

A Novel Intrusion Detection System to Enhance Internet of Things Security, Based on the Integration of Evolutionary Approach, Blockchain, and Federated Learning

Mahdi Bazargani¹, Nafiseh Fareghzadeh^{2*}, and Mohammad Hossein Najafi³

1- Department of Computer Engineering, Za.C., Islamic Azad University, Zanzan, Iran.

2*- Department of Computer Engineering, Khod.c., Islamic Azad University, Zanzan, Iran.

3- Department of Computer Engineering, Za.C., Islamic Azad University, Zanzan, Iran.

¹mbzirn@iaau.ac.ir, and ^{2*}fareghzadeh@iaau.ac.ir, and ³m.najafi.research@gmail.com

Abstract- In this research, a novel and combinational intrusion detection system has been developed to enhance security in the Internet of Things (IoT) scope, which is designed based on the combination of blockchain, federated learning, and crow evolutionary algorithms. In addition to enhancing security, the proposed approach also addresses the challenges of data privacy and proposes a federated learning-based framework for anomaly detection that is superior to non-federated learning methods. In the operational process of the proposed system, data is first divided into independent blocks and, by using a combination of federated learning and blockchain approaches, samples are extracted for model training and recognition tasks. The proposed system is evaluated using supervised and unsupervised deep learning models, and the performance of the models in enhancing IoT network security is evaluated and compared. Also, the robustness of the proposed approach against a diverse range of attack vectors is carefully tested, and different federated learning algorithms are replaced with counterparts to reduce the computational load and optimize communication costs. Simulation results presented that the integration of dynamic thresholding method with federated learning model provides very high accuracy in classifying network traffic windows as normal or abnormal for detecting DDoS attacks. Based on the evaluation results, the proposed system has achieved an accuracy of 98.7%, a true positive rate of 0.987, a false positive rate of 0.026, and a recall of 0.987, which shows superior performance compared to the related approaches.

Keywords- Intrusion Detection, Security, IoT, Blockchain, Federated Learning, Crow Evolutionary Algorithm

سیستم تشخیص نفوذ نوین جهت افزایش امنیت اینترنت اشیا مبتنی بر تجمیع رویکرد تکاملی، بلاکچین و یادگیری فدرال

مهدی بازرگانی^۱، نفیسه فارغ زاده^{۲*}، محمد حسین نجفی^۳

۱- دانشکده مهندسی کامپیوتر، دانشگاه آزاد اسلامی، واحد زنجان، زنجان، ایران.

۲*- دانشکده مهندسی کامپیوتر، دانشگاه آزاد اسلامی، واحد خدابنده، زنجان، ایران.

۳- دانشکده مهندسی کامپیوتر، دانشگاه آزاد اسلامی، واحد زنجان، زنجان، ایران.

¹mbzirn@iau.ac.ir, ^{2*}fareghzadeh@iau.ac.ir, and ³m.najafi.research@gmail.com

* نشانی نویسنده مسئول: نفیسه فارغ زاده، زنجان، خدابنده، دانشگاه آزاد اسلامی واحد خدابنده، دانشکده مهندسی کامپیوتر.

چکیده- در این پژوهش، یک سیستم نوین تشخیص نفوذ ترکیبی در جهت افزایش امنیت اینترنت اشیا (IoT) ارائه شده است. هسته سیستم پیشنهادی بر تلفیق بلاکچین، یادگیری فدرال و الگوریتم تکاملی کلاغ استوار است. رویکرد پیشنهادی ضمن بهبود امنیت، حفظ حریم خصوصی داده ها را نیز تضمین می کند. بدین منظور، چارچوبی مبتنی بر یادگیری فدرال برای تشخیص ناهنجاری ها توسعه یافته است. در فرآیند اجرا، داده ها ابتدا به بلوک های مستقل تقسیم می شوند. سپس نمونه ها با ترکیب یادگیری فدرال و بلاکچین، جهت آموزش مدل و شناسایی استخراج می گردند. ارزیابی سیستم با استفاده از مدل های یادگیری عمیق نظارت شده و نظارت نشده انجام شده است. عملکرد چارچوب پیشنهادی در مقایسه با روش های غیرفدرال، برتری دارد. همچنین کارایی مدل ها در ارتقای امنیت شبکه IoT مقایسه گردید. استحکام رویکرد پیشنهادی در برابر بردارهای حمله متنوع و سطوح مختلف مقیاس پذیری سنجیده شد. به منظور کاهش بار محاسباتی و هزینه های ارتباطی، الگوریتم های یادگیری فدرال با نمونه های کارآمد جایگزین شدند. نتایج شبیه سازی نشان می دهد که ترکیب آستانه گذاری پویا با یادگیری فدرال، دقت طبقه بندی ترافیک شبکه را برای تشخیص حملات DDos افزایش می دهد. بر پایه نتایج ارزیابی، سیستم پیشنهادی به دقت ۹۸٫۷٪ دست یافته است. نرخ مثبت واقعی (TPR) و فراخوانی (Recall) برابر با ۰٫۹۸۷ و نرخ مثبت کاذب (FPR) برابر با ۰٫۰۲۶ محاسبه شد، که این مقادیر در مقایسه با پژوهش های مرتبط، نشان دهنده عملکرد برتر سیستم پیشنهادی است.

واژه های کلیدی: تشخیص نفوذ، امنیت، اینترنت اشیا، بلاکچین، یادگیری فدرال، الگوریتم تکاملی کلاغ

۱- مقدمه

پیوسته اشاره دارد که با حسگرها، نرم افزارها و سایر فن آوری های دیگری تجمیع شده است تا امکان جمع آوری و تبادل داده ها از طریق اینترنت یا سایر شبکه های ارتباطی را ایجاد نماید. اینترنت اشیا قادر به ایجاد شبکه ای یکپارچه و هوشمند است که در آن اشیاء فیزیکی بتوانند به نحوی با یکدیگر و با انسان ها ارتباط برقرار کنند که منجر به فرآیندهای کارآمد، خودکار،

امروزه اینترنت اشیا با ارائه فرصت های بی شمار و بهبود خدمات، راحتی و ایمنی در جنبه های مختلف زندگی، دستاوردهای قابل توجهی به دست آورده اند. اینترنت اشیا به شبکه وسیعی از دستگاه های فیزیکی، وسایل نقلیه، وسایل و اشیاء مختلف به هم

تصمیم گیری بهبود یافته و تجربه کاربری پیشرفته گردد.

با این حال در حوزه اینترنت اشیا، چالش‌هایی مانند مسائل امنیتی، حفظ حریم خصوصی، و هزینه‌های بالای پیاده‌سازی نیز وجود دارند که باید به آنها توجه کرد. به عنوان نمونه، یک حمله‌ی DOS با پر کردن ترافیک سرویس‌دهنده، خدمات را تحت‌الشعاع قرار داده و وب‌سایت یا سایر منابع را از دسترس خارج می‌کند. حمله‌ی DDoS^۱ نیز شباهت زیادی به تهدید قبلی دارد، اما از سیستم‌های بیشتری استفاده می‌کند تا سیل ترافیک را تشدید نموده و با بارگذاری بیش از حد خدمات را مختل کند. این چالش‌ها با این واقعیت تشدید می‌شوند که دستگاه‌های اینترنت اشیا اغلب داده‌های حساس را جمع‌آوری و انتقال می‌دهند و آنها را به اهداف جذابی برای حملات سایبری تبدیل می‌کنند. با وجود چالش‌های امنیتی و حریم خصوصی که مانع پذیرش بلاشروط اینترنت اشیا در جامعه مبتنی بر فناوری می‌شود، دستاوردهای اینترنت اشیا نسبت به چالش‌های آن بیشتر است. جهت ایمن‌سازی بهتر این سیستم‌ها و شبکه‌ها در برابر نقض احتمالی اصول امنیتی، تلاش‌ها به طور مداوم در حال انجام است. تکنیک‌های جدید مبتنی بر هوش مصنوعی، از جمله یادگیری ماشینی و یادگیری عمیق، مورد استقبال گسترده قرار گرفته‌اند و برای تضمین امنیت شبکه‌های IoT استفاده می‌شوند [۱].

در حوزه امنیت، زنجیره‌بلوکی یا بلاکچین به عنوان یک فناوری رمزنگاری پیشرفته، انتخابی مناسب برای پشتیبانی از نیازمندی‌های پایه امنیتی در شبکه‌های اینترنت اشیا است. این فناوری در جهت حفظ اعتماد و امنیت در ارتباطات بین دستگاه‌ها، بهبود پایداری شبکه‌ها و کاهش احتمال تخلیه باتری در دستگاه‌های اینترنت اشیا، می‌تواند بسیار مفید باشد [۲]. نقش اساسی یک بلاکچین این است که یک روش رمزنگاری امن برای به دست آوردن یک بلوک و مجموعه‌ای از رکوردهای مرتب شده زمانی را ارائه دهد که به طور عمومی قابل تأیید و تغییر نیست. از آنجایی که بلاکچین‌ها در میان شبکه‌ای از هم‌تایان به اشتراک گذاشته شده و همگام می‌شوند، اغلب به عنوان یک پایگاه داده باز و رکوردهای جهانی استفاده می‌شوند. علاوه بر این، تراکنش‌های بلاکچین کپی شده و در یک شبکه هم‌تای به هم‌تای کاملاً غیرمتمرکز ذخیره می‌شوند و احتمال از دست رفتن داده‌ها را کاهش می‌دهند [۲]. با این وجود، برای استفاده کامل از قابلیت‌های فناوری بلاکچین، چندین مشکل

وجود دارد که باید حل شود. چالش‌های اساسی شامل مقیاس-پذیری، امنیت و حریم خصوصی است [۳]. دامنه فناوری بلاکچین فراتر از ارزش‌های دیجیتال است و کاربرد بالقوه آن شامل ساخت سیستم‌های امن در محیط‌های مجازی است. اخیراً با میلیاردها دستگاه IoT یا اینترنت اشیا صنعتی یا IIoT^۲، تبادل داده بسیار پراکنده، و تعامل غیرمتمرکز داده، اینترنت اشیا مسائل جدیدی را در ساخت یک IDS^۳ قابل اجرا ارائه کرده است. لذا برای اطمینان از امنیت IoT/IIoT، طراحی IDS برای شبکه‌های مرتبط ضروری است. به طور کلی، IDS به جلوگیری از ترافیک مضر کمک می‌کند زیرا برای شناسایی فعالیت شبکه، به ویژه برای تمایز بین ترافیک عادی و مخرب استفاده می‌شود. رویکرد IDS ترکیبی، مزایای بلاکچین و یادگیری فدرال در شبکه‌های IoT/IIoT را با استفاده از تکنیک‌های تشخیص مختلف در میان دستگاه‌های شبکه لایه پایین‌تر و بالا به ترتیب، ترکیب می‌کند. بلاکچین نیز عموماً به دلیل تغییر ناپذیری آن و رویه‌های توافقی که توسط گره‌ها دنبال می‌شود، برای افزایش امنیت در شبکه‌های IDS استفاده می‌شود. دستگاه‌های مبتنی بر اینترنت اشیا IoT یا IIoT دارای آسیب‌پذیری‌هایی هستند که عوامل مخرب به دلیل ماهیت بهم پیوسته و منابع محدودشان می‌توانند برای سود مالی، سرقت داده یا خرابکاری از آنها سوء استفاده کنند. حملات امنیتی علیه IoT/IIoT از ضعف‌های احراز هویت، پروتکل‌های رمزگذاری و آسیب‌پذیری‌های میان‌افزار سوءاستفاده می‌کنند که منجر به دسترسی غیرمجاز، دستکاری دستگاه، نقض داده‌ها و اختلالات شبکه می‌شود [۴]. اجرای اقدامات امنیتی قوی، استانداردسازی و ارزیابی ریسک، بهترین شیوه‌ها برای رسیدگی به این چالش‌ها است. این رویکرد خطرات را کاهش می‌دهد و از رشد و پیشرفت مداوم فناوری IoT حمایت می‌کند [۵].

