستگاه‌ها بر دستگاه‌های صنعتی موجود در شرکت نظارت کرده، با آن‌ها تعامل داشته و به‌صورت بلادرنگ نسبت به جمع‌آوری و تجزیه‌وتحلیل داده‌ها اقدام کرده و همچنین ثبت تمام رویدادهایی که در دستگاه‌های صنعتی اتفاق می‌افتد را انجام می‌دهند. سیستم کنترل نظارت و اکتساب داده[۴](SCADA) بزرگ‌ترین زیرمجموعه از دستگاه‌های کنترل صنعتی است. این سیستم یک رابط کاربری گرافیکی را از طریق رابط انسانی-دستگاه HMI[۵] فراهم می‌کند. HMI مشاهده وضعیت سیستم، برقرار ارتباط با دستگاه‌های IIoT و دریافت هشدار که نشان‌دهنده رفتارهای غیرطبیعی است را برای کاربران امکان‌پذیر می‌سازد. طرح کلی دستگاه‌های SCADA به همراه یک سیستم تشخیص نفوذ در شکل ۱ نشان داده‌شده است [۵].

[1] Industrial Internet of Things
[2] Cyber Physical System
[3] Embedded Devices
[4] Supervisory Control and Data Acquisition
[5] Human Machine Interface



شکل ۱. طرح کلی سیستم SCADA

همان‌گونه که در شکل ۱ نشان داده‌شده است ساختار SCADA شامل چهار زیرسیستم مختلف است. شبکه I/O، مرکز کنترل و نظارت، شبکه کنترل و شبکه شرکت. شبکه I/O شامل دستگاه‌های IIoT مستقرشده متشکل از حس‌گرها و محرک‌ها در فرآیند صنعتی است. کنترل نظارتی زیرسیستم مسئول اصلی امنیت، کنترل و نظارت بر دستگاه‌های IIoT است. شبکه کنترل شامل کنترل‌کننده‌های منطقی قابل‌برنامه‌ریزی (PLCs)[1] است که به‌طور مستقیم دریافت داده و مدیریت فرآیندهای فیزیکی را انجام می‌دهد. ازآنجاکه حس‌گرها و محرک‌ها نمی‌توانند به‌طور مستقیم ارتباط برقرار کنند، PLC ها برای جمع‌آوری داده‌ها و ارسال دستورات به محرک‌ها استفاده می‌شوند. درنهایت، شبکه شرکت شامل سرورها، کامپیوترها و دیگر کاربران متصل به شبکه برای سایر خدمات عمومی مانند انتقال فایل، میزبانی وب‌سایت، ایمیل، برنامه‌ریزی منابع و غیره است [۵]. در این مقاله، یک سیستم تشخیص نفوذ (IDS)[2] مبتنی بر یادگیری ماشین که با بررسی بسته‌ها و ارتباطات ایجادشده به دنبال کشف حملات فعال در شبکه‌های IIoT است ارائه می‌شود. توضیح اینکه سیستم تشخیص نفوذ معمولاً در لبه شبکه داخلی بعد از دیواره آتش و در اتصال با سوییچ هسته شبکه است.

این مقاله در ادامه به‌صورت زیر سازمان‌دهی می‌شود. در بخش دوم تحقیقات مرتبط با امنیت در شبکه IIoT موردبررسی قرار می‌گیرد. در بخش سوم طرح پیشنهادی یک IDS در شبکه IIoT ارائه می‌شود. در بخش چهار ارزیابی از روش پیشنهادی ارائه و در بخش پنجم جمع‌بندی و پیشنهادهایی برای تحقیقات آینده آورده می‌شود.

۲- مرور ادبیات

معماری امنیتی IIoT با ساختار امنیتی در شبکه‌های سازمانی متفاوت است. معماری IIoT دارای یک ساختار سلسله مراتبی از فناوری‌های مختلف شبکه و پروتکل‌های ارتباطی که دستگاه‌های سخت‌افزاری را با نرم‌افزار کنترلی به هم متصل می‌کند تشکیل‌شده است. مسائل امنیتی مطرح‌شده در IIoT بیشتر مربوط به طراحی پروتکل‌های ارتباطی ناامن هستند که باوجود چنین معماری، پیشگیری از حملات سایبری در شبکه IIoT را پیچیده می‌کند [۶]. در ادامه با بررسی تحقیقات پیشین روش‌های مختلف امن‌سازی شبکه‌های IIoT را بررسی خواهیم کرد.

آرزم و براتی [۷] در مقاله خود به ارائه روشی برای احراز هویت در اینترنت اشیاء مبتنی بر موقعیت گیرنده‌های Wi-Fi و فناوری زنجیره بلوکی پرداختند. روش پیشنهادی، شامل دو مرحله است. در مرحله اول، هر دستگاه اثبات‌کننده، مقداردهی اولیه می‌شود که برای اجرای پروتکل تصدیق و احراز هویت به آن نیاز است. در مرحله دوم، برای احراز هویت و حفظ حریم خصوصی، پروتکل زنجیره بلوکی استفاده می‌شود. با توجه به نتایج حاصل‌شده با تزریق خطاهای ناشی از تهدیدات امنیتی، نشان داد که تهدیدات امنیتی گذرا باعث بروز پنج خطا و تزریق تهدیدات امنیتی پایدار باعث بروز ۱۰۰ خطا در سیستم می‌شوند. بدون اعمال تمهیدات امنیتی، ۱۰۵ نمونه از داده‌های اینترنت اشیاء مورد تهدید قرار گرفت. با توجه به تعداد کل داده‌های قرائت‌شده در پنج ثانیه شبیه‌سازی که برابر با ۵۰۰ نمونه است، میزان داده‌های مورد تهدید قرارگرفته برابر با ۲۱ درصد است.

سید ترابی و پهلوان [۸]، در مقاله خود به رمزگذاری در احراز هویت دستگاه‌های اینترنت اشیاء پرداختند. در این مقاله روشی جهت حفظ امنیت اطلاعات مبتنی بر احراز هویت به کمک رمزنگاری و امضای دیجیتال و همچنین شکست فایل به بخش‌های کوچک‌تر ارائه‌شده است. آزمایش‌ها بر اساس تعداد مختلف بسته‌های انتقالی برای روش پیشنهادی نشان می‌دهد که این روش به دلیل استفاده از رمزنگاری تصادفی باعث شده است که معیار عملکرد خوبی به ازای بسته‌های مختلف ازنظر معیارهای نظیر مصرف انرژی، دقت و غیره داشته باشد. همچنین بر اساس روش پیشنهادی ضمن احراز هویت کاربران، سرویس محرمانگی و دقت اطلاعات نیز به نحوی مناسب در روش پیشنهادی مورد توجه قرارگرفته شده است.

