دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

زمانبندی وظایف در سیستم‌های توزیع شده لایه مه و ابر محاسباتی با استفاده از الگوریتم بهینه‌سازی سوسک سرگین

رضا عزیزی[۱]، محسن اقبالی[۲]*

[۱] استادیار گروه مهندسی کامپیوتر، واحد میبد، دانشگاه آزاد اسلامی، میبد، ایران، ***azizi.reza@maybodiau.ac.ir***

[۲] دانشجوی دکتری مهندسی کامپیوتر، گروه مهندسی کامپیوتر، واحد میبد، دانشگاه آزاد اسلامی، میبد، ایران، ***m.eghbali@maybodiau.ac.ir***

چکیده

در چند سال گذشته اینترنت اشیاء رشد قابل توجهی داشته‌است و تعداد زیادی شیء هوشمند به آن متصل شده‌است. رایانش ابری به عنوان یک سیستم پردازش داده‌ها در اینترنت اشیاء است با این حال، سرورها در الگوی محاسبات ابری معمولاً در یک فاصله فیزیکی طولانی از دستگاه‌های اینترنت اشیاء قرار دارند و تأخیر زیاد ناشی از فواصل طولانی نمی‌تواند به طور مؤثر برنامه‌های اینترنت اشیاء بلادرنگ را برآورده کند. به دلیل این مسائل، محاسبات لبه و مه به عنوان فناوری محاسباتی محبوب در زمینه اینترنت اشیا ظاهر شده است. یکی از چالش‌های مهم اینترنت اشیا، مسئله زمان‌بندی وظایف در لایه مه و ابر است. در روش پیشنهادی برای تخصیص منابع آزاد از شبکه عصبی LSTM استفاده می‌شود و برای زمانبندی بهینه وظایف در لایه ابر و مه از الگوریتم بهینه‌سازی سوسک سرگین استفاده می‌شود. آزمایشات نشان می‌دهد که در مجموعه داده HPC2N دقت، حساسیت و صحت روش پیشنهادی برای پیش‌بینی وضعیت منابع به ترتیب برابر ۹۴/۷۲ درصد، ۹۳/۲۱ درصد و ۹۱/۶۴ درصد است. در مجموعه داده NASA دقت، حساسیت و صحت پیش‌بینی روش پیشنهادی در تخصیص منابع به ترتیب برابر ۹۵/۶۸ درصد، ۹۴/۶۱ درصد و ۹۲/۳۷ درصد است. روش پیشنهادی نسبت به روش‌های RNN، 1DCNN،MLP دقت بیشتری در تخصیص منابع برای زمان‌بندی دارد. شاخص Makespan روش پیشنهادی نسبت به روش‌های AO_AVOA، AVOA، PSO، HHO و FA مقدار کمتری و بهتری را در زمان‌بندی وظایف نشان می‌دهد.

کلید واژه- اینترنت اشیاء، الگوریتم بهینه‌سازی سوسک سرگین، زمانبندی وظایف، لایه ابر، لایه مه

۱-مقدمه

با تکامل فناوری‌های دیجیتال، حجم عظیمی از داده‌ها از منابع متعدد تولید می‌شود. چنین داده‌هایی را می‌توان با استفاده از راه‌حل‌های رایانش ابری ذخیره و پردازش کرد. با این حال، رایانش ابری نمی‌تواند تحرک و امنیت اینترنت اشیاء[۱] را پشتیبانی کند[۱]. اینترنت اشیاء از گره‌های توزیع شده عظیمی تشکیل شده است که در آنها هر گره متشکل از حسگرهای مختلفی است. هر حسگر در داخل یک گره متوجه تفاوت‌هایی می‌شود که در ناحیه محصور در پهنای باند مجاز خود رخ داده است. متعاقباً پس از خوشه‌بندی و ارزیابی داده‌های جمع‌آوری‌شده، تصمیمات سنتی و اقدامات ترتیب داده شده اتخاذ می‌شود. اینترنت اشیا سیستمی است که در کاربردهای مختلف مانند تولید، مراقبت‌های بهداشتی، کشاورزی، مناطق شهری و غیره کشف شده‌است[۲].