اخیراً، بهره‌گیری از راه‌حل‌های ترکیبی و بهره‌گیری از مزایا و دستاوردهای تکنیک‌های برتر در جهت ارتقای امنیت در حوزه اینترنت اشیا بسیار موثر بوده است. یادگیری فدرال به عنوان رویکردی نوظهور در حوزه فناوری‌های هوشمند، محبوبیت بالایی به دست آورده است. هدف این تکنیک ارتقای استفاده از هوش مصنوعی، در عین حفظ حریم خصوصی است و در طیف متنوعی از برنامه‌ها کاربرد دارد. مفهوم اساسی یادگیری فدرال حول تمرکززدایی ذخیره سازی داده‌های کاربر، انتقال مسئولیت اصلی سرویس دهنده به جمع‌آوری پارامترهای مدل از همه مشتریان، و به‌روزرسانی مدل جهانی نهایی به صورت محلی در سرویس دهنده می‌چرخد. تکنیک‌های یادگیری

بر اینترنت اشیا، ضروری است ابتدا روش‌ها و رویکردهای موجود در این حوزه مورد بررسی قرار گیرند. لذا در این بخش، مرور و مقایسه‌ای بر تحقیقات و سیستم‌های مطرح‌شده مرتبط در ادبیات موضوعی اخیر انجام شده است.

در حوزه پژوهش، نویسندگان در مقاله [۷] تجزیه و تحلیل ظریفی از همگرایی اینترنت اشیا و بلاکچین ارائه کرده‌اند و یک معماری سه لایه جدید را پیشنهاد کرده‌اند که محاسبات شبیه و محاسبات ابری را برای غلبه بر برخی چالش‌های موجود در حوزه‌های امنیت، مقیاس‌پذیری، کارایی و تأخیر ادغام می‌کند. علیرغم نقاط قوت سیستم پیشنهادی آنها، نویسندگان اذعان دارند که حساسیت تحمل خطای بیزانسی به حملات سایبیل همچنان یک نگرانی بالقوه در سیستم پیشنهادی است. در مقاله [۸] یک چارچوب کنترل دسترسی نوآورانه، ولی محدود به لحاظ سطح اجرا، برای اینترنت اشیا ارائه شده است و از فناوری بلاکچین برای رفع نیاز روزافزون به امنیت در چشم‌انداز در حال گسترش اینترنت اشیا استفاده می‌شود. در مقاله [۹] نویسندگان با معرفی یک طرح مبتنی بر بلاکچین به افزایش امنیت اینترنت اشیا در شبکه‌های 5G کمک کرده‌اند. تحقیقات آنها شواهد دقیقی از صحت و جعل‌ناپذیری این طرح ارائه می‌دهد که رویکرد ایمن و کارآمدی را برای یکپارچگی داده‌ها ایجاد می‌کند. در مقاله [۱۰] یک معماری یکپارچه بلاکچین و اینترنت اشیا را برای کاهش ناکارآمدی‌هایی مانند ورود دستی داده‌ها و به‌روزرسانی‌های تأخیری پیشنهاد می‌کنند. نویسندگان در مقاله [۱۱] با پیشنهاد یک چارچوب مبتنی بر بلاکچین که از رویکرد هایپرلجر فابریک برای مدیریت دفتر کل محلی استفاده می‌کند. پژوهش [۱۳] یک مدل L2D2 مبتنی بر شبکه عصبی حافظه طولانی کوتاه مدت را برای تشخیص نفوذ چند کلاسه در محیط‌های اینترنت اشیا حوزه پزشکی پیشنهاد می‌دهد. مطالعه [۱۴] عملکرد الگوریتم‌های مختلف یادگیری ماشین و یادگیری عمیق را در تشخیص حملات انکار سرویس توزیع شده در محیط‌های اینترنت اشیا مقایسه می‌کند. استفاده از زمان پنجره مبتنی بر آنتروپی نیز برای افزایش کارایی تشخیص پیشنهاد شده است. تحقیق [۱۵] یک مدل یادگیری عمیق ترکیبی را با ترکیب ترانسفورماتور و شبکه عصبی کانولوشن، همراه با تکنیک‌های مدیریت عدم تعادل کلاس و پیش‌پردازش داده‌های بهبود یافته، برای ساخت یک سیستم تشخیص نفوذ مؤثر پیشنهاد می‌نماید. تحقیق [۱۶] مدل‌های هوش مصنوعی و فناوری بلاکچین را برای تشخیص ناهنجاری‌های امنیت سایبری و ارائه مدیریت داده‌های ایمن و ضد دستکاری در

گروهی مانند رویکرد جنگل تصادفی نیز قادرند در مدیریت داده‌های حجیم این حوزه توانمند عمل کنند. همچنین، استفاده از رویکردهای تکاملی مانند الگوریتم بهینه‌سازی کلاغ نیز در جهت ایجاد یک راه حل امنیتی جامع سودمند می‌باشند [۶].

به منظور مواجهه با چالش‌های بیان شده و در جهت افزایش امنیت در حوزه اینترنت اشیا، رویکرد پیشنهادی این پژوهش توسعه یک سیستم تشخیص نفوذ نوین مبتنی بر تجمیع ساختارمند بلاکچین، یادگیری فدرال و الگوریتم بهینه‌سازی کلاغ می‌باشد و هدف نهایی ارائه یک راه‌حل عملیاتی و کاربردی است که به طور مؤثر در سیستم‌های تشخیص نفوذ مستقر باشد. لذا با توجه به ضرورت و محاسن بکارگیری رویکرد ترکیبی اتخاذ شده در جهت افزایش ایمن‌سازی شبکه‌های اینترنت اشیا، اهداف اصلی و دستاوردهای این پژوهش به شرح زیر است:

- ۱- توسعه یک سیستم تشخیص نفوذ ترکیبی نوین مبتنی بر تجمیع بلاکچین، یادگیری فدرال و رویکرد تکاملی جهت افزایش امنیت در اینترنت اشیا.
 - ۲- بهبود مستمر دقت مدل و کارایی سیستم پیشنهادی از طریق یک مکانیسم تکاملی انگیزشی با استفاده از الگوریتم بهینه‌سازی جستجوی کلاغ و استراتژی‌های مدیریت گره منطقی.
 - ۳- ارائه مکانیزمی مؤثر جهت ارزیابی استحکام و دقت سیستم پیشنهادی تحت سناریوهای مختلف حمله.
- اهداف فوق در راستای طراحی یک سیستم ترکیبی ایمن، قدرتمند و مقاوم در برابر حملات رایج شبکه‌ای DoS و DDoS روی زیرساخت اینترنت اشیا توسعه داده شده است.
- در ادامه مقاله، ابتدا تحقیقات وابسته مرور و با رویکرد پیشنهادی این پژوهش مقایسه می‌شوند. سپس، در بخش سوم سیستم پیشنهادی، فازهای اجرایی، جریان‌های کاری و جزئیات توصیف می‌گردد. بخش چهارم مراحل پیاده‌سازی سیستم طراحی شده با نرم‌افزارهای مربوطه و مقایسه با روش‌های موجود را مورد بررسی قرار می‌دهد. در نهایت، در بخش پنجم مقاله حاضر نتایج به‌دست‌آمده جمع‌بندی و مقایسه شده و پیشنهادهایی در جهت توسعه بکارگیری و افزایش امنیت در شبکه‌های مبتنی بر زیرساخت اینترنت اشیا ارائه خواهند شد.

۲- تحقیقات مرتبط و مقایسه

با توجه به اهمیت چالش‌های امنیتی در حوزه شبکه‌های مبتنی

اینترنت اشیا در حوزه سلامت ادغام می‌کند. تحقیق [۱۷] از یک رویکرد مبتنی بر یادگیری ماشین برای تشخیص حملات چند لایه در محیط‌های IoT استفاده می‌کند. تحقیق [۱۸] از چندین الگوریتم یادگیری ماشین برای تشخیص نفوذ در شبکه‌های اینترنت اشیا استفاده می‌کند. الگوریتم‌های مورد بررسی شامل الگوریتم‌های یادگیری نظارت‌شده و بدون نظارت هستند. تحقیق [۱۹] یک مدل تشخیص نفوذ مبتنی بر گروه را برای خانه‌های هوشمند متصل به اینترنت اشیا پیشنهاد می‌کند. این مدل با استفاده از مجموعه داده DS2OS آموزش داده و آزمایش شده و بر اساس چندین معیار عملکرد ارزیابی شده است. درواقع هدف، تشخیص انواع مختلف حملات و رفتار عادی در محیط یک خانه هوشمند است [۲۰]. براساس جنبه های امنیتی بررسی شده در تاریخچه پژوهش، مقایسه ابعاد اصلی رویکردهای مرتبط در جدول زیر خلاصه شده است.

علاوه بر موارد فوق، نویسندگان در [۲۴] مدل تشخیص بلاکچین توزیع‌شده مبتنی بر جامعه را بر اساس الگوی پیشنهاد-انتخاب-تنظیم پیشنهاد دادند. نتایج مطالعات آنها نشان می‌دهد که چگونه الگوریتم به طور مؤثر عمل می‌کند و

ساختارهای گروهی را در شبکه اعتماد پیچیده بیت‌کوین با موفقیت تشخیص می‌دهد. در مقاله [۲۵] استفاده از بلاکچین مبتنی بر سیستم تشخیص نفوذ مشارکتی به رفع برخی مشکلات موجود سیستم تشخیص نفوذ مشارکتی، از جمله محرمانگی و ایجاد اجماع، کمک می‌کند و تبادل امضاها، تشکیل بین میزبان‌ها و سیستم تشخیص نفوذ مشارکتی را تسهیل می‌کند. در [۲۶] در پاسخ به نگرانی‌های فعلی سیستم تشخیص نفوذ معمولی، الکسوپولوس و همکارانش [۲۷] یک سیستم تشخیص نفوذ مبتنی بر بلاکچین را پیشنهاد کردند. فناوری بلاکچین برای بهبود یا اصلاح مدیریت اعتماد بین گره‌های سیستم تشخیص نفوذ استفاده می‌شود. آنها از هشدارهای خام استفاده کرده‌اند که می‌توانند توسط گره‌های سیستم تشخیص نفوذ به عنوان یک تراکنش در فناوری بلاکچین در بین گره‌های نقطه ورود غیرقانونی تکثیر شوند. در [۲۸] نویسندگان مقاله یک رویکرد تشخیص نفوذ مشارکتی مبتنی بر بلاکچین را بدون هیچ گونه مداخله شخص ثالث مورد اعتماد یا مرکزی برای تشخیص نفوذ توزیع شده در سیستم‌های میکروگرید معرفی کرده‌اند.

جدول ۱: مقایسه پارامتریک از تحقیقات مرتبط اخیر

مرجع	چالش امنیتی	رویکرد	محدودیت
[۱۳]	نیاز به یک سیستم تشخیص نفوذ چند کلاسه دقیق و کارآمد که قادر به شناسایی حملات بلادرنگ باشد.	تحلیل و بررسی مدل‌های یادگیری عمیق جایگزین و تکنیک‌های بهینه‌سازی برای افزایش بیشتر عملکرد و کاهش سربار محاسباتی.	رویکرد محاسباتی پرهزینه براساس شبکه‌های عصبی گراف.
[۱۴]	آسیب‌پذیری زیرساخت‌های اینترنت اشیا مبتنی بر ابر در برابر حملات انکار سرویس توزیع‌شده.	توسعه مکانیسم‌های تشخیص تطبیقی و بلادرنگ برای حملات انکار سرویس توزیع شده در محیط اینترنت اشیا مبتنی بر ابر.	اتکا به محیط عملیاتی محدود و مجموعه داده‌های خاص تعمیم‌پذیری یافته‌ها را به سایر محیط‌های اینترنت اشیا یا سناریوهای حمله محدود می‌نماید.
[۱۵]	تشخیص نفوذ در سیستم‌های شبکه، به ویژه با تمرکز بر محدودیت‌های روش‌های موجود در مدیریت عدم تعادل کلاس.	بررسی استراتژی‌های بهینه‌سازی بیشتر برای کاهش پیچیدگی محاسباتی مدل، که امکان استقرار آن را در محیط‌های با محدودیت منابع فراهم می‌کند.	محدودیت‌های بالقوه مربوط به پیچیدگی محاسباتی.
[۱۶]	مسئله ناهنجاری‌های امنیت سایبری در اینترنت اشیا مراقبت‌های بهداشتی، با تمرکز بر نیاز به مدیریت داده‌های امن.	بررسی پلتفرم‌های بلاکچین جایگزین، بررسی استفاده از مدل‌ها و تکنیک‌های پیشرفته تر هوش مصنوعی.	سربار محاسباتی تراکنش‌های بلاکچین و چالش‌های مقیاس‌پذیری.
[۱۷]	چالش بهینه‌سازی انتخاب ویژگی برای مدل‌های یادگیری ماشینی که برای تشخیص حملات چندلایه طراحی شده‌اند.	بررسی سایر روش‌های انتخاب ویژگی، ارزیابی روش پیشنهادی در طیف وسیع‌تری از حملات و مجموعه داده‌ها.	کارایی و اثربخشی روش انتخاب ویژگی پیشنهادی ممکن است بسته به الگوریتم یادگیری ماشین خاص متفاوت باشد.
[۱۸]	پرداختن به نگرانی فزاینده تهدیدات امنیتی که دستگاه‌ها و شبکه‌های اینترنت اشیا را هدف قرار می‌دهند.	توسعه مدل‌های یادگیری ماشین پیچیده، ایجاد سیستم‌های قوی و سازگار.	محدودیت‌های بالقوه مربوط به محیط و زیرساخت پیاده سازی و اجرایی.
[۱۹]	گسترش سطح حمله ناشی از ماهیت به هم پیوسته دستگاه‌های خانه هوشمند و محدودیت‌های راه‌حل‌های امنیتی سنتی	توسعه سیستم تشخیص نفوذ تطبیقی و پویا، ادغام مدل پیشنهادی با سایر سازوکارهای امنیتی.	محدودیت‌های بالقوه مربوط به پیچیدگی محاسباتی و اتکا به یک مجموعه داده خاص.