عیسی‌لو و سلیمانی [۹] در مقاله خود به ارائه روشی کم‌بار برای

[1] Programmable Logic Controllers

[2] Intrusion Detection System

احراز هویت اشیاء در اینترنت اشیاء پرداختند. در این مقاله به بررسی مدل احراز هویت EAP پرداخته و با سبک‌بار نمودن زیرشاخه‌ای از آن یعنی EAP-PSK روشی کم‌بار برای احراز هویت اشیاء در اینترنت اشیاء را ارائه کرده‌اند. روش ارائه‌شده با ابزار شبیه‌سازی Cooja مورد آزمایش قرارگرفته و در مقایسه با روش پایه در شاخص‌های مانند مجموع تعداد پیام‌های احراز هویت، زمان احراز هویت، نرخ موفقیت احراز هویت و نرخ مصرف انرژی بهبودیافته است.

IDS یک چهارچوب دفاعی امیدوارکننده در امنیت سایبری است که برای تشخیص حملات در شبکه استفاده می‌شود. در سال‌های اخیر این مکانیسم در شبکه‌های IIoT نیز موردتوجه قرارگرفته است. در ادامه سیستم‌های تشخیص نفوذ ارائه‌شده برای شبکه IIoT که مبتنی بر الگوریتم‌های یادگیری ماشین هستند موردبررسی قرار می‌گیرند. هاب و همکاران، یادگیری زیر فضای تصادفی و الگوریتم نزدیک‌ترین همسایه[1] (KNN) را برای بهبود دقت تشخیص KNN ترکیب کرده‌اند. نتایج به‌دست‌آمده بهبود عملکرد KNN را نشان می‌دهد [۱۰].

ردی و همکاران، یک مطالعه تجربی با استفاده از چندین روش یادگیری ماشین برای تشخیص ناهنجاری و طبقه‌بندی حمله ارائه کردند. آن‌ها یک چهارچوب یادگیری با استفاده از XGBoost ارائه نمودند. نتایج نشان می‌دهد که XGBoost یک رویکرد امیدوارکننده برای تشخیص نفوذ در طبقه‌بندی حملات است [۱۱].

زانگ و همکاران [۱۲] با استفاده از روش‌های دسته‌بندی KNN و RF[2] سیستم تشخیص نفوذ و قابل‌اعتماد برای تشخیص حملات مردمیانی و منع سرویس توزیع‌شده را در مقاله خود ارائه نموده‌اند.

ایگنر و همکاران [۱۳] در مقاله خود به شناسایی حملات مردمیانی در شبکه‌های کنترل صنعتی پرداختند. روش پیشنهادی از تشخیص ناهنجاری با توسعه مدلی از رفتار عادی شبکه سیستم کنترل صنعتی استفاده می‌کند. یک سیستم صنعتی ساده، متشکل از حس‌گرها و محرک‌ها، با کنترل‌کننده‌هایی که به‌طور گسترده در صنعت استفاده می‌شوند، راه‌اندازی شده است. سپس از یک رویکرد یادگیری ماشین بر اساس الگوریتم k- نزدیک‌ترین همسایه با واگرایی برگمن برای تعریف مدلی از رفتار عادی (معتبر) استفاده‌شده است. پس‌ازآن حملات مردمیانی علیه سیستم انجام‌شده و رفتار آن در حین حمله با مدل رفتاری معتبر مقایسه گردیده است. نتایج نشان می‌دهد که رویکرد اتخاذ شده قادر به شناسایی حملات با دقت رضایت بخشی است.

زولانواری و همکارانش [۱۴] در مقاله خود بر روی استفاده از یادگیری ماشین (ML) برای افزایش امنیت دستگاه‌های صنعتی اینترنت اشیا (IIoT) تمرکز دارد. به دلیل پیامدهای مخرب احتمالی حملات سایبری، بر نیاز حیاتی به ایمن‌سازی دستگاه‌های IIoT تأکید می‌کند. نویسندگان بر اهمیت ML و تجزیه‌وتحلیل داده‌های بزرگ در تجزیه‌وتحلیل و ایمن‌سازی فناوری IIoT تأکید می‌کنند و اینکه چگونه این تکنیک‌ها می‌توانند امنیت سیستم‌های IIoT را بهبود بخشند. این مطالعه شامل مروری بر پروتکل‌های رایج IIoT و آسیب‌پذیری‌های مرتبط با آن‌ها، ارزیابی آسیب‌پذیری سایبری و بررسی ادبیات راه‌حل‌های تشخیص نفوذ موجود با استفاده از مدل‌های ML است. علاوه بر این، نویسندگان یک مطالعه موردی شامل اجرای یک بستر آزمایشی در دنیای واقعی برای انجام حملات سایبری و طراحی یک سیستم تشخیص نفوذ[3] (IDS) با استفاده از تشخیص ناهنجاری مبتنی بر ML ارائه می‌کنند. این مطالعه همچنین عملکرد IDS را از طریق معیارهای نماینده برای ارزیابی اثربخشی روش‌ها ارزیابی می‌کند. پس‌زمینه مقاله مروری جامع از آسیب‌پذیری‌ها در سیستم‌های IIoT، استفاده از ML برای تشخیص نفوذ و توسعه یک بستر آزمایشی در دنیای واقعی برای انجام حملات سایبری و طراحی IDS ارائه می‌کند. هدف این مطالعه پرداختن به چالش‌های امنیتی در سیستم‌های IIoT و نشان دادن پتانسیل راه‌حل‌های مبتنی بر ML در افزایش امنیت این سیستم‌ها است.