رایانش ابری[۲] یک مرحله دسترسی به داده‌ها برای پشتیبان‌گیری و پردازش داده‌ها است زیرا دارای ذخیره‌سازی و پردازش بالایی است. داده‌های گسترده تولید شده توسط دستگاه‌های اینترنت اشیا برای پردازش(تخلیه کار) به محاسبات ابری ارائه می‌شود. با این حال، انتقال مستقیم داده‌ها به ابر ممکن است باعث هزینه‌های شبکه یا به عبارتی سربار پهنای باند شود. ابر معمولا تاخیر پردازش داده‌های آپلود شده را به دلیل زمان پاسخ

[1] Internet of Things (IoT)
[2] Cloud computing(CC)

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

طولانی به تاخیر می‌اندازد[۳]. در نتیجه، یک جایگزین جدید، یعنی محاسبات مه[۳] برای رسیدگی به توانایی‌های دستگاه‌های اینترنت اشیا معرفی شد. بنابراین، محاسبات مه توسط سیسکو در سال ۲۰۱۲ برای حل این چالش‌ها ارائه شد. محاسبات مه، خدمات ابری را تا لبه شبکه گسترش می‌دهد. با انجام عملیات مربوطه با استفاده از منابع محلی که در نزدیکی دستگاه‌های لبه اینترنت اشیاء وجود دارد، زمان و حجم انتقال داده را کاهش داد. بنابراین استفاده از منابع محلی باعث کاهش هزینه‌ها، کاهش تاخیر، افزایش سطح محرمانگی و امنیت و کاهش بار ترافیک شبکه می‌شود. وقتی منابع موثری در محاسبات مه وجود نداشته‌باشد، از منابع ابری با هزینه‌های بالاتر استفاده می‌شود. محاسبات مه و محاسبات لبه اغلب به جای یکدیگر در مطالعات برای اشاره به همان مفهوم توزیع منابع محاسباتی نزدیک‌تر به کاربران نهایی استفاده می‌شوند[۴].

با اینترنت اشیا، اتصال به اینترنت فراتر از دستگاه‌های هوشمند سنتی مانند گوشی‌های هوشمند و تبلت‌ها به طیف متنوعی از دستگاه‌ها و موارد دیگر(حسگرها، ماشین‌ها، وسایل نقلیه و غیره) گسترش می‌یابد. اینترنت اشیاء حجم عظیمی از داده‌ها را تولید می‌کند که نیاز به ذخیره، پردازش و تجزیه و تحلیل برای به دست آوردن اطلاعات ارزشمند به‌منظور برآوردن اهداف و نیازهای کاربر دارند. علاوه بر این، تعداد و مقیاس برنامه‌ها و خدمات نیز به سرعت در حال افزایش است بنابراین، نیازمند قابلیت پردازشی است تا آنجا که حتی قدرتمندترین دستگاه‌های هوشمند در حال حاضر نمی‌توانند آن را برآورده کنند[۵]. محیط ابری که به عنوان یک مرکز منبع بزرگ شناخته می‌شود و امکان دسترسی همه‌جانبه به اشتراک‌گذاری و ارائه منابع به کاربران را به صورت انعطاف‌پذیر از طریق مکانیزم مجازی‌سازی فراهم می‌کند، می‌تواند یک پلتفرم بالقوه برای پشتیبانی از پیشرفت‌های اینترنت اشیا باشد. محدودیت‌های دستگاه‌های هوشمند موجود(عمر باتری، قدرت پردازش، ظرفیت ذخیره‌سازی و منابع شبکه) را می‌توان با انجام کارهای وقت‌گیر و پرمحصول به یک پلتفرم محاسباتی قدرتمند مانند محاسبات ابری به حداقل رساند و در همان حال وظایف ساده را برای دستگاه‌های هوشمند واگذار کرد[۶]. مدل محاسباتی مه- ابر دارای مزایای متعددی از جمله کاهش تأخیر، کاهش ترافیک شبکه و افزایش کارایی انرژی است، با این حال، این مدل جدید با مجموعه‌ای از چالش‌ها نیز همراه است[۷].