واقعی IoT داشته باشد.

در مقایسه با مدل های سنتی، در روش پیشنهادی به افزایش دقت و کارایی دست یافتیم و با انباشته شدن داده های اضافی، سود بیشتری پیش بینی می شود. ارزیابی ها و تحلیل ها یافته ها را تایید می کند. سیستم پیشنهادی به یک رویکرد پیشگام در به کارگیری فناوری بلاکچین و یادگیری فدرال می انجامد و به ویژه برای برنامه هایی که نیاز به یادگیری فدرال دارند، مانند امنیت شهر هوشمند، تشخیص شبکه وسایل نقلیه و طبقه بندی و ردیابی اطلاعات IoT مفید است. طرح پیشنهادی با تجمیع موثر بلاکچین و یادگیری فدرال، نه تنها قابلیت اطمینان یادگیری فدرال را تقویت می کند، بلکه توان عملیاتی بلاکچین را نیز افزایش داده و دارای کاربردهای ملموس و موثر است. در بخش بعدی ابعاد دیدگاه پیشنهادی، فازهای اجرایی و جریان های کاری مرتبط را توصیف می نمائیم.

۳- سیستم پیشنهادی

پژوهش حاضر یک چارچوب زنجیره ای جامع و سبک وزن جدید را معرفی می کند، که می تواند مدیریت گره ها را در چارچوب یادگیری فدرال مبتنی بر بلاکچین تسهیل کند. علاوه بر این، این پژوهش مکانیسم های تشویقی و اصول اثبات کار را در ساختار بلاکچین ادغام می کند تا کارایی عملیاتی زنجیره را افزایش دهد. همزمان، یک مؤلفه مدیریتی برای رسیدگی به حضور گره های کم اثر توسط الگوریتم بهینه سازی جستجوی کلاغ توسعه می یابد، در نتیجه اختلال احتمالی ناشی از گره های غیرفعال یا غیرقابل اعتماد در فرآیند یادگیری را کاهش می دهد. شرکت کنندگان در شبکه بلاکچین معمولاً به عنوان ماینر شناخته می شوند و مسئولیت اصلی آنها تولید، اعتبار سنجی و انتشار تراکنش های بلوکی است. با استفاده از رمزنگاری و فناوری قراردادهای هوشمند، هر رکورد تراکنش غیرقابل تغییر می شود و از یکپارچگی و ماهیت متوالی بلاکچین اطمینان حاصل می کند. در دیدگاه پیشنهادی برای محافظت از انتقال داده ها با تجمیع ساختارمند فناوری بلاکچین با یادگیری فدرال و یادگیری ماشین، مکانیسم های اشتراک IDS بهینه شده است، در نتیجه کیفیت جمع آوری داده ها و کارایی کلی سیستم را افزایش می دهد.

لذا با توجه به ضرورت، در پژوهش حاضر یک سیستم تشخیص نفوذ نوین جهت افزایش امنیت در حوزه اینترنت اشیا توسعه یافته است که بر مبنای ترکیب فناوری های بلاکچین، یادگیری فدرال و الگوریتم تکاملی کلاغ طراحی شده است. رویکرد پیشنهادی علاوه بر ارتقای امنیت، به چالش های حفظ حریم خصوصی داده ها نیز

لازم به ذکر است، تحقیقات وابسته بینش های ارزشمندی را در مورد تکنیک ها، چالش ها و راه حل های مرتبط با افزایش امنیت در شبکه های مبتنی بر زیرساخت اینترنت اشیا ارائه کرده اند. سیستم های پیشنهادی مزایای متعددی از جمله افزایش حریم خصوصی و انعطاف پذیری در برابر حملات سایبری بالقوه را ارائه می دهند که آنها را به یک راه حل امیدوارکننده برای پیاده سازی امنیت این شبکه های مدرن تبدیل می کند. با این حال، تجزیه و تحلیل ها چندین چالش جدید را نشان داد که ضرورتاً باید به آنها پرداخته شود، از جمله مقیاس پذیری، دشواری تضمین حریم خصوصی داده ها بدون دقت پایین، و اهمیت ارتباط و همکاری مناسب بین گره های شرکت کننده. علاوه بر این، ادغام تکنیک های پیشرفته مانند یادگیری انتقال فدرال و پروتکل های تجمیع امن می تواند دقت، کارایی و استحکام سیستم های تشخیص و پیشگیری از نفوذ (IDPS) فدرال را افزایش دهد. این همان رویکرد پیشنهادی و دستاورد اصلی پژوهش حاضر است که تکمیل کننده رویکردهای مرتبط قبلی، در جهت افزایش امنیت در حوزه شبکه های مبتنی بر اینترنت اشیا می باشد. در مقایسه با تحقیقات مرتبط لازم به ذکر است، ضرورت پژوهش و انگیزه های تحقیق حاضر ناشی از چالش های جدی امنیت و حریم خصوصی در اکوسیستم اینترنت اشیا است که به دلیل تنوع و پویایی بالای بردارهای حمله، افزایش دستگاه های متصل و محدودیت های پردازشی موجود در این شبکه ها، روش های محدود و کلاسیک تشخیص نفوذ را ناکارآمد ساخته است. علاوه بر این، انتقال و تجمیع داده ها در سرویس دهنده مرکزی به منظور آموزش مدل های یادگیری ماشین، ریسک افشای اطلاعات حساس را افزایش می دهد و نیازمند رویکردی ترکیبی مبتنی بر حفظ پراکندگی داده ها و حریم خصوصی است. در این راستا، رویکرد پیشنهادی این پژوهش با الهام از پتانسیل بالای یادگیری فدرال برای آموزش مدل های توزیع شده بدون نیاز به اشتراک گذاری داده خام، و همچنین استفاده از بلاکچین جهت تضمین شفافیت و امنیت تراکنش های شبکه، سیستمی چندمنظوره، جامع و ترکیبی در جهت تشخیص نفوذ در حملات رایج در حوزه اینترنت اشیا مانند حملات DoS و DDoS ارائه می دهد. اهمیت کاربردی رویکرد پیشنهادی در کاهش نرخ هشدار کاذب، ارتقای پایداری و مقیاس پذیری IDS های IoT و امکان استقرار آن در محیط های توزیعی و چنددینفع نهفته است. همچنین بهره گیری از رویکرد تکاملی و الگوریتم بهینه سازی کلاغ برای غلبه بر مشکلات همگرایی و افزایش دقت شناسایی، موجب شده است تا سیستم توسعه یافته قابلیت مقابله مؤثر با حملات متنوع را داشته و با هزینه محاسباتی و ارتباطی بهینه، قابلیت استقرار عملی را در محیط های

- بلاکچین سبک‌وزن با الگوریتم اجماع اثبات سهام PoS^۵ اصلاح‌شده برای ایجاد لایه اعتماد، تضمین یکپارچگی داده‌ها و ردیابی معتبر تراکنش‌های امنیتی.
- تکنیک یادگیری فدرال با معماری Client-Server که امکان آموزش غیرمتمرکز مدل‌های یادگیری عمیق را فراهم کرده و حریم خصوصی داده‌ها را حفظ می‌کند.
- الگوریتم جستجوی کلاغ بهبودیافته برای انتخاب بهینه ویژگی‌ها و تنظیم خودکار پارامترهای شبکه‌های عمیق.

سیستم پیشنهادی یک چارچوب زنجیره یادگیری فدرال سبک با پشتیبانی مدیریت گره‌ها را پیشنهاد می‌نماید که افزایش دقت مدل در چارچوب پیشنهادی، از طریق مدیریت موثر اجزای شبکه و گره‌ها و توسعه مکانیسم‌های اثبات کار صورت می‌پذیرد. روش پیشنهادی بهبود مستمر دقت مدل و کارایی استفاده از دستگاه را از طریق یک مکانیسم انگیزشی، با استفاده از الگوریتم بهینه‌سازی جستجو کلاغ و استراتژی‌های مدیریت گره منطقی محقق می‌کند. به منظور ارزیابی سیستم پیشنهادی آزمون‌های استحکام و دقت تحت سناریوهای مختلف حمله انجام می‌شود. برای ساده‌سازی بلاکچین و کاهش تراکم، در این رویکرد سیستم فایل بین سیاره‌ای برای مدیریت بلاک‌های بیش از حد گنجانده شده است. براساس شکل ۱، فرایند عملیاتی سیستم شامل مراحل زیر تعریف می‌گردد:

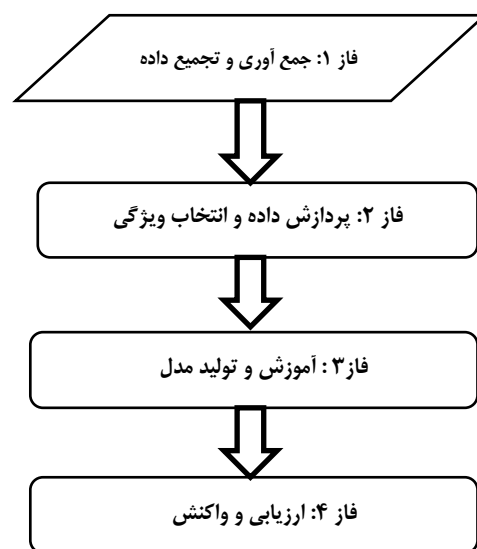
- فاز اول: جمع‌آوری و تجمیع داده
در این فاز، براساس مجموعه داده سیستمی، جمع‌آوری داده‌ها در گره‌های IoT (Edge) شامل داده‌های ترافیک شبکه، سنسورها، لاگ‌های دستگاهی صورت می‌پذیرد.
- فاز دوم: پردازش داده و انتخاب ویژگی
در فاز دوم ابتدا پیش‌پردازش داده‌ها شامل عملیات پاک‌سازی نویز و نرمال‌سازی و تجمیع اطلاعات انجام شده و سپس عملیات استخراج ویژگی‌ها توسط الگوریتم تکاملی کلاغ برای انتخاب ویژگی مؤثر صورت می‌پذیرد.
- فاز سوم: آموزش و تولید مدل
در فاز سوم آموزش مدل فدرال با ساختار سلسله‌مراتبی انجام می‌شود. هر گره مدل محلی می‌سازد، سپس وزن‌ها با سرویس دهنده تجمیع شده (FedAvg) و نسخه ارتقا یافته مدل (ثبت بلوک به بلوک در بلاکچین برای تأیید و اطمینان از عدم تغییر غیرمجاز) ثبت می‌گردد.
- فاز چهارم: ارزیابی و واکنش
در آخرین فاز اجرایی ارزیابی و واکنش به حملات، شامل تشخیص نوع حمله، محاسبه معیارهای دقت، نرخ خطا،

توجه کرده و چارچوبی عملیاتی بر اساس یادگیری فدرال برای تشخیص ناهنجاری‌ها پیشنهاد می‌دهد.

لازم به ذکر است، امنیت در این پژوهش به معنای تضمین تشخیص به موقع و دقیق حملات و ناهنجاری‌ها در شبکه‌های اینترنت اشیاء با کمترین نرخ هشدارهای مثبت کاذب است، همراه با حفظ محرمانگی و صحت داده‌ها در شبکه. به عبارت دیگر در پژوهش حاضر، هدف ارتقای امنیت در دو حوزه اصلی است:

- افزایش دقت در شناسایی رفتارهای غیرعادی یا مخرب (تشخیص نفوذ مؤثر)
- حفظ حریم خصوصی داده‌های پراکنده در شبکه بدون نیاز به اشتراک‌گذاری داده خام بین گره‌ها.