یعقوب کایوده سعید و سانجی میسرادر [۱۵] در مقاله خود تحت عنوان «مدل‌های یادگیری گروهی مبتنی بر بهینه‌ساز گرگ خاکستری برای تشخیص نفوذ در اینترنت اشیا» از تکنیک‌های بهره اطلاعاتی IG[4] و PCA[5] برای انتخاب و استخراج مجموعه‌ای از ویژگی‌ها بهره می‌برند و سپس از روش گرگ خاکستری برای بهینه‌سازی پارامترهای یک مدل ماشین ترکیبی از چهار ماشین دیگر شامل درخت تصمیم، جنگل تصادفی، شبکه عصبی و k نزدیک‌ترین همسایه استفاده می‌کنند.

مائولی و همکاران [۱۶] در این مقاله خود، مسائل امنیتی اینترنت اشیاء صنعتی را از سه جنبه بررسی کرده‌اند شامل: (۱) تهدیدات امنیتی و مکانیسم‌های حمله آن‌ها برای نشان دادن آسیب‌پذیری اینترنت صنعتی اشیاء. (۲) روش‌های تشخیص نفوذ از دیدگاه

[1] K Nearest Neighbor
[2] Random Forest
[3] Intrusion Detection Systems
[4] Information Gain (IG)
[5] Principal Component Analysis (PCA)

شناسایی حمله و (۳) برخی از استراتژی‌های دفاعی به طور جامع خلاصه شده است. در پایان این مقاله چندین نکته پایانی و دستورالعمل برای تحقیقات در آینده ارائه شده است.

فتانی و همکاران [۱۷] در این مقاله خود یک مدل IDS جدید مبتنی بر ترکیب روش‌های یادگیری عمیق و بهینه‌سازی را پیشنهاد کرده‌اند. در این مقاله ابتدا یک روش استخراج ویژگی مبتنی بر CNN توسعه داده شده است. سپس، یک روش انتخاب ویژگی جدید بر اساس نسخه اصلاح شده بهینه‌سازی رشد[1] (GO) به نام MGO استفاده می‌شود. آن‌ها از الگوریتم بهینه‌سازی نهنگ[2] (WOA) برای تقویت روند جستجوی GO استفاده کرده‌اند. ارزیابی و مقایسه‌های گسترده‌ای برای ارزیابی کیفیت روش پیشنهادی با استفاده از مجموعه داده‌های عمومی محیط‌های ابری و اینترنت اشیا (IoT) انجام شده است. آن‌ها به این نتیجه رسیده‌اند که MGO در تمام مقایسه‌های تجربی بهتر از چندین روش قبلی ذکر شده در مقاله عمل کرده است.

سلطانی و همکاران [۱۸] در نیز مقاله خود تحت عنوان «تشخیص نفوذ قوی برای ارتباطات شبکه در اینترنت اشیا: یک رویکرد یادگیری ماشین ترکیبی» یک رویکرد یادگیری ماشین ترکیبی (با استفاده از k-نزدیک‌ترین همسایگان و جنگل‌های تصادفی به عنوان طبقه‌بندی‌کننده نظارت‌شده) برای افزایش دقت سیستم‌های تشخیص نفوذ و به حداقل رساندن خطر حملات احتمالی ارائه کرده‌اند. همچنین در این مقاله، از الگوریتم‌های حذف به عقب و تحلیل تفکیک خطی برای کاهش ویژگی و کاهش هزینه‌های محاسباتی استفاده شده است. عملکرد مدل پیشنهادی در چارچوب برنامه‌نویسی پایتون، با استفاده از مجموعه داده‌های CICIDS 2017، NSL-KDD و TON-IoT ارزیابی شده.

۳- روش پیشنهادی

در این بخش سیستم تشخیص نفوذ پیشنهادی برای تشخیص حملات سایبریIIoT ارائه شده است که در این روش از ترکیب الگوریتم فرا ابتکاری گرگ خاکستری (GWO) [۱۹] و الگوریتم‌های یادگیری ماشین KNN، DT[3] و ANN[4] استفاده‌شده است [۲۰]. در ادامه ابتدا مجموعه داده استانداردی که برای تشخیص حملات در IIoT استفاده می‌شود معرفی خواهد شد سپس به معرفی روش پیشنهادی خواهیم پرداخت.

۳-۱- مجموعه داده

مجموعه داده‌های متفاوت مانند IoTID20، NSL-KDD، UNSW-NB15، CICIDS2017 و Kddcup۹۹ برای بررسی امنیت IIoT ارائه‌شده است. در این مقاله با توجه به تعداد ویژگی‌ها، تعداد نمونه‌های زیاد و گستردگی استفاده در تحقیقات پیشین از مجموعه داده KDDcup۹۹ استفاده می‌شود [۲۱]. این مجموعه داده دارای ۴۱ ویژگی برابر جدول ۱ است که این ویژگی‌ها به ویژگی‌های اساسی، ویژگی محتوایی و ویژگی ترافیک تقسیم‌بندی می‌شوند.

جدول ۱. دسته‌بندی ویژگی‌ها

شماره	نام ویژگی	تعریف ویژگی	طبقه اختصاص یافته
۱	duration	طول اتصال	ویژگی‌های پایه‌ای
۲	protocol_type	نوع پروتکل (TCP، UDP ...)	
۳	service	سرویس مقصد (ftp، telnet ...)	
۴	Flag	وضعیت اتصال	
۵	src_bytes	شماره B از مبدأ تا مقصد	
۶	dst_bytes	شماره B از مقصد تا مبدأ	
۷	land	اگر آدرس مبدأ و مقصد همان land=1/ اگرنه، ۰ است	
۸	wrong_fragment	تعداد قطعات اشتباه	
۹	urgent	تعداد بسته‌های فوری	
۱۰	hot	تعداد نشانگرهای داغ	ویژگی‌های محتوایی
۱۱	num_failed_logins	تعداد تلاش‌های ناموفق برای ورود	
۱۲	logged_in	اگر وارد شوید=۱/اگر ورود ناموفق بود ۰	
۱۳	num_compromised	تعداد کشورهای در معرض خطر	
۱۴	root_shell	اگر یک مفسر دستور با حساب ریشه در حال اجرا1=shell / اگرنه، ۰ باشد	
۱۵	su_attempted	اگر دستور su تلاش شد1=su tryed / اگرنه، سپس ۰ (ورود موقت به سیستم با سایر اطلاعات کاربری کاربر)	
۱۶	num_rot	تعداد دسترسی‌های ریشه	
۱۷	num_file_creations	تعداد عملیاتی که فایل‌های جدید ایجاد می‌کند	
۱۸	num_shells	تعداد مفسران فرمان فعال	
۱۹	num_access_files	تعداد عملیات ایجاد فایل	
۲۰	num_outbound _cmds	تعداد دستورات خروجی در یک جلسهftp	
۲۱	is_host_login	اگر لاگین در لیست ورود میزبان باشد، 1=loginاست/اگرنه، ۰ است	
۲۲	is_guest_login	اگر مهمان وارد سیستم شده باشد،	