درواقع، انگیزه پژوهش بهبود تشخیص حملات جدید و ناشناخته، کاهش خطاهای مثبت کاذب (FP) و تأخیر، حفظ حریم خصوصی دستگاه‌های IoT و مقیاس‌پذیری به ویژه در سناریوهای ارزیابی شده است. همچنین، در سیستم پیشنهادی بهره‌گیری از یادگیری فدرال و بلاکچین سبب افزایش مقاومت در برابر حملات داخلی و حملات تغییر داده‌ها می‌شود، چرا که این فناوری‌ها صحت داده‌ها، جلوگیری از دستکاری، و حفظ محرمانگی را فراهم می‌کنند. شکل زیر فرایند عملیاتی سیستم پیشنهادی را نمایش می‌دهد:



شکل ۱: فرایند عملیاتی سیستم پیشنهادی و فازهای اجرایی

مطابق شکل بالا، سیستم پیشنهادی این پژوهش بر مبنای یک معماری چندلایه و یکپارچه شامل چهار فاز اجرایی و جریان‌های کاری مرتبط، طراحی شده است که در آن سه تکنولوژی کلیدی به‌صورت هم‌افزا و ترکیبی به شرح زیر ادغام می‌گردند:

FN و FP، و بررسی عملکرد تحت سناریوهای حمله صورت می پذیرد.

در فرایند اجرایی، ابتدا داده‌های ترافیک شبکه از دستگاه‌های IoT جمع‌آوری و پس از پیش‌پردازش اولیه، بر روی بلاکچین محلی ثبت می‌شوند؛ سپس هر گره مشارکت‌کننده با استفاده از الگوریتم کلاغ، فرآیند انتخاب ویژگی چندهدفه را بر روی مجموعه داده خود اجرا کرده و زیرمجموعه بهینه ویژگی‌ها را شناسایی می‌کند؛ در ادامه، مدل‌های یادگیری عمیق (شامل DNN، CNN و LSTM) به‌صورت محلی آموزش داده شده و وزن‌های آموخته شده، بدون انتقال داده‌های خام، در چارچوب یادگیری فدرال تجمیع می‌شوند تا مدل عام بهینه حاصل شود. نهایتاً و در آخرین فاز براساس تشخیص نوع حمله واکنش مناسب صورت می‌پذیرد.

لازم به ذکر است، استفاده از الگوریتم کلاغ در این پژوهش به دلیل کارایی بالا در مسائل پیچیده، توانایی بهینه‌سازی سریع، انعطاف‌پذیری و سازگاری با محیط‌های پویا و تکنولوژی‌های نوین مانند بلاکچین و یادگیری فدرال انتخاب شده است. این الگوریتم با داشتن مزایای متعدد نسبت به سایر روش‌های بهینه‌سازی، گزینه مناسبی برای بهبود امنیت در اینترنت اشیا و پیشگیری از نفوذ است. الگوریتم تکاملی بهینه‌سازی کلاغ برای اجرا باید تعداد تکرار مشخصی را برای اتمام به الگوریتم در نظر بگیرد. این تعداد باید عددی مناسبی باشد، زیرا این تعداد در صورت زیاد بودن زمان اجرای روش را غیرمنطقی می‌نماید و در صورتی که این تعداد بسیار کم باشد به مقدار بهینه مورد نظر دست نیافته و نتایج نهایی مناسب نخواهند بود. احتمال آگاهی به عنوان یک AP نامیده می‌شود، که اختلاف تشدید و تنوع را کاهش می‌دهد. مقادیر منتخب برای پارامترهای الگوریتم تکاملی بهینه‌سازی کلاغ در بخش پیاده‌سازی و ارزیابی سیستم پیشنهادی توصیف خواهند شد.

همچنین، در سیستم پیشنهادی تابع برازش نقش اصلی را در ارزیابی کیفیت جواب در الگوریتم تکاملی ایفا می‌کند. در مدل پیشنهادی ما که امنیت اینترنت اشیا مبتنی بر ترکیب بلاکچین و یادگیری فدرال طراحی شده است، در تابع برازش باید معیارها و پارامترهای کلیدی زیر را در نظر گرفته شود:

- دقت تشخیص نفوذ: معیار اصلی ارزیابی عملکرد مدل، دقت در شناسایی حملات و نفوذهای احتمالی است.
- هزینه ارتباطی: با توجه به یادگیری فدرال، میزان داده‌های منتقل‌شده بین گره‌ها بسیار مهم است. تابع برازش باید تلاش کند هزینه ارتباط را به حداقل برساند.

- زمان اجرا: تابع برازش باید بازده زمانی را ارزیابی کند تا مدل در زمان کوتاه با دقت بالا پاسخ دهد.

- مصرف انرژی: در اینترنت اشیا، مصرف انرژی باید حداقل شود تا عمر دستگاه‌ها افزایش پیدا کند.

بنابراین به منظور ارزیابی دقیق، یک ترکیب خطی وزنی از معیارهای کلیدی به صورت رابطه زیر تعریف می‌شود:

$$\begin{aligned} \text{Fitness} = & w_1 \times \text{Accuracy} \\ & - w_2 \times \text{CommunicationCost} \\ & - w_3 \times \text{ExecutionTime} \\ & - w_4 \times \text{EnergyConsumption} \end{aligned} \quad (1)$$

در رابطه (۱)، متغیرهای w_1, w_2, w_3, w_4 وزن‌هایی هستند که بسته به درجه اهمیت هر معیار در سیستم اجرایی و محیط پیاده‌سازی تعیین می‌شوند. در سیستم پیشنهادی، استفاده از الگوریتم تکاملی و بهینه‌سازی کلاغ (CSA)، با هدف انتخاب ویژگی‌های بهینه از داده‌های حجیم و پیچیده‌ی ترافیک اینترنت اشیا، جهت بهبود عملکرد سیستم تشخیص نفوذ (IDS) انجام شده است. در این پژوهش، CSA برای بهینه‌سازی هاپیرپارامترهای مدل یادگیری فدرال (FL) تشخیص نفوذ به کار گرفته می‌شود. هدف بهینه یافتن مجموعه‌ای از هاپیرپارامترها است که منجر به حداکثر صحت و دقت و حداقل نرخ خطای مثبت کاذب^۶ در اینترنت اشیا می‌شود. بنابراین مسئله به‌صورت یک مسئله بهینه‌سازی چندهدفه مدل شد، که اهداف آن بیشینه‌سازی دقت مدل^۷، کمینه‌سازی هزینه ارتباطی در آموزش فدرال، کمینه‌سازی زمان اجرا و کاهش مصرف انرژی دستگاه‌ها است. هدف دیگر تابع برازندگی کمینه‌سازی هزینه می‌باشد، لذا همگرایی نمودار به سمت مقدار کمتر مورد انتظار است.

این اهداف در قالب یک تابع برازندگی به صورت رابطه (۱) در مقاله تعریف شدند. هر کلاغ در جمعیت الگوریتم نشان‌دهنده‌ی یک بردار انتخاب ویژگی است. به عبارت دیگر، هر موقعیت کلاغ معرف مجموعه‌ای از ویژگی‌های فعال یا غیرفعال در مدل IDS می‌باشد. برای مثال، اگر داده اولیه شامل M ویژگی باشد، موقعیت کلاغ یک بردار دودویی به طول M است که در آن مقدار «۱» نشان‌دهنده انتخاب و مقدار «۰» نشان‌دهنده حذف آن ویژگی است. به منظور ارزیابی کیفیت هر بردار ویژگی، مدل یادگیری فدرال با مشارکت گره‌های IoT آموزش داده می‌شود و چهار معیار کلیدی شامل دقت، هزینه ارتباطی، زمان اجرا و مصرف انرژی محاسبه می‌گردند. سپس با استفاده از وزن‌دهی (w_1, w_2, w_3, w_4) یک مقدار برازندگی نهایی محاسبه می‌شود.

در مدل پیشنهادی، چهار وزن w_1 تا w_4 بر اساس اهمیت عملیاتی

نقش محوری در افزایش کارایی و دقت مدل تشخیص نفوذ پیشنهادی ایفا می‌کند.

مطابق توضیحات بالا، یکی از اهداف تابع برازندگی فوق کمینه‌سازی هزینه است. بنابراین نمودار به‌درستی به سمت کمینه همگرا می‌شود. بدلیل آنکه تابع هزینه تعریف‌شده یک تابع مینیمم‌سازی است، همگرایی نمودار به سمت مقدار کمتر مورد انتظار است. در سیستم پیشنهادی انتخاب ویژگی جهت تشخیص نفوذ مبتنی بر الگوریتم کلاغ است، سپس بعد از طبقه‌بندی ویژگی‌های مناسب جهت تشخیص نوع ترافیک سالم و غیرسالم با استفاده از تکنیک یادگیری فدرال انجام می‌شود. در گام بعدی، ویژگی‌ها براساس وزن اختصاص داده شده مرتب و انتخاب می‌شوند و ویژگی‌های انتخاب شده برای تولید مدل به یادگیری فدرال داده می‌شود. تشخیص نوع ترافیک سالم و غیرسالم با تکنیک یادگیری فدرال انجام می‌شود. در آخرین فاز، سیستم بر اساس معیارهای کلیدی ارزیابی می‌گردد.

لازم به ذکر است، همه فعالیت‌های مرتبط با جمع‌آوری داده و نگهداری داده بر بستر بلاکچین صورت می‌پذیرد. حذف بلاکچین از ساختار پیشنهادی می‌تواند به کاهش امنیت و قابلیت اعتماد سیستم منجر گردد، حتی اگر بطور مستقیم در مقادیر دقت تشخیص نمود نیابد. همچنین، بلاکچین در این سیستم به عنوان یک لایه امنیتی مکمل و محافظت‌کننده در کنار یادگیری فدرال و الگوریتم بهینه‌سازی کلاغ عمل می‌کند. در بخش بعد جزئیات پیاده‌سازی و ارزیابی سیستم پیشنهادی تشریح گردیده است.

۴- ارزیابی

در این بخش جزئیات پیاده‌سازی رویکرد پیشنهادی توصیف می‌گردد و عملکرد سیستم مورد ارزیابی قرار می‌گیرد. به منظور پیاده‌سازی سیستم پیشنهادی از نرم‌افزارهای Matlab [۲۹] و Weka [۳۰] استفاده شده است. سیستم پیشنهادی در عمل از مدل‌های یادگیری عمیق بدون نظارت و نظارت شده را استفاده می‌کند که از طریق آموزش فدرال برای توسعه سیستم‌های تشخیص نفوذ برای دستگاه‌های اینترنت اشیا آموزش دیده‌اند.

در این پژوهش، برای ارزیابی عملکرد مدل پیشنهادی در تشخیص نفوذ و حملات مختلف در محیط اینترنت اشیا، از مجموعه داده N-BaIoT به عنوان یک مجموعه داده معتبر و استاندارد استفاده شده است [۳۱]. مجموعه داده N-BaIoT یکی از پایگاه‌های داده اطلاعاتی معتبر و متنوع برای ارزیابی سیستم‌های تشخیص نفوذ در محیط اینترنت اشیا (IoT) است. N-BaIoT دارای ۱۱۵ ویژگی اصلی است. پس از استخراج ویژگی، الگوریتم کلاغ ۱۳ ویژگی

معیارهای اصلی (دقت، هزینه مخابراتی، زمان آموزش و انرژی مصرفی) تعیین و مقدار آن‌ها به ترتیب ۰.۴۵، ۰.۲۰، ۰.۲۰ و ۰.۱۵ انتخاب شده است. این وزن‌ها مطابق استانداردهای طراحی توابع چندهدفه در FL، نیازمندی‌های محیط IoT/IDS و کالیبراسیون تجربی اولیه تنظیم شده‌اند. برای سنجش پایداری، ۱۲ سناریوی مختلف وزن‌دهی آزمایش شد که حداکثر افت دقت تنها ۰.۶٪ بود؛ بنابراین مدل نسبت به تغییر وزن‌ها حساسیت غیرعادی نداشته و نتایج قابل تکرار هستند. در هر تکرار، کلاغ‌ها موقعیت خود را بر اساس دو اصل رفتاری الهام‌گرفته از طبیعت به‌روزرسانی می‌کنند:

- کشف منبع غذا: حرکت به سمت موقعیت بهینه‌تر بر اساس اطلاعات ذخیره‌شده در حافظه.
- اجتناب از شکارچی: تغییر مسیر تصادفی برای جلوگیری از همگرایی زودرس.