[1] Growth Optimizer (GO)
[2] Whale Optimization Algorithm
[3] Decision Tree
[4] Artificial Neural Network

		ورود مهمان=۱/ اگرنه، ۰ است	
۲۳	Count	تعداد اتصالات به همان میزبان به‌عنوان اتصال فعلی در یک بازه زمانی معین	ویژگی‌های ترافیکی
۲۴	srv_count	تعداد اتصالات به همان سرویس بااتصال فعلی در یک بازه زمانی معین	
۲۵	serror_rate	%از اتصالات با خطاهایSYN	
۲۶	srv_serror_rate	%از اتصالات با خطاهایSYN	
۲۷	rerror_rate	%از اتصالات با خطاهایREJ	
۲۸	srv_rerror_rate	%از اتصالات با خطاهایREJ	
۲۹	same_srv_rate	%از اتصالات به همان سرویس	
۳۰	diff_srv_rate	%از اتصالات به سرویس‌های مختلف	
۳۱	srv_diff_host _rate	%از اتصالات به هاست‌های مختلف	
۳۲	dst_host_count	تعداد اتصالات به همان مقصد	
۳۳	dst_host_ srv_count	تعداد اتصالات به یک مقصد که از همان سرویس استفاده می‌کنند	
۳۴	dst_host_same_ srv_rate	%از اتصالات به یک مقصد که از یک سرویس استفاده می‌کنند	
۳۵	dst_host_diff_ srv_rate	%از اتصالات به هاست‌های مختلف در یک سیستم	
۳۶	dst_host_same_ src_port_tate	%از اتصالات به یک سیستم با پورت منبع یکسان	
۳۷	dst_host_srv_ diff_host_rate	%از اتصالات به یک سرویس که از میزبان‌های مختلف می‌آید	
۳۸	dst_host_serro r_rate	%از اتصالات به یک میزبان با خطایS0	
۳۹	dst_host_srv_ serror_rate	%از اتصالات به هاست و سرویس مشخص‌شده با خطایS0	
۴۰	dst_host_rerrr_ rate	%از اتصالات به هاست و سرویس مشخص‌شده با خطایS0	
۴۱	dst_host_srv_ rerror_rate	%از اتصالات به هاست و سرویس مشخص‌شده با خطای S0	

ویژگی‌های اساسی مانند Protocol_type، Duration و Service و... شامل خصوصیاتی می‌باشند که از سرآیند یک اتصال TCP/IP استخراج می‌گردند. 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 تجزیه و تحلیل تعداد ارتباط برقرارشده و بر اساس دو نمونه قبلی محاسبه می‌شوند. این مجموعه داده شامل ۲۲ حمله در چهار دسته DOS، U2R، R2L و Probe است. داده‌های ترافیک نرمال شبکه دارای برچسب Normal هستند. در جدول ۲ جزئیات مجموعه داده KDDcup۹۹ بیان‌شده است.

جدول ۲. جزئیات مجموعه داده KDD Cup ۹۹

دسته‌بندی حملات	نام حمله	تعداد رکورد
DoS (Denial of Service)	back, land, neptune, pod, smurf, teardrop	۵۴۲۹۴
R2L (Remote to Local)	ftp_write, guesspasswd, imap, multihop, phf, spy, warezlient, warezmaster	۵۳۹۴
U2R (User to Root)	buffer_overflow, loadmodule, perl, rootkit	۱۲۶
Probe	ipsweep, nmap, portsweep, satan	۱۳۸۵۹

۳-۱-۱- پیش‌پردازش داده‌ها

ویژگی‌های مجموعه داده در جدول ۱ گروه‌بندی شده‌اند. با توجه به اینکه ویژگی‌ها دارای مقادیر غیر عددی می‌باشند بایستی این ویژگی‌ها را به روش نگاشت همانی با اعداد حقیقی جایگزین کرد. همچنین گروه ویژگی‌های دارای مقدار صفر در تصمیم‌گیری بی‌تأثیر بوده‌اند و از مجموعه داده حذف می‌شوند. مجموعه داده، ممکن است برای یک نمونه خاص، یک ویژگی که لازم باشد را نداشته باشد و باعث ایجاد خطا در روند برنامه شود؛ بنابراین حداکثر مقدار ممکن آن ویژگی در جایگاه خالی قرار خواهد گرفت. مقادیر ویژگی‌های مجموعه داده در دامنه متفاوتی قرار دارند و احتمال بروز خطا در نتایج افزایش می‌یابد، بنابراین ۲۲ ویژگی باقیمانده با استفاده از رابطه (۱) در مقیاس [۰،۱] نرمال‌سازی می‌شود.

$$\bar{X} = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (۱)$$

در این رابطه X معرف داده‌های اولیه، $\bar{X}$ مقدار ویژگی نرمالیزه شده، X_{min} و X_{max} به ترتیب کمترین و بیشترین مقدار ویژگی در مجموعه داده هستند. برچسب‌های داده‌ها به دو گروه حمله و نرمال تقسیم می‌شود. IDS حملات مختلف را در یک گروه کلی در نظر خواهد گرفت. پس از آماده‌سازی مجموعه داده‌ها به دو بخش آموزش شامل ۸۰ درصد نمونه‌ها و ارزیابی شامل ۲۰ درصد نمونه‌ها تقسیم می‌شوند. داده‌های بخش آموزش برای تولید مدل ماشین مورداستفاده قرار می‌گیرند. داده‌های بخش ارزیابی برای بررسی دقت مدل تولیدشده با داده‌های آموزش مورد استفاده قرار می‌گیرد.

۳-۲- استخراج ویژگی

بعد از آماده‌سازی مجموعه داده، استخراج ویژگی با استفاده از الگوریتم گرگ خاکستری[1](GWO) بر روی داده‌ها انجام می‌شود [۲۲]. هرکدام از دسته گرگ‌ها به‌عنوان اجزای یک راه‌حل در نظر گرفته می‌شوند. همهٔ گرگ‌های مهاجم که در یک راه‌حل می‌باشند به‌عنوان یک واحد کلی در نظر گرفته می‌شوند که به سمت یک

[1] Gery Wolf Algorithm

هدف حرکت می‌نمایند. نحوه شکار گرگ خاکستری به این شکل است که دسته گرگ‌ها به چهار نوع تقسیم می‌شوند:

الف- گرگ آلفا که رهبر گروه است و همیشه بهترین موقعیت را نسبت به هدف دارد.

ب- گرگ بتا که پایین‌تر از گرگ آلفا و بالاتر از سایرین است و تابع دستورات آلفا است و درواقع نقش مشاور را دارد.

ج- گرگ امگا که از دودسته بالا پایین‌تر و از سایرین بالاتر است و نقش محافظ در گله را بر عهده دارد.

د- گرگ‌های پیر و سالخورده دسته که گرگ دلتا نامیده می‌شوند.

در این الگوریتم بهترین موقعیت برای راه‌حل مسئله به ترتیب مربوط به گرگ آلفا، گرگ بتا، گرگ دلتا و سایر راه‌حل‌ها نیز مربوط به گرگ امگا است. الگوریتم گرگ خاکستری بیشتر بر اساس موقعیت گرگ آلفا، بتا و دلتا جستجو انجام می‌دهد. در مدل‌سازی ریاضی برای محاسبه هدف از معادله (۲)، (۳) و (۴) استفاده می‌شود.

$$\vec{X}(t+1) = \vec{X}_p(t) - \vec{A}.\vec{D} \quad (۲)$$

$$\vec{D}(t) = \left|\vec{C}.\vec{X}_p(t) - \vec{X}(t)\right|, t = 1,2,\dots,t_{max} \quad (۳)$$

$$\vec{A} = 2\vec{a}.\vec{r_1} - \vec{a}, \vec{C} = 2.\vec{r}_2 \quad (۴)$$

در این معادلات t تکرار فعلی را بیان می‌کند، $\vec{X}(t+1)$ موقعیت گرگ خاکستری در تکرار t+۱ ، $\vec{X}_p(t)$ موقعیت هدف در تکرار t، بردار $\vec{D}$ فاصله هر کدام از گرگ‌های آلفا، بتا یا دلتا از گرگ امگا یا سایر شکارچیان گله است، بردارهای ضرایب $\vec{A}$ و $\vec{C}$ با توجه به تعریف، از $\vec{r}_1$ و $\vec{r}_2$ که بردارهای تصادفی بوده و به‌صورت تصادفی بین [۰،۱] هستند، به دست می‌آیند. همچنین a یک مؤلفه کاهشی است که به‌صورت خطی در بازه [۰،۲] کاهش پیدا می‌کند. با توجه به اینکه هیچ اطلاعاتی در مورد موقعیت هدف خود نداریم موقعیت هدف همان موقعیت گرگ آلفا در نظر گرفته می‌شود. به‌روزرسانی موقعیت گرگ خاکستری طبق روابط (۵)، (۶)، (۷) و (۸) انجام می‌شود [۱۵].

$$\vec{X}_1(t) = \vec{X}_\alpha(t) - \vec{A}_1.\vec{D}_\alpha(t), \vec{D}_\alpha(t) = \left|\vec{C}_1.\vec{X}_\alpha(t) - \vec{X}(t)\right| \quad (۵)$$

$$\vec{X}_2(t) = \vec{X}_\beta(t) - \vec{A}_2.\vec{D}_\beta(t), \vec{D}_\beta(t) = \left|\vec{C}_2.\vec{X}_\beta(t) - \vec{X}(t)\right| \quad (۶)$$

$$\vec{X}_3(t) = \vec{X}_\delta(t) - \vec{A}_3.\vec{D}_\delta(t), \vec{D}_\delta(t) = \left|\vec{C}_3.\vec{X}_\delta(t) - \vec{X}(t)\right| \quad (۷)$$

$$\vec{X}(t+1) = \frac{\left(\vec{X}_1(t)+\vec{X}_2(t)+\vec{X}_3(t)\right)}{3} \quad (۸)$$

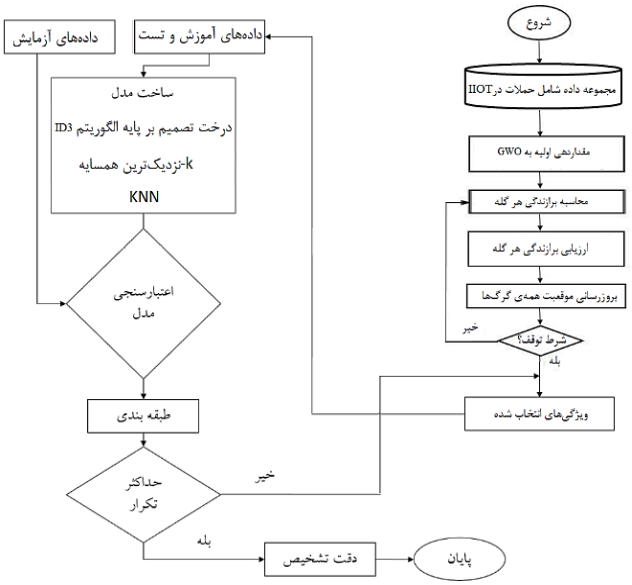
$\vec{X}_\alpha(t)$، $\vec{X}_\beta(t)$ و $\vec{X}_\delta(t)$ موقعیت گرگ‌های آلفا، بتا و دلتا در زمان t و X_1، X_2، X_3 موقعیت هرکدام از گرگ‌ها بر اساس این سه گرگ است [۱۸].