این دو مکانیزم باعث ایجاد تعادل میان اکتشاف و بهره‌برداری در فضای جستجو می‌شود. پارامترهای الگوریتم از جمله اندازه جمعیت (۳۰ کلاغ)، تعداد تکرارها (۲۰ تکرار)، احتمال پرواز (۰.۸) و حد آستانه (۱-۰.۱)، معرفی و در آزمایش‌ها کالیبره شدند. در پایان فرایند، بهترین موقعیت کلاغ به عنوان مجموعه ویژگی بهینه انتخاب می‌شود و این مجموعه به فاز اجرایی بعدی (آموزش مدل تشخیص نفوذ با یادگیری فدرال) منتقل می‌شود. در فاز بعد، آموزش مدل‌های یادگیری عمیق (DNN، CNN، LSTM) به صورت محلی روی هر دستگاه انجام می‌شود. وزن‌ها و پارامترهای آموزش دیده بدون انتقال داده‌های خام به سرویس دهنده مرکزی ارسال می‌شوند. با استفاده از الگوریتم FedAvg، پارامترهای مدل‌های محلی تجمیع و مدل سراسری ارتقا می‌یابد. اطلاعات مربوط به پارامترها و وزن‌های مدل‌ها روی بلاکچین ثبت می‌شود تا از دستکاری جلوگیری شود. الگوریتم CSA برای بهینه‌سازی هایپرپارامترهای مدل و انتخاب بهترین مجموعه ویژگی‌ها در این فاز نیز استفاده می‌گردد. الگوریتم زمانی متوقف می‌شود که حداکثر تعداد تکرارها انجام شده باشد، یا به یک معیار همگرایی مشخص (مثلاً عدم بهبود قابل توجه در F1-Score در چندین دور متوالی) رسیده باشد. در پایان اجرای CSA، بهترین موقعیت یافت شده توسط تمام کلاغ‌ها که منجر به بالاترین F1-Score شده است، به عنوان هایپرپارامترهای بهینه برای مدل نهایی تشخیص نفوذ انتخاب می‌شود. این هایپرپارامترهای بهینه سپس برای آموزش نهایی و استقرار مدل در محیط شبیه‌سازی شده مورد استفاده قرار می‌گیرند. الگوریتم کلاغ به‌طور مستقیم نقش مدل‌سازی مسئله انتخاب ویژگی و بهینه‌سازی چندهدفه را ایفا می‌کند که منجر به افزایش دقت و کاهش هزینه‌ها در سیستم پیشنهادی شد. درواقع این الگوریتم به عنوان یک ابزار بهینه‌سازی،

منتخب را از میان ۱۱۵ ویژگی انتخاب کرده است. این مجموعه داده شامل داده‌های ترافیکی واقعی از بیش از ۹ نمونه مختلف دستگاه IoT است که رفتارهای عادی و حملات متنوع مانند حملات DDoS، نفوذهای غیرمجاز و سایر تهدیدات شبکه‌ای را در بر می‌گیرد. داده‌ها شامل ویژگی‌های آماری، رفتاری و زمانی از ترافیک شبکه هستند و به دلیل تنوع دستگاه‌ها و نوع حملات، امکان ارزیابی جامع و دقیق مدل‌های یادگیری ماشین برای تشخیص ناهنجاری را فراهم می‌کنند. N-BaIoT دارای ویژگی‌هایی مانند حجم مناسب برای آموزش مدل‌های عمیق، تعادل نسبی بین نمونه‌های عادی و حمله و استاندارد بودن فرمت داده است که آن را به گزینه‌ای ایده‌آل جهت آزمایش و اعتبارسنجی سیستم‌های امنیتی تبدیل می‌کند. در مقاله حاضر، مجموعه داده N-BaIoT به عنوان منبع اصلی داده‌ها برای آموزش و ارزیابی سیستم تشخیص نفوذ استفاده شده است. ابتدا داده‌ها از نظر کیفیت بررسی و پیش‌پردازش شامل نرمال‌سازی و حذف نویز انجام شد. سپس با استخراج ویژگی‌های مرتبط و بهینه‌سازی شده توسط الگوریتم جستجوی کلاغ، مجموعه ویژگی‌های با ابعاد کمتر ولی اطلاعات‌دار بدست آمد که برای آموزش مدل‌های محلی یادگیری فدرال استفاده شد. بدین ترتیب، مدل‌ها توانستند رفتارهای ناهنجار و حملات مختلف را با دقت بالا تشخیص دهند. در این پژوهش، برای ارزیابی دقیق مدل پیشنهادی در تشخیص نفوذ و حملات مختلف، از مجموعه داده معتبر و استاندارد استفاده شده است N-BaIoT استفاده شده است که ویژگی‌های منحصر به فردی دارد و فرایند ارزیابی جامعی را تبیین می‌نماید.

به منظور تقسیم‌بندی و آماده‌سازی داده‌ها، برای هر مجموعه داده فرآیند پیش‌پردازش جامعی شامل موارد زیر انجام شده است: حذف رکوردهای تکراری و ناکامل، نرمال‌سازی ویژگی‌های عددی با استفاده از روش Min-Max Scaling، کدگذاری ویژگی‌های طبقه‌ای با One-Hot Encoding، تعادل‌بخشی کلاس‌ها با تکنیک SMOTE، و تقسیم داده‌ها به نسبت ۷۰٪ آموزش، ۱۵٪ اعتبارسنجی و ۱۵٪ تست. همچنین برای شبیه‌سازی محیط یادگیری فدرال، هر مجموعه به ۱۰ زیرمجموعه غیرهمپوشان (client) تقسیم شده که هر کدام نمایانگر یک گره مشارکت‌کننده در شبکه IoT است.

لازم به ذکر است، یادگیری فدرال پایه بر Non-IID بودن است. درکل، از آنجائی که در محیط IoT داده‌ها برچسب‌گذاری نشده‌اند، نرخ تولید متفاوت است و سناریوهای ترافیکی بسیار نامتوازن هستند، تقسیم یک به یک دستگاه به گره یک حالت پایه است و اضافه کردن گره دهم یک سناریوی غیرخطی واقعی‌تر و با پیچیدگی بیشتر ایجاد می‌نماید. همچنین مدل پیشنهادی نیازمند ارزیابی در شرایط ناهمگون است. مدل AutoEncoder (AE) فدرال باید بتواند

عدم توازن، اختلاف حجم و رفتارهای دستگاه‌محور را مدیریت نماید. لذا، گره دهم دقیقاً برای سنجش مقاومت مدل در مقابل Non-IID شدید تعریف شده است. لذا، عملکرد مدل‌های آموزش‌دیده با آموزش فدرال با مدل‌هایی که از طریق غیرآموزش فدرال با استفاده از مجموعه داده N-BaIoT، شامل نه دستگاه اینترنت اشیا آموزش داده شده‌اند، مقایسه می‌شود. همچنین جهت بهبود دقت مدل‌های یادگیری عمیق، یک بهینه‌سازی فراپارامتر با جستجوی تصادفی انجام می‌شود. معیارهای عملکرد مختلفی برای ارزیابی نتایج پیش‌بینی استفاده می‌شود. بر اساس آزمایش مدل‌های آموزش‌دیده آموزش فدرال و غیر آموزش فدرال بر روی هر نه دستگاه اینترنت اشیا، نتایج نشان می‌دهد که مدل AE بدون نظارت که از طریق آموزش فدرال آموزش داده شده است، از نظر همه معیارها بهترین گزینه است. برای سهولت در مدیریت مجموعه داده، ما ۵٪ از مجموعه داده اصلی را با استفاده از جستجوهای MySQL انتخابی استخراج کردیم. ۵٪ استخراج شده از ۴ فایل با حجم کلی ۱،۰۷ گیگابایت و حدود ۳ میلیون رکورد تشکیل شده است [۳۱و۱۶].

به منظور پردازش داده و انتخاب ویژگی از الگوریتم تکاملی بهینه‌سازی کلاغ استفاده شده است. الگوریتم کلاغ (CSA) در مدل پیشنهادی دو نقش مستقل و مکمل دارد: اول، انتخاب ویژگی که ریشه در ساختار استاندارد یک بردار دودویی $S = \{f_1, f_2, \dots, f_n\}$ ، $f_i \in \{0, 1\}$ دارد و دوم، بهینه‌سازی هایپر پارامترهای مدل فدرال. در این مرحله، الگوریتم CSA برای پیدا کردن تنظیمات بهینه آموزش به کار می‌رود تا سرعت همگرایی افزایش یابد، سربار ارتباطی کاهش یابد و مدل دقت بالاتری کسب کند. بردار جستجوی الگوریتم کلاغ برای تنظیم هایپر پارامترها به صورت $X = [lr, batch, nlayers, dropout, C]$ تعریف می‌شود. در الگوریتم کلاغ هر عامل جستجو شامل یک بردار پیوسته شامل هایپر پارامترهای کلیدی مدل یعنی نرخ یادگیری (۰،۰۱-۰،۰۰۱)، اندازه Batch (۱۶-۱۲۸)، تعداد لایه‌های AutoEncoder (۲-۶)، نرخ Dropout (۰،۴-۰) و ضریب هزینه ارتباطی در FL (۱-۲) را نمایش می‌دهد. در هر تکرار، مدل فدرال با مجموعه پارامترهای پیشنهادی اجرا، و دقت، سربار ارتباطی، زمان آموزش و انرژی مصرفی اندازه‌گیری می‌شود تا تابع برازندگی ارزیابی گردد. سپس موقعیت عامل‌ها براساس قواعد پرواز CSA و احتمال آگاهی به‌روزرسانی شده و از جستجوی مؤثر و عدم گرفتارشدن در کمینه‌های محلی اطمینان حاصل می‌شود. در نهایت، بردار با کمترین مقدار تابع برازندگی به عنوان مجموعه بهینه انتخاب می‌گردد. این الگوریتم برای اجرا باید تعداد تکرار مشخصی را برای اتمام به الگوریتم در نظر بگیرد و این تعداد باید عددی مناسبی

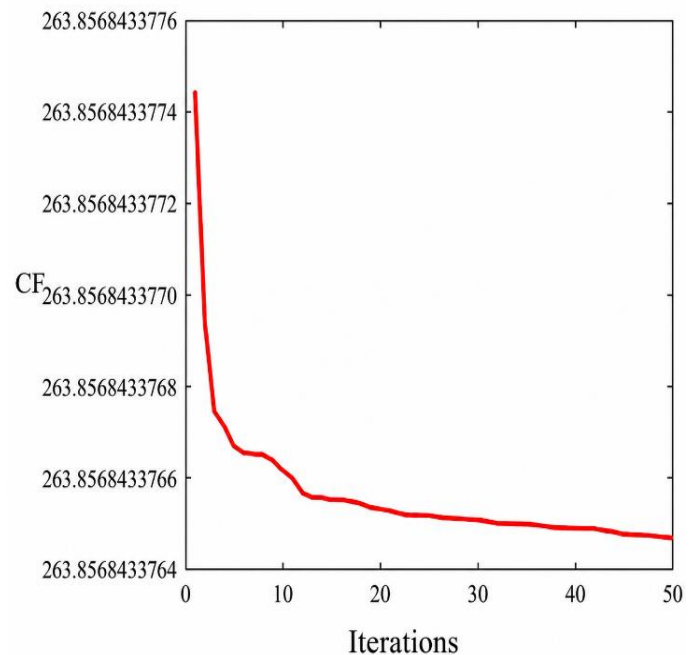
دست نیافته و کیفیت نتایج کاهش می‌یابد. براساس نتایج ارزیابی تابع هزینه، در قسمت الف شکل ۲، با افزایش تعداد تکرار به ۵۰، تغییر چشمگیری در بهبود نتایج همگرایی مشاهده نمی‌شود، در حالی که بار محاسباتی و زمان اجرا بسیار افزایش پیدا می‌کند. همانگونه که در قسمت ب، شکل ۲ نمایش داده شده است، مقادیر در تعداد ۲۰ تکرار و حد پایین و بالای ۱- و ۱ به همگرایی لازم رسیده و نیاز به تکرار بیشتر نیست و با این تعداد می‌توان به انتخاب مناسب دست یافت. بنابراین، انتخاب ۲۰ تکرار به عنوان نقطه مناسب، تعادلی بهینه بین دقت مدل و هزینه زمانی محاسبات برقرار می‌سازد و از لحاظ عملی به صرفه‌تر و کاراتر است. همچنین، پارامترهای الگوریتم کلاغ از جمله اندازه جمعیت، نرخ احتمال پرواز و پارامتر حافظه نیز به گونه‌ای تنظیم شده‌اند که با تعداد تکرار ۲۰ بهترین عملکرد را داشته باشند. تست‌های اعتبارسنجی با روش K-Fold با معیار (K=10) نیز موفقیت این انتخاب را تایید کرده‌اند و مدلی با دقت و استحکام بالا بر اساس این پارامترها ساخته شده است. بنابراین براساس نتایج، مقدار ۲۰ به عنوان عددی مناسب جهت تعداد تکرار انتخاب گردید.