با استفاده از الگوریتم گرگ خاکستری زیرمجموعه‌ای از ویژگی‌ها که به بهینه‌ترین مقدار منجر می‌شوند، انتخاب می‌گردد. تابع برازندگی برای انتخاب ویژگی از هر دسته گرگ طبق رابطه (۹) تعریف می‌شود. در رابطه (۹)، $|n|$ تعداد کل ویژگی‌ها و $|s|$تعداد ویژگی‌های انتخاب‌شده است. پارامتر Accuracy درصد دقت و مقدار پارامترهای δ و ρ ثابت هستند و مقدار آن‌ها به ترتیب برابر با ۰،۷ و $\delta - 1$ است.

$$Fitness = \delta.Accuracy + \rho.\frac{|n|-|s|}{|n|} \quad (۹)$$

استخراج ویژگی با ابعاد بالا امکان ایجاد راه‌حل‌های جدید را فراهم می‌کند اما در هنگام محاسبات چالش‌های ایجاد می‌شود. چالش‌های مانند اینکه در اکثر مواقع همه ویژگی‌ها برای یافتن جواب مناسب که در آن‌ها نهفته است مهم نیستند. راهکار مناسب برای انتخاب ویژگی، حذف آن قسمت از زیرمجموعهٔ ویژگی‌های ورودی است که اطلاعات کمی در خود دارند. این امر باعث ایجاد ویژگی‌های باقیمانده که در افزایش کارایی طبقه‌بندی مؤثر هستند و باعث بالا رفتن دقت می‌شود می‌گردند.

هنگامی‌که با استفاده از الگوریتم گرگ خاکستری ویژگی‌های جدید از مجموعه داده استخراج‌شده، با استفاده از الگوریتم‌های یادگیری ماشین مانند DT، KNN و ANN مدل تصمیم‌گیری را آموزش می‌دهیم. در این مدل نتایج بررسی هرکدام از بسته‌ها با استفاده از مدل به دو گروه کلی ترافیک حمله و نرمال تقسیم‌بندی می‌شوند. نتایج به‌دست‌آمده در بخش بعدی مورد ارزیابی و مقایسه قرار می‌گیرد. فلوچارت اجرایی این روش در شکل ۲ ارائه شده است.



شکل ۲. فلوچارت پیشنهادی

۴- ارزیابی نتایج

در این بخش نتایج الگوریتم‌ها به‌صورت جداگانه ارائه می‌شود. برای آموزش مدل پیشنهادی از سخت‌افزار با مشخصات پردازنده Intel(R) Core(TM) i5 - U۴۵۰۰ و رم ۸,۰۰ GB استفاده‌شده است. با استفاده از ماتریس درهم‌ریختگی[1] نتایج پیش‌بینی‌شده توسط مدل‌های DT، KNN و ANN توسط داده‌های استخراج‌شده از الگوریتم گرگ خاکستری بررسی می‌شود. ماتریس درهم‌ریختگی یک روش ارزیابی بازدهی مدل‌های طبقه‌بندی یادگیری ماشین است که به‌صورت جدول ۳ است و میزان انطباق مقادیر پیش‌بینی‌شده با مقادیر واقعی را نشان می‌دهد.

جدول ۳. ماتریس درهم‌ریختگی

		مقادیر واقعی	
		۰	۱
مقادیر پیش‌بینی‌شده	۰	TN	FP
	۱	FN	TP

برای به دست آوردن معیارهای Accuracy، Precision، Recall و F-measure با استفاده از ماتریس درهم‌ریختگی مطابق روابط (۱۰)، (۱۱)، (۱۲) و (۱۳) عمل می‌شوند.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (۱۰)$$

$$Precision = \frac{TP}{TP+FP} \quad (۱۱)$$

$$Recall = \frac{TP}{TP+FN} \quad (۱۲)$$

$$F-measure = \frac{2\times Precision\times Recall}{Precision\times Recall} \quad (۱۳)$$

Accuracy یا دقت یکی از معیارهای ارزیابی یک مدل است که نشان‌دهنده تعداد پیش‌بینی‌های صحیح است. Precision یا صحت نشان‌دهنده درصد حملاتی است که مدل به‌صورت صحیح نسبت به تمام حملات تشخیص داده است. Recall نسبت هشدارهای صحیح به تمام هشدارهای سیستم را نشان می‌دهد؛ و F-measure یک معیار هماهنگ است که به‌عنوان میانگین هندسی Precision و Recall مورداستفاده قرار می‌گیرد.

حال با توجه به معیارهای ارزیابی نتایج به‌دست‌آمده در هرکدام از الگوریتم‌ها را نشان خواهیم داد. برای الگوریتم GWO-DT نتایج مطابق جدول ۴ به‌دست‌آمده است. جدول ۴ نشان می‌دهد مدل GWO-DT میزان دقت ۹۰/۷۲۷۳ درصد و میزان خطای ۰/۰۹۲۷ با کاهش تعداد ویژگی‌های موجود در مجموعه داده توانسته وقوع یا عدم وقوع حمله را به شکل مطلوبی پیش‌بینی نماید؛ اما وجود خطا ۹/۲۷ درصد نیازمند بهبود مدل پیش‌بینی است.

جدول ۴. نتایج GWO-DT

پارامتر	مقدار
TP	۲۸۳۲
TN	۱۱۶۰
FP	۲۰۳
FN	۲۰۵
Precision	۹۳/۳۱۱۳ درصد
Recall	۹۳/۲۴۹۹ درصد
Accuracy	۹۰/۷۲۷۳ درصد
F-Measure	۹۳/۲۸۰۶ درصد

مرحله بعد، به‌جای الگوریتم DT، از الگوریتم KNN جهت ارائه مدل استفاده می‌شود. نتایج این الگوریتم، در جدول ۵ ارائه‌شده است.

جدول ۵. نتایج GWO-KNN

پارامتر	مقدار
TP	۲۶۰۴
TN	۱۰۸۶
FP	۵۲۸
FN	۱۸۲
Precision	۸۳/۱۴۱۸ درصد
Recall	۹۳/۴۶۷۳ درصد
Accuracy	۸۳/۸۶۳۶ درصد
F-Measure	۸۸/۰۰۲۷ درصد

همان‌طور که در جدول ۵ مشاهده می‌شود، اعمال الگوریتم KNN به الگوریتم GWO در راستای کاهش ویژگی و شناسایی حمله به میزان دقت ۸۳/۸۶۳۶ درصد رسیده که در مقایسه با الگوریتم GWO-DT که ۹۰/۷۲۷۳ درصد است، در حدود ۷ درصد کمتر است. از طرفی میزان خطای به‌دست‌آمده توسط الگوریتم تلفیقی GWO-KNN برابر با ۰/۱۶۱۳ است.