در کلیه آزمایش‌ها از تست صحت به روش K-Fold با پارامتر رایج و نرمال K=10 استفاده شده است. در این نوع اعتبارسنجی داده‌ها به K زیرمجموعه افراز می‌شوند. از این K زیرمجموعه، هر بار یکی برای اعتبارسنجی و K-1 زیرمجموعه دیگر برای آموزش بکار می‌روند. این روال K بار تکرار می‌شود و همه داده‌ها دقیقاً یکبار برای آموزش و یکبار برای اعتبارسنجی بکار می‌روند. در نهایت میانگین نتیجه این K بار اعتبارسنجی به عنوان یک تخمین نهایی برگزیده می‌شود. در ادامه روش پیشنهادی با روش‌های مطرح یادگیری ماشین و معیارهای کلیدی ارزیابی مورد مقایسه قرار می‌گیرد. اولین معیار کیفی ارزیابی شده، معیار صحت است. صحت یک مدل با نسبت پیش‌بینی‌های دقیق تولید شده در تمام انواع پیش‌بینی‌ها ارزیابی می‌شود، یعنی اثربخشی نمونه‌های طبقه‌بندی شده براساس تعداد کلی رخدادها ارزیابی می‌گردد. اندازه‌گیری صحت به ویژه در مواردی ارزشمند است که توزیع کلاس‌ها در متغیر هدف به طور یکنواخت در سراسر مجموعه داده پخش شود. رابطه صحت ارزیابی شده در معادله (۲) بیان شده است [۳۱]:

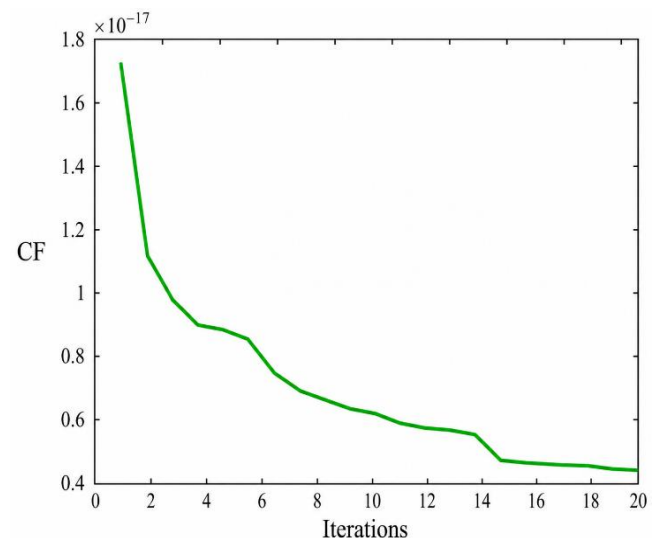
$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

در رابطه (۲)، TP تعداد نمونه‌های حمله است که سیستم به درستی به عنوان «حمله» تشخیص داده است. TN تعداد نمونه‌های نرمال است که سیستم به درستی به عنوان «نرمال» تشخیص داده است. FP تعداد نمونه‌های نرمال است که سیستم به اشتباه به عنوان

باشد. در شکل ۲. ارزیابی نتایج با تعداد تکرار و حدود آستانه مختلف در این الگوریتم نشان داده شده است.



الف) تعداد ۵۰ تکرار و حد پایین و بالای آزاد



ب) تعداد ۲۰ تکرار و حد آستانه ۱- و ۱

شکل ۲: مقادیر تابع هزینه CF الگوریتم تکاملی بهینه سازی کلاغ با تعداد تکرار مختلف

اساسا الگوریتم تکاملی کلاغ به منظور تشخیص نفوذ در شبکه اینترنت اشیا، جهت انتخاب ویژگی‌ها به تعداد مشخصی از تکرارها نیاز دارد؛ این تعداد باید به گونه‌ای باشد که الگوریتم به همگرایی مناسب برسد و به بهترین جواب بهینه نزدیک شود. اگر تعداد تکرارها بسیار زیاد باشد، زمان اجرای روش به شکل غیرمنطقی افزایش یافته و هزینه زمانی بالایی به سیستم تحمیل می‌شود. از سوی دیگر، اگر تعداد تکرارها پایین باشد، الگوریتم به مقدار بهینه

«حمله» تشخیص داده است. همچنین FN تعداد نمونه‌های حمله که سیستم به اشتباه به عنوان «نرمال» تشخیص داده است.

معیار ارزیابی بعدی، معیار FP Rate یا نرخ مثبت کاذب می باشد. این معیار به معنای نسبت نمونه‌های نرمالی است که به اشتباه به عنوان نمونه‌های آسیب دیده شناسایی شده‌اند. این معیار یکی از ارکان ارزیابی مدل‌های شناسایی نفوذ است که نشان می‌دهد چند درصد از نمونه‌های منفی واقعی (نرمال) به اشتباه مثبت تشخیص داده شده‌اند. این معیار یکی از بنیادی‌ترین شاخص‌ها در ارزیابی کیفیت سیستم‌های شناسایی نفوذ است، نشان می‌دهد سیستم تا چه حد هشدارهای مثبت کاذب (که می‌توانند باعث اتلاف منابع و اعتماد کمتر به سیستم شوند) را به حداقل رسانده است. این معیار بر اساس عناصر ماتریس درهم‌ریختگی تعریف می‌شود و براساس رابطه زیر قابل محاسبه می‌باشد:

$$FPRate = \frac{FP}{FP + TN} \quad (3)$$

در رابطه (۳)، FP Rate نشان‌دهنده نسبت نمونه‌های نرمالی است که به اشتباه حمله تشخیص داده شده‌اند. FP تعداد نمونه‌های منفی است که به اشتباه مثبت تشخیص داده شده‌اند. TN نیز تعداد نمونه‌های منفی که به درستی منفی تشخیص داده شده‌اند. عموماً این رابطه در کنار معیارهای دیگر ارزیابی مانند True Positive Rate یا Recall، Precision و Accuracy استفاده می‌شود تا عملکرد کلی مدل را نشان دهد. در مطالعه حاضر، ابتدا داده‌ها به دو دسته کلی مثبت (نمونه‌های حمله) و منفی (نمونه‌های نرمال) تقسیم شده‌اند. سپس پس از اجرای مدل شناسایی، ماتریس درهم‌ریختگی به دست آمده است که شامل شمارش موارد TP, FP, TN, FN می‌باشد. برای نمونه، اگر تعداد کل نمونه‌های نرمال (منفی واقعی) برابر با N_{neg} باشد و از این مقدار N_{neg} تعداد FP نمونه به اشتباه توسط مدل مثبت (حمله) تشخیص داده شده باشند، و مابقی $TN = N_{neg} - FP$ درست منفی شناخته شده باشند، آنگاه FP Rate مطابق فرمول بالا قابل محاسبه می‌باشد. بر اساس نتایج بدست آمده از ارزیابی، نرخ FP Rate برابر با ۰,۰۲۶ است که به معنی ۲,۶ درصد نمونه‌های نرمال اشتباه شناسایی شده است. این مقدار از تقسیم تعداد نمونه‌های FP بر مجموع FP و TN به دست آمده است. این محاسبه از طریق شمارش عناصر ماتریس درهم‌ریختگی حاصل از داده‌های تست صورت گرفته است.

معیار عملکردی مهم دیگر، یادآوری یا Recall می‌باشد. یادآوری تعداد رخدادهایی را نشان می‌دهد که به عنوان معیوب طبقه‌بندی شده‌اند که به‌طور دقیق توسط مدل پیش‌بینی شده‌اند. معادله زیر

نشان دهنده نسبت نمونه‌های مشکل‌ساز است که به‌طور دقیق توسط مدل شناسایی شده‌اند:

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

بعلاوه معیار دقت، دقت پیش‌بینی‌های مثبت انجام‌شده توسط مدل را اندازه‌گیری می‌کند و نسبت نمونه‌های پیش‌بینی‌شده به‌عنوان مثبت و مثبت واقعی را نشان می‌دهد و با استفاده از فرمول زیر محاسبه می‌شود [۳۲]:

$$precision = \frac{TN}{TN + FP} \quad (5)$$

نهایتاً امتیاز F_1score ، معیاری است که دقت و یادآوری را در یک مقدار واحد ترکیب می‌کند. این به‌ویژه زمانی مفید است که توزیع کلاس ناهموار (به‌معنی کلاس‌های نامتعادل) وجود داشته باشد. این معیار میانگین هارمونیک دقت و یادآوری است و با استفاده از فرمول زیر محاسبه می‌شود [۳۳]:

$$F_1score = 2 \times \frac{precision \times recall}{precision + recall} \quad (6)$$

نتایج بدست آمده در ارزیابی‌های عملیاتی سیستم پیشنهادی براساس معیارهای کلیدی توصیف شده، به شرح زیر است:

- دقت کلی سیستم یا Accuracy: معیار دقت سراسری نشان‌دهنده نسبت نمونه‌های درست به کل نمونه‌ها است. دقت مدل در شناسایی درست تراکنش‌های شبکه ۹۸,۷٪ گزارش شده است. این مقدار بالای دقت، ناشی از ترکیب بهینه الگوریتم یادگیری فدرال با انتخاب ویژگی‌های بهینه توسط الگوریتم تکاملی کلاغ و تضمین صحت داده‌ها توسط بلاک‌چین است. به دلیل هم‌افزایی این تکنولوژی‌ها، مدل توانسته است بیشترین داده‌های صحیح را به درستی طبقه‌بندی کند.
- نرخ مثبت‌های واقعی^۸: این معیار درصد حملات درست شناسایی شده از کل حملات واقعی را نشان می‌دهد. سیستم قادر به شناسایی ۹۸,۷٪ از حملات واقعی است که نشان‌دهنده حساسیت بالا به تهدیدات و قدرت بالای تشخیص است. عملکرد قوی ناشی از آموزش سلسله‌مراتبی مدل‌ها در یادگیری فدرال و انتخاب دقیق ویژگی‌ها است.
- نرخ مثبت‌های کاذب (FP rate): این معیار درصد نمونه‌های نرمال که به اشتباه به عنوان حمله شناسایی شده‌اند را نشان می‌دهد. در مطالعه حاضر این نرخ ۰,۰۲۶ گزارش شده است که مقدار پایینی است و بیانگر کاهش هشدارهای نادرست و

لازم به ذکر است، هم‌سویی مقادیر بدست آمده از ارزیابی سیستم، بیانگر قوت سیستم در مدیریت پویایی و پیچیدگی داده‌های ترافیکی در IoT است، که علاوه بر توان تشخیص دقیق، قابلیت تطبیق و به‌روزرسانی مستمر دارد. به طور کلی، مقادیر این معیارها نشان‌دهنده کارایی و توانمندی سیستم پیشنهادی در مقابله با پیچیدگی‌ها و پویایی‌های شبکه‌های اینترنت اشیا بوده، به طوری که با حفظ تعادل بین دقت، حساسیت به تهدیدات و کاهش خطاهای کاذب، مدل عملکرد عملیاتی موثری ارائه می‌دهد. دربخش بعدی به منظور تاکید بر دستاوردهای این پژوهش نتایج بدست آمده از ارزیابی را با تحقیقات مرتبط مقایسه خواهیم نمود.

۵- بحث و مقایسه

به منظور ارزیابی جامع سیستم پیشنهادی در این بخش، نتایج بدست آمده از معیارهای عملیاتی آزموده شده را با معیارهای مشابه در تحقیقات مرتبط مقایسه نموده‌ایم. بر اساس ارزیابی معیارهای عملکردی در سیستم تشخیص نفوذ پیشنهادی و مقایسه با معیارهای مشابه در تحقیقات وابسته، نتایج بدست آمده مطابق جدول ۲. قابل توصیف می‌باشد:

مطابق ارزیابی‌های تجمیع شده در جدول بالا، معیار صحت با مقدار ۹۸٫۷٪ در روش پیشنهادی بهترین عملکرد را از خود به ثبت رسانده است. متد بیزین و J48 نزدیک‌ترین نتایج را به روش پیشنهادی داشته است و مقادیر معیارها در سایر روش‌ها پایین تر از روش پیشنهادی است و روش AdaBoost پایین‌ترین مقدار معیار صحت را دارا بوده است. در معیار TP Rate روش پیشنهادی بالاترین مقدار را (معادل ۰٫۹۸۷) از خود نشان داده است و مقادیر مرتبط در روش‌های دیگر پایین‌تر از روش پیشنهادی بوده است. همچنین روش پیشنهادی پایین‌ترین مقدار را (معادل ۰٫۰۲۶) در معیار FP Rate داشته است و در مقابل روش آدا بوست بالاترین مقدار در روش‌های مورد آزمایش بوده است. با توجه به بالا بودن TP و صحت

افزایش اعتماد به سیستم است. استفاده از الگوریتم انتخاب ویژگی بهینه و آستانه‌گذاری پویا موجب شده تا میزان مثبت‌های کاذب به حداقل برسد.