در ادامه، از الگوریتم شبکه عصبی MLP استفاده می‌شود. در این مدل شبکه عصبی، الگوریتمی برای یادگیری نظارتی شبکه عصبی با استفاده از گرادیان کاهشی که در آن خطا نسبت به وزن‌های شبکه عصبی محاسبه می‌شود، تعیین می‌گردد. در جدول ۶ معماری شبکه عصبی با توجه به بیشترین ضریب همبستگی و بالاترین کارایی ازلحاظ تعداد لایه و تعداد نرون، برای دستیابی به‌دقت بالاتر در آموزش بررسی‌شده است.

بر مبنای معماری‌های اشاره‌شده در جدول ۶، معماری مناسب برای شبکه عصبی مصنوعی به‌منظور تعیین عملکرد طبقه‌بندی حملات برابر با ۱۰ لایه و ۲۰ نرون برای هر لایه است؛ زیرا دارای بیشترین

[1] Confusion matrix

کارایی و بالاترین ضریب همبستگی در بین سایر معماری‌های بررسی‌شده است. در جدول ۷ متغیرهای اولیه شبکه عصبی ارائه‌شده است. با ارائه مدل شبکه عصبی مصنوعی مناسب، مدل GWO-ANN ارائه می‌شود؛ که نتایج آن در جدول ۸ ارائه‌شده است.

همان‌طور که در جدول ۸ مشاهده می‌شود، اعمال الگوریتم ANN به الگوریتم گرگ خاکستری در راستای کاهش ویژگی و شناسایی حمله به میزان دقت ۹۳/۲۲۷۳ درصد رسیده که در مقایسه با الگوریتم GWO-KNN که ۹۰/۷۲۷۳ درصد است، حدود ۳ درصد افزایش داشته است.

جدول ۶. مقایسه معماری‌های مختلف شبکه عصبی MLP

معماری		کارایی	ضریب همبستگی
تعداد لایه	تعداد نرون		
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
۵	۱۰	۰/۰۰۷	۰/۹۰۸۰
۱۲	۸	۰/۰۰۴۳	۰/۸۳۸۸
۱۰	۱۵	۰/۰۰۳۰	۰/۹۵۸۰
۹	۲۵	۰/۰۰۱۹	۰/۹۶۸۶

جدول ۷. پارامترهای شبکه عصبی

پارامتر	مقدار
Traning	Levenberg-Marquardt
Performance	Mean Squared Error
Epoch	۱۰۰۰

جدول ۸. نتایج GWO-ANN

پارامتر	مقدار
TP	۲۹۵۱
TN	۱۱۵۱
FP	۱۵۳
FN	۱۴۵
Precision	۹۵/۰۷۰۹ درصد
Recall	۹۵/۳۱۶۵ درصد
Accuracy	۹۳/۲۲۷۳ درصد
F-Measure	۹۵/۱۹۳۵ درصد

به‌منظور مقایسه بهتر الگوریتم‌های بکار رفته تمامی نتایج ازنظر دقت و زمان حل در جدول ۹ ارائه می‌شود.

همچنین در جدول ۱۰ مقایسه‌ای بین میزان خطای به‌دست‌آمده از تشخیص حملات سایبری توسط الگوریتم گرگ خاکستری در ترکیب با طبقه‌بندی‌های DT، KNN و ANN ارائه‌شده است.

بر اساس نتایج به‌دست‌آمده ازنظر دقت در انتخاب ویژگی‌ها و میزان تشخیص حملات الگوریتم GWO-ANN دارای بهترین عملکرد است. پس از الگوریتم ANN، الگوریتم DT در مرتبه دوم قرار می‌گیرد و الگوریتم KNN از دقت کمتری نسبت به دو الگوریتم ANN و DT برخوردار است.

در میان سه الگوریتم بررسی‌شده GWO-ANN دارای بهترین دقت و کمترین خطا است ازاین‌رو نتایج به‌دست‌آمده الگوریتم را با کارهای مشابه که توسط عیسی، اورمان و بریفکانی [۲۳] که از ترکیب الگوریتم گربه‌ماهی[1] و DT در سیستم تشخیص نفوذ استفاده کرده‌اند مقایسه می‌کنیم. آن‌ها از روش‌های مختلف در راستای انتخاب تعداد ویژگی‌های مجموعه داده استفاده نموده‌اند که بهترین دقت برای ۱۰ ویژگی است. در جدول ۱۱ می‌توان نتایج مقایسه دقت و همچنین نرخ تشخیص الگوریتم GWO-ANN را با نتایج ارائه‌شده در مقاله اشاره‌شده مشاهده نمود.

جدول ۹. مقایسه دقت الگوریتم پیشنهادی با الگوریتم‌های استفاده‌شده

الگوریتم	دقت	تعداد ویژگی انتخاب‌شده
GWO – DT	۹۰/۷۲۷۳	۷
GWO – KNN	۸۳/۸۶۳۶	۹
GWO - ANN	۹۳/۲۲۷۳	۸

جدول ۱۰. میزان خطای به‌دست‌آمده از الگوریتم‌های مبتنی بر گرگ خاکستری

الگوریتم	خطا
GWO – DT	۰/۰۹۲۷
GWO – KNN	۰/۱۶۱۳
GWO - ANN	۰/۰۶۷۷

جدول ۱۱. مقایسه دقت و نرخ تشخیص

الگوریتم	دقت (٪)	نرخ تشخیص (٪)	تعداد ویژگی انتخاب‌شده
GWO-ANN	۹۳/۲۲۷۳	۹۵/۰۷۰۹	۸
CFA-DT	۹۲/۸۳۷	۹۲/۰۵۱	۱۰

[1] Cuttle Fish

همان‌طور که در جدول فوق نیز مشاهده می‌شود، روش پیشنهادی GWO-ANN ازنظر دقت و نرخ تشخیص همچنین انتخاب ویژگی کمتر عملکرد بهتری داشته است.