- دقت مثبت: مقدار Precision درصد نمونه‌های که به درستی به عنوان حمله تشخیص داده شده‌اند، از کل نمونه‌هایی است که مدل به عنوان حمله اعلام کرده است. در سیستم پیشنهادی Precision بالای بدست آمده بیانگر توانایی دقیق در شناسایی نمونه‌های مخرب و کاهش هشدارهای مثبت کاذب می‌باشد. این دقت بالا ناشی از انتخاب ویژگی بهینه توسط الگوریتم جستجوی کلاغ و استفاده از یادگیری فدرال است که باعث می‌شود مدل فقط نمونه‌های مرتبط با حملات را شناسایی کند و نمونه‌های نرمال را اشتباه نگیرد.
- نرخ بازبایی یا حساسیت (Recall): نرخ نمونه‌های واقعی حمله است که مدل به درستی تشخیص داده است. مقدار بالای Recall نشان‌دهنده تشخیص موثر همه یا اکثریت نمونه‌های حمله در داده‌های IoT است. این معیار اهمیت بالایی دارد چون در شبکه‌های امنیتی از دست دادن حملات خطرناک False Negative می‌تواند منجر به خسارت شود. استفاده از مدل‌های یادگیری عمیق و تجمیع آنان در یادگیری فدرال باعث افزایش حساسیت مدل نسبت به رفتارهای غیرعادی می‌شود و این امر Recall بالای مدل ما را بیان می‌نماید.
- میانگین هارمونیک (F1-Score): این معیار تعادلی بین Precision و Recall ایجاد می‌کند و بیانگر عملکرد کلی مدل در شناسایی حملات است. مقدار بالای F1-Score در سیستم نشان می‌دهد که تعادل مناسبی بین کمترین هشدارهای کاذب و بیشترین تشخیص واقعی ایجاد شده است. دلیل برتری این معیار در سیستم ما ترکیب بهم افزای سه فناوری بلاک‌چین، الگوریتم تکاملی و یادگیری فدرال است که باعث شده هر دو جنبه دقت و حساسیت بالا و بهینه باشد.

جدول ۲: مقایسه عملکرد روش پیشنهادی با رویکردهای مرتبط

	J48 [۳۱و۳۴]	NaiveBayes [۲۷و۳۳]	RandomForest [۶و۳۵]	AdaBoost [۴و۱۳]	سیستم پیشنهادی
Accuracy	۹۶٫۲	۹۳٫۲	۹۲٫۸	۸۲٫۸	۹۸٫۷
TP Rate	۰٫۹۶۲	۰٫۹۳۰	۰٫۹۲۸	۰٫۸۲۸	۰٫۹۸۷
FP Rate	۰٫۰۹۹	۰٫۰۷۱	۰٫۳۱۲	۰٫۷۴۸	۰٫۰۲۶
Precision	۰٫۹۶۱	۰٫۹۴۰	۰٫۹۳۳	۰٫۸۵۸	۰٫۹۸۷
Recall	۰٫۹۶۲	۰٫۹۳۰	۰٫۹۲۸	۰٫۸۲۸	۰٫۹۸۷
F ₁ score	۰٫۹۶۱	۰٫۹۳۳	۰٫۹۶۲	۰٫۷۶۳	۰٫۹۸۷

در آزمون مقیاس پذیری و شرایط تغییر و افزایش گره‌های IoT سیستم، نتایج زیر مشاهده شد:

- در آزمون تعداد گره‌ها به میزان متغیر و برابر ۱۰۰ و ۵۰۰ و ۲۰۰۰ افزایش یافت و دقت و زمان آموزش ارزیابی گردید.
- با تغییر شرایط نهایتاً آموزش مدل جهانی تنها ۱۲۷ ثانیه افزایش یافت (از ۱۸۳ ثانیه به ۳۱۰ ثانیه).
- سربر حافظه هر گره تقریباً ثابت است (۱۵ MB $\pm$ ۲,۳).

همچنین، به علت بهره‌گیری از معماری سلسله‌مراتبی FL با سرویس دهنده‌های تجمیع منطقه‌ای، عملکرد سیستم با افزایش حجم داده‌ها، کمتر از ۸٪ افت دقت نشان داد و این نتایج نشان می‌دهند که سیستم توانایی مقیاس‌پذیری معقولی در شبکه‌های IoT دارد، بدون آنکه دقت تشخیص به شکل محسوسی افت کند.

همچنین در ارزیابی دیگر، اثر واقعی به‌کارگیری بلاک‌چین بر هزینه ارتباطی و محاسباتی در مدل پیشنهادی ارزیابی و مقایسه شد. نتایج نشان داد استفاده از بلاک‌چین تنها منجر به افزایش قابل‌مدیریت در سربر ارتباطی (+۱۵,۵٪) و تأخیر تبادل دوره‌های متحدسازی، برابر ۷ میلی ثانیه می‌شود، که عمدتاً ناشی از افزوده‌شدن هش مدل، امضاهای دیجیتال و متادیتای بلاک است. زمان ثبت هر بلاک نیز ۳,۱ میلی ثانیه اندازه‌گیری شد که در محدوده استاندارد سامانه‌های FL قرار دارد. در مقابل، بهبودهای امنیتی حاصل از بلاک‌چین شامل تضمین یکپارچگی وزن‌ها، جلوگیری از حملات Poisoning و Backdoor و فراهم شدن قابلیت پایش کامل، به‌طور معناداری ارزش افزوده ایجاد می‌کند. این نتایج نشان می‌دهد استفاده از بلاک‌چین در چارچوب پیشنهادی از نظر هزینه-امنیت کاملاً توجیه‌پذیر و عملیاتی است. نهایتاً نتایج نشان می‌دهد مدل FL-AE با دقت ۹۸,۷٪، مقدار F1 برابر ۰,۹۸۷ و نرخ FP معادل ۰,۰۲۶ عملکرد بهتری نسبت به مدل نظارت‌شده FL-DNN (با دقت ۹۷,۲٪) و نسخه متمرکز AE داشته است. دلیل علمی این برتری نیز سازگاری ذاتی AE با داده‌های بدون‌برچسب، پایداری بیشتر در شرایط Non-IID فدرال، و توانایی بهتر در شناسایی انحراف از الگوی رفتار عادی که در محیط IoT و IDS اهمیت بالایی دارد.

لازم به ذکر است، در این مطالعه، از بلاک‌چین به‌عنوان یک سیستم توزیع‌شده و غیرقابل تغییر برای تسهیل فرآیند یادگیری فدرال استفاده شده است. این فناوری به‌ویژه در مدیریت مشارکت‌کنندگان، ثبت مراحل آموزش، تأیید صحت داده‌ها و مدل‌های ارسالی، و افزایش شفافیت در فرآیند ادغام مدل‌ها نقش کلیدی ایفا می‌کند. به‌عنوان مثال، هر بار که یک دستگاه محلی مدل

روش پیشنهادی می‌توان این مقدار را برتر دانست. همچنین، روش پیشنهادی بالاترین مقدار (معادل ۰,۹۸۷) را در معیار Precision از خود به ثبت رسانده است. همچنین معیار Recall در روش پیشنهادی (معادل ۰,۹۸۷) به نسبت کلیه روش‌ها نتایج بهتری از خود ثبت رسانده است. نزدیکترین روش به روش پیشنهادی J48 است. پایین‌ترین مقادیر مربوط به روش بیزین است. همچنین AdaBoost و جنگل تصادفی نیز از روش پیشنهادی پایینتر بوده است. در نهایت در معیار F_1 score روش پیشنهادی به نسبت روش‌های دیگر در معیار F_1 score دارای مقدار بهتری بوده است و نزدیکترین مقدار به روش پیشنهادی روش J48 است. براساس تفسیر نتایج قابل به ذکر است، در تمامی ارزیابی‌ها روش AdaBoost پایین‌ترین مقدار را داشته است. به دلیل استفاده از قوانین مشخص و عملکرد قانون بیس این روش نتایج بسیار پایینی دارد و از آنجایی که قوانین و مرزهای تشخیص نفوذ در IoT بسیار نزدیک به یکدیگر بوده و مرز مشخصی را نمی‌توان برای آن در نظر گرفت این روش پایین‌ترین نتایج را داشته است. همچنین، در جدول ۳. مقایسه معیار صحت در سیستم پیشنهادی با دو پژوهش مرتبط در این زمینه انجام شده است. براساس ارزیابی نتایج، روش پیشنهادی به نسبت روش LSTM ۳,۷ درصد بهبود و به نسبت روش DNN مقدار ۱,۵ درصد بهبود ایجاد نموده است.

جدول ۳: مقایسه صحت روش پیشنهادی و روش‌های مرتبط

روش پیاده سازی	مقدار صحت (درصد)
روش پیشنهادی	۹۸,۷
LSTM [۳۱]	۹۵
DNN [۱]	۹۷,۲

همچنین به منظور بررسی مقیاس‌پذیری، آزمون مقیاس پذیری سیستم با تغییر تعداد گره‌های IoT انجام شد. برای ارزیابی مقیاس‌پذیری، حجم داده‌ها با استفاده از سه تکنیک داده‌افزایی مبتنی بر سری زمانی، SMOTE-TS^۱ و تزریق نویز موجکی، به‌ترتیب تا ۵۰ گیگابایت برای ۱۰۰ گره، ۲۸۰ گیگابایت برای ۲۰۰ گره و ۱,۲ ترابایت برای ۵۰۰ گره افزایش داده شد. نتایج بدست آمده از آزمون در جدول ۴ نمایش داده شده است:

جدول ۴: آزمون مقیاس پذیری

دقت-F1 (Score)	زمان آموزش جهانی	تعداد گره‌ها	حجم داده‌های آموزشی
۰,۹۳۲	۱۸۳s	۱۰۰	۵۰GB
۰,۹۲۸	۲۱۴s	۵۰۰	۲۸۰GB
۰,۹۲۶	۳۱۰s	۲۰۰۰	۱,۲ TB

مرتبط است. درخصوص افزایش استحکام و قابلیت اطمینان با توسعه و بکارگیری چارچوب عملیاتی پیشنهادی لازم به ذکر است، استحکام در این پژوهش به توانایی مدل در حفظ عملکرد پایدار در شرایط چالش برانگیز یادگیری فدرال، از جمله توزیع غیرهمگن داده‌ها (Non-IID)، نوسان داده‌های محلی گره‌ها و تغییر تنظیمات تابع برازندگی، اشاره دارد. برای این منظور، آزمایش‌های تابع برازندگی، تحلیل پایداری دقت در سناریوهای مختلف، و بررسی نرخ خطای مثبت کاذب در شرایط فدرال انجام و نتایج کمی آن‌ها گزارش شده است. همچنین، قابلیت اطمینان سیستم از طریق مقایسه عملکرد مدل پیشنهادی با مدل‌های متمرکز و فدرال فاقد سازوکارهای امنیتی بررسی شده است. کاهش معنادار نرخ FP، پایداری دقت در طول دوره‌های فدرال و افت محدود عملکرد در سناریوهای مقیاس‌پذیری (افزایش تعداد گره‌ها و حجم داده‌ها) مبین افزایش قابلیت اطمینان سیستم می‌باشد. براساس نتایج ارزیابی‌ها، در سیستم پیشنهادی معیارهای صحت با مقدار ۹۸٫۷٪، TP Rate با مقدار ۰٫۹۸۷، FP Rate با مقدار ۰٫۰۲۶، و معیار Recall با مقدار ۰٫۹۸۷، بهترین عملکرد را نسبت به رویکردهای مرتبط از خود به ثبت رسانده‌اند. ادغام تکنیک‌های پیشرفته مانند یادگیری انتقال فدرال و پروتکل‌های تجمیع امن توانست عملکرد و استحکام IDPS فدرال را افزایش دهد. همچنین روش آستانه‌گذاری پویا که در پژوهش حاضر بکار گرفته شده، به دقت تقریباً کاملی در طبقه‌بندی پنجره‌های ترافیک شبکه به عنوان عادی یا غیرعادی برای تشخیص حمله DDoS دست می‌یابد. استفاده این سیستم براساس آستانه‌های پویا به آن اجازه می‌دهد تا به خوبی با شرایط نوسان شبکه سازگار شود و به طور قابل توجهی نرخ مثبت و منفی کاذب را کاهش دهد. مدل یادگیری فدرال به ویژه در تفسیر ساختارهای داده پیچیده ماهر است و آن را در شناسایی تمایزات ظریف بین ترافیک عادی و حملات DDoS پیچیده با نرخ پایین بسیار موثر می‌سازد. این رویکرد برای ماهیت پویا و پیچیده ترافیک شبکه طراحی شده است و راه حلی بسیار دقیق و تطبیقی برای تمایز بین رفتار عادی و غیرعادی در پنجره‌های ترافیکی با امتیاز دقت نزدیک به عدد ۱ و مقدار ایده آل ارائه می‌دهد. به عنوان راهکارهای آتی پژوهش، در نظر گرفتن انواع آزمون امنیت و عملکرد موثر سیستم در محیط‌های عملیاتی با تنظیمات مختلف می‌تواند تضمین کند که رویکرد پیشنهادی به مؤثرترین و ایمن‌ترین شکل ممکن قابل توسعه می‌باشد.

مراجع

[۱] سنگبران، مجید، جلالی بیدگلی، امیر، ارائه یک مدل محاسبه اعتماد در شبکه اینترنت اشیا ناهمگون جهت افزایش استحکام در برابر حملات،

خود را به مرکز ارسال می‌کند، این اقدام در بلاک‌چین ثبت می‌شود و امکان بازبینی و تأیید آن فراهم می‌گردد. این رویکرد به کاهش خطر تقلب، افزایش اعتماد بین شرکت‌کنندگان و بهبود قابلیت اطمینان سیستم کمک می‌کند. همچنین، روش آستانه‌گذاری پویا، زمانی که با مدل یادگیری فدرال ادغام می‌شود، به دقت تقریباً کاملی در طبقه‌بندی پنجره‌های ترافیک شبکه به عنوان عادی یا غیرعادی برای تشخیص حمله DDoS دست می‌یابد. استفاده این سیستم براساس آستانه‌های پویا به آن اجازه می‌دهد تا با نوسان شبکه سازگار شود و نرخ مثبت و منفی کاذب را کاهش دهد. مدل یادگیری فدرال به ویژه در تفسیر ساختارهای داده پیچیده ماهر است و آن را در شناسایی تمایزات ظریف بین ترافیک عادی و حملات DDoS پیچیده با نرخ پایین بسیار موثر می‌سازد. این رویکرد برای ماهیت پیچیده ترافیک شبکه‌های توزیعی طراحی شده است و راه حلی پویا و تطبیقی برای تمایز بین رفتار عادی و غیرعادی در پنجره‌های ترافیکی با امتیاز دقت نزدیک به ایده آل ارائه می‌دهد. چنین رویکردی می‌تواند به عنوان نقطه عطفی در حوزه سیستم‌های تشخیص نفوذ در شبکه‌های مبتنی بر زیرساخت پویای اینترنت اشیا و به عنوان پلی بین دستاوردهای قبلی و ایمن‌سازی موثرتر این سیستم‌ها قابل کاربرد باشد.

۶- نتیجه‌گیری

در این پژوهش یک سیستم تشخیص نفوذ امنیتی نوین مبتنی بر تلفیق و تجمیع ساختارمند رویکردهای بلاکچین، یادگیری فدرال و الگوریتم تکاملی کلاغ، جهت افزایش امنیت شبکه‌های مبتنی بر بستر اینترنت اشیا طراحی و توسعه یافته است. هدف نهایی سیستم پیشنهادی، ارائه یک راه‌حل عملیاتی و کاربردی است که به طور موثر در حوزه سرویس‌های تشخیص نفوذ مستقر باشد. با احترام به دستاوردهای قبلی در حوزه پژوهش، تحقیقات مرتبط نیز بینش‌های ارزشمندی را در مورد تکنیک‌ها، چالش‌ها و راه حل‌های مرتبط با این موضوع ارائه کرده‌اند. با این حال، تجزیه و تحلیل تحقیقات مرتبط همچنین چندین چالش را نشان داد که باید به آنها پرداخته می‌شد، از جمله مقیاس‌پذیری، دشواری تضمین حریم خصوصی داده‌ها بدون دقت پایین، و اهمیت ارتباط و همکاری مناسب بین مشتریان شرکت‌کننده. روش پیشنهادی بهبود مستمر دقت مدل و کارایی استفاده از دستگاه را از طریق یک مکانیسم انگیزشی با استفاده از الگوریتم بهینه‌سازی جستجو کلاغ و استراتژی‌های مدیریت گره تطابقی و منطقی محقق می‌کند. ارزیابی‌های استحکام و دقت تحت سناریوهای مختلف حمله انجام شده‌اند و نتایج ارزیابی‌ها اثبات‌کننده برتری سیستم پیشنهادی درمقایسه با راهکارهای

- [19] D. Rani, et al., "Design of an intrusion detection model for IoT-enabled smart home", *IEEE Access* 11: pp. 52509-52526, 2023.
- [20] J. Hassan, et al., "A survey on blockchain-based intrusion detection systems for IoT", *VAWKUM Transactions on Computer Sciences* 11(1): pp. 138-151, 2023.
- [21] C. Liang, et al., "Intrusion detection system for the internet of things based on blockchain and multi-agent systems", *Electronics* 9(7): pp. 1120, 2020.
- [22] K. Ambili, and J. Jose, "Trust based intrusion detection system to detect insider attacks in IoT systems", in *Information Science and Applications: ICISA 2019*, Springer, pp. 631-638, 2019.
- [23] M. Nemati, N. Fareghzadeh, Investigating the applicable security encryption algorithms, *Arman Process Journal (APJ)*, 2(3):21-9, 2021.
- [24] Y. Chen, and J. Liu, "Distributed community detection over blockchain networks based on structural entropy", In *Proceedings of the 2019 ACM International Symposium on Blockchain and Secure Critical Infrastructure*, 2019.
- [25] G.D. Putra, et al., "Towards scalable and trustworthy decentralized collaborative intrusion detection system for IoT", In *2020 IEEE/ACM Fifth International Conference on Internet-of-Things Design and Implementation (IoTDI)*, 2020.
- [26] D. Laufenberg, et al., "An architecture for blockchain-enabled collaborative signature-based intrusion detection system", In *Proceedings of the 20th Annual SIG Conference on Information Technology Education*, 2019.
- [27] N. Alexopoulos, et al., "Towards blockchain-based collaborative intrusion detection systems. in *Critical Information Infrastructures Security: 12th International Conference*", CRITIS 2017, Lucca, Italy, October 8-13, 2017.
- [28] B. Hu, et al., "A collaborative intrusion detection approach using blockchain for multimicrogrid systems", *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 49(8): pp. 1720-1730, 2019.
- [29] The MathWorks Inc. MATLAB version 9.13.0 (R2022b). Natick, Massachusetts: The MathWorks Inc.; 2022. Available from: <https://www.mathworks.com>.
- [30] E. Frank, M. Hall, I.H. Witten, "The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016. Available from: https://waikato.github.io/weka-wiki/citing_weka.
- [31] N. Fareghzadeh, MA. Seyyedi, M. Mohsenzadeh, "Dynamic performance isolation management for cloud computing services", *The Journal of Supercomputing*, 74(1), pp. 417-55, 2018. <https://doi.org/10.1007/s11227-017-2135-2>.
- [32] J. Kim, M. Shim, S. Hong, Y. Shin, and E. Choi, "Intelligent detection of iot botnets using machine learning and deep learning", *Applied Sciences* 10(19):7009, 2020.
- [33] M. Vishwakarma, et al. "A new two-phase intrusion detection system with Naïve Bayes classifier", *Journal of Information Security and Applications*, 2023. <https://doi.org/10.1016/j.dajour.2023.100233>.
- [34] R. Panigrahi, S. Borah, "Rank allocation to J48 group of decision tree classifiers using binary and multiclass intrusion detection datasets", *Procedia computer science*, 2018. <https://doi.org/10.1016/j.procs.2018.05.186>.
- [35] M. Hasan, Kotov, A.I. Carcone, M. Dong, S. Naar, K.B. Hartlieb, "A study of the effectiveness of machine learning methods for classification of clinical interview fragments into a large number of categories", *Journal of biomedical informatics*, 62, pp. 21-31, 2016. <http://dx.doi.org/10.1016/j.jbi.2016.05.004>.
- مجله علمی رایانش نرم و فناوری اطلاعات، جلد ۱۰، شماره ۳، پاییز ۱۴۰۰.
- [۲] جعفری، زهرا، رنجبر، وحید، مصطفوی، سید اکبر، ارائه یک سازوکار مدیریت اعتماد مبتنی بر بلاکچین برای اینترنت اشیا، *مجله علمی رایانش نرم و فناوری اطلاعات*، جلد ۱۲، شماره ۴، زمستان ۱۴۰۲.
- [3] J. Liu, A. Wu, "Comprehensive Survey on Blockchain Technology and Its Applications", *Highlights in Science, Engineering and Technology*, 85: pp. 128-138, 2024.
- [4] N. Fareghzadeh, MA. Seyyedi, M. Mohsenzadeh, "Toward holistic performance management in clouds: taxonomy, challenges and opportunities", *Journal of Supercomputing*, 75(1), p. 272, 2019. <https://doi.org/10.1007/s11227-018-2679-9>.
- [5] N. Sharma, M. Shamkuwar, I. Singh, "The history, present and future with IoT", In *Internet of things and big data analytics for smart generation*, Cham: Springer International Publishing, pp. 27-51, 2018.
- [6] H. M. Shirazi, N. Fareghzadeh, and A. Seyyedi, "A combinational approach to service identification in SOA", *Journal of Applied Sciences Research*, 5(10), pp. 1390-1397, 2009.
- [7] A. Al Sadawi, M.S. Hassan, and M. Ndiaye, "A survey on the integration of blockchain with IoT to enhance performance and eliminate challenges", *IEEE Access* 9, pp. 54478-54497, 2021.
- [8] A. Ouaddah, A. Abou Elkalam, and A. Ait Ouahman, "FairAccess: a new Blockchain-based access control framework for the Internet of Things", *Security and communication networks* 9(18): pp. 5943-5964, 2016.
- [9] H. Wang, et al., "RDIC: A blockchain-based remote data integrity checking scheme for IoT in 5G networks", *Journal of Parallel and Distributed Computing* 152: pp. 1-10, 2021.
- [10] S.B. Rane, and Y.A.M. Narvel, "Data-driven decision making with Blockchain-IoT integrated architecture: a project resource management agility perspective of industry 4.0", *International Journal of System Assurance Engineering and Management* 13(2): pp. 1005-1023, 2022.
- [11] S. Sun, et al., "Blockchain-based IoT access control system: towards security, lightweight, and cross-domain", *IEEE Access* 9: pp. 36868-36878, 2021.
- [12] A. Amrullah, "A Review and Comparative Analysis of Intrusion Detection Systems for Edge Networks in IoT", *Intellithings Journal* 1(1): pp. 1-10, 2025.
- [13] G. Akar, et al., "L2D2: A Novel LSTM Model for Multi-Class Intrusion Detection Systems in the Era of IoMT", *IEEE Access*, 2025.
- [14] M.E. Manaa, et al., "DDoS attacks detection based on machine learning algorithms in IoT environments", *Inteligencia Artificial* 27(74): pp. 152-165, 2024.
- [15] H. Kamal, and M. Mashaly, "Advanced Hybrid Transformer-CNN Deep Learning Model for Effective Intrusion Detection Systems with Class Imbalance Mitigation Using Resampling Techniques", *Future Internet* 16(12): pp. 481, 2024.
- [16] O.P. Olawale, and S. Ebadinezhad, "Cybersecurity anomaly detection: Ai and ethereum blockchain for a secure and tamperproof ioht data management", *IEEE Access*, 2024.
- [17] B.A. Sukhni, et al., "Extracting Optimal Number of Features for Machine Learning Models in Multilayer IoT Attacks", *Sensors* 24(24): pp. 8121, 2024.
- [18] M. Esmaeili, et al., "Machine Learning-Assisted Intrusion Detection for Enhancing Internet of Things Security", *arXiv preprint arXiv:2410.01016*, 2024.

پاورقی‌ها:

- ¹ Distributed Denial of Service Attack (DDos)
- ² Industrial Internet of Things (IIoT)
- ³ Intrusion Detection System (IDS)
- ⁴ Intrusion Detection and Prevention System
- ⁵ Proof-of-stake consensus algorithm
- ⁶ False Positive Rate
- ⁷ Accuracy
- ⁸ True Positive Rate
- ⁹ Synthetic Minority Oversampling Technique for Time Series (SMOTE-TS)