۵- جمع‌بندی

در شبکه‌های اینترنت اشیاء صنعتی حملات فعال باعث ایجاد ضرر و زیان فراوانی می‌شوند ازاین‌رو نیاز به یک سیستم تشخیص نفوذ برای تشخیص حملات ضروری است. در این مقاله با استفاده از مجموعه داده KDDCup۹۹ اقدام به ارائه یک سیستم تشخیص نفوذ مبتنی بر یادگیری ماشین شد. این مدل با استفاده از الگوریتم گرگ خاکستری به‌عنوان یک واحد استخراج‌کننده ویژگی در مجموعه داده عمل می‌کند و ویژگی‌های مستقلی را از مجموعه داده استخراج و به ورودی الگوریتم‌های یادگیری ماشین می‌دهد. الگوریتم‌های که موردبررسی قرارگرفته‌اند شامل KNN، DT و ANN است که از این میان الگوریتم ANN دارای عملکرد بهتری است. همچنین روش پیشنهادی GWO-ANN در مقایسه با روش CFA-ANN حدود ۱/۱۷۶۳ درصد بهبود را نشان می‌دهد.

مراجع

[1] X.Fei and G. Tian, "Fault Identification and Analysis of Communication NetworkBased on Deep Learning," 2022.

[2] J. Wan, J. Li, M. Imran and D. Li, "A blockchain-based solution for enhancing security and privacy in smart factory," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 6, pp. 3652-3660, 2019.

[3] R. Sadeghi, C. Wachsmann and M. Waidner, "Security and privacy challenges in industrial internet of thing," *In 2015 52nd ACM/EDAC/IEEE Design Automation Conference (DAC)*, pp. 1-6, June 2015.

[4] J. Sengupta, 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*, p. 149, 2020.

[5] M. Zolanvari, M. A. Teixeira, L. Gupta and K. Khan, "Machine learning-based network vulnerability analysis of industrial Internet of Things," *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 6822-6834, 2019.

[6] Deshpande, P. Pitale and S. Sanap, "Industrial automation using Internet of Things (IOT)," *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, vol. 5, no. 2, pp. 266-269, 2016.

[۷] ر. آرزم و ع. براتی، "ارائه روشی برای احراز هویت در اینترنت اشیاء مبتنی بر موقعیت گیرنده‌های Wi-Fi و فناوری زنجیره بلوکی"، پنجمین اجلاس ملی محاسبات توزیعی و پردازش داده‌های بزرگ، ص ۱۸ ,۱۳۹۸.

[۸] ع. سید ترابی و ر. پهلوان، "رمزگذاری در احراز هویت دستگاه‌های اینترنت اشیا"، چهارمین اجلاس ملی ایده‌های نوین در فنی و مهندسی، ص ۲۴، ۱۳۹۸.

[۹] ح. عیسی لو و ع. سلیمانی، "ارائه روشی کم‌بار برای احراز هویت اشیاء در اینترنت اشیا"، کنگره ملی تحقیقات بنیادین در مهندسی کامپیوتر و فناوری اطلاعات، ص ۱۵، ۱۳۹۸.

[10] Derhab, M. Guerroumi, A. Gumaei, L. Maglaras, M. A. Ferrag, M. Mukherjee, et al., "Blockchain and random subspace learning-based ids for sdn-enabled industrial iot security", Sensors, vol. 19, no. 14, pp. 3119, 2019.

[11] D. K. K. Reddy, H. Behera, J. Nayak, B. Naik, U. Ghosh and P. K. Sharma, "Exact greedy algorithm based split finding approach for intrusion detection in fog-enabled iot environment", Journal of Information Security and Applications, vol. 60, pp. 102866, 2021.

[12] F. Zhang, H. A. D. E. Kodituwakku, J. W. Hines and J. Coble, "Multilayer data-driven cyber-attack detection system for industrial control systems based on network system and process data", IEEE Transactions on Industrial Informatics, vol. 15, no. 7, pp. 4362-4369, 2019.

[13] O. Eigner, P. Kreimel, P. Tavolato and P. Kreimel, "Detection of man-in-the-middle attacks on industrial control networks," in In 2016 International Conference on Software Security and Assurance (ICSSA), 2016.

[14] M. Zolanvari, . M. A. Teixeira, L. Gupta, . K. M. Khan and R. Jain, "Machine Learning-Based Network Vulnerability Analysis of Industrial Internet of Things," Internet of Things Journal, vol. 6, no. 4, pp. 6822-6833, 2019.

[15] Saheed, Yakub Kayode, and Sanjay Misra. "A voting gray wolf optimizer-based ensemble learning models for intrusion detection in the Internet of Things." International Journal of Information Security (2024): 1-25.

[16] Wang, M.; Sun, Y.; Sun, H.; Zhang, B. Security Issues on Industrial Internet of Things: Overview and Challenges. Computers 2023, 12, 256. https://doi.org/ 10.3390/computers12120256.

[17] Fatani, A.; Dahou, A.; Abd Elaziz, M.; Al-qaness, M.A.A.; Lu, S.; Alfadhli, S.A.; Alresheedi, S.S. Enhancing Intrusion Detection Systems for IoT and Cloud Environments Using Growth Optimizer Algorithm and Conventional Neural Networks. Sensors 2023, 23, 4430. https://doi.org/10.3390/s23094430.

[18] Soltani, Nasim & Rahmani, Amir & Bohlouli, Mahdi & Hosseinzadeh, Mehdi. Robust intrusion detection for network communication on the Internet of Things: a hybrid machine learning approach. Cluster Computing. (2024). 1-17. https://doi.org/10.1007/s10586-024-04483-7.

[19] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf Optimizer," Advances in Engineering Software, vol. 69, pp. 46–61, Mar. 2014.

[20] C. M. Bishop, Pattern Recognition and Machine Learning. Springer, 2016.

[21] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set." 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications, 2009.

[۲۲] ل. عجمی بختیاروند و ز. بهشتی، "روشی نوین برای خوشه‌بندی داده‌ها با استفاده از الگوریتم بهینه‌سازی چهارگرگ خاکستری"، *نشریه مهندسی برق و مهندسی کامپیوتر ایران، ب- مهندسی کامپیوتر*، ص.ص ۲۷۴-۲۶۱،شماره ۴، سال ۱۹، ۱۴۰۰.

[23] S. Eesa, Z. Orman and. A. M. A. Brifcani, "A novel feature-selection approach based on the cuttlefish optimization algorithm for intrusion detection systems. Expert systems with applications," Elsevier, vol. 42, no. 5, pp. 2670-2679, 2015.