یکی از چالش‌های مطرح در اینترنت اشیا، تخصیص منابع[۴][۸] و زمان بندی وظایف[۵][۹] است. یک سیستم بسیار توزیع شده مانند محاسبات مه-ابر یک پلت فرم ایده آل برای استقرار برنامه‌های کاربردی است؛ آن دسته از برنامه‌های موازی که وظایف آنها مستقل از یکدیگر است. برنامه‌های کاربردی در بسیاری از سناریوها ظاهر می‌شوند، از جمله داده‌کاوی، جست‌وجوهای گسترده (مانند شکستن کلید)، محاسبات فراکتال، زیست‌شناسی محاسباتی، تصویربرداری کامپیوتری، رمزگذاری/رمزگشایی ویدیو و برنامه‌های کاربردی مختلف اینترنت اشیاء[۹]. چالش اصلی، زمان‌بندی وظایف در مجموعه گره‌های پردازشی از جمله گره‌های ابری (مانند سرورها یا ماشین‌های مجازی) و گره‌های مه است. هدف از زمان‌بندی وظایف در سیستم مه-ابر به نفع کاربران یا ارائه‌دهندگان خدمات است. از طرف کاربران، آنها نگران برخی از معیارهای ساخت، بودجه، مهلت، امنیت و هزینه هستند. از سوی دیگر، هدف ارائه‌دهندگان خدمات، متعادل کردن بار، استفاده از منابع و بهره‌وری انرژی است. برای تضمین QoS، زمان پاسخگویی نقش مهمی در زمانی که مستقیماً بر تجربه کاربر تأثیر می‌گذارد، ایفا می‌کند[۱۰]. علاوه بر این، هزینه پیاده‌سازی نیز جنبه‌ای است که بسیار مورد علاقه و اقبال کاربران بوده‌است. یک برنامه زمانبندی کار که زمان تکمیل را به حداقل می‌رساند و در هزینه های پولی صرفه‌جویی می‌کند، توافق نامه سطح خدمات[۶] امضا شده با کاربران را برآورده می‌کند[۱۱]. مدیریت وظایف[۷] در معماری‌های ابر-مه یک موضوع مهم است. استفاده بهینه از منابع ابر مه برای افزایش پارامترهای کیفی مختلف مانند زمان اجرای کار، هزینه عملیات و مصرف انرژی موضوع مهمی است. زمان‌بندی کار مناسب در محیط مه، هزینه‌ها و تاخیرهای پردازش/ارتباطات را کاهش می‌دهد. یکی از مشکلات محققان، انتخاب یک روش زمانبندی کارآمد است[۱۲].

مسئله زمانبندی کار، دشواری چند جمله‌ای غیر قطعی[۸] و بهینه‌سازی چالش برانگیز است. امروز، الگوریتم فرابتکاری کاربردی می‌تواند برای حل مسئله زمانبندی وظایف، مورد استفاده قرار گیرد. از جمله روش‌های فرابتکاری که برای زمانبندی وظایف مورد استفاده قرار گرفته‌است، می‌توان

[3] Fog Computing (FC)
[4] Resource allocation
[5] Tasks scheduling
[6] Service Level Agreement (SLA)
[7] Task management
[8] NP-hard

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

به الگوریتم ژنتیک[۱۳]، الگوریتم بهینه‌سازی ذرات[۱۴]، الگوریتم بهینه‌سازی وال[۱۵] و الگوریتم بهینه‌سازی شاهین[۱۶] اشاره کرد. هدف از این مقاله، ارایه یک رویکرد جدید برای در نظر گرفتن منابع ابر و مه برای زمانبندی کارها و وظایف در اینترنت اشیا است. ارایه یک تابع هدف جدید در زمانبندی برای کاهش تاخیر اجرای وظایف از دیگر اهداف تحقیق به‌شمار می‌رود. در روش پشنهادی برای زمانبندی وظایف از الگوریتم بهینه‌سازی سوسک سرگین استفاده می‌شود تا تاخیر اجرای وظایف کاهش داده شود. ارایه یک نسخه گسسته از الگوریتم سوسک سرگین[۱۷] برای زمانبندی وظایف در لایه مه و ابر، نخستین نوآوری مقاله است. نوآوری دیگر مقاله ارایه یک تابع هدف کارآمد برای زمانبندی وظایف است. به‌کارگیری منابع در لایه مه و ابر برای زمانبندی بهینه، دیگر نوآوری این مقاله به‌شمار می‌رود. پیش‌بینی بار محاسباتی منابع با استفاده از شبکه عصبی LSTM برای اجرای وظایف روی منابع آزاد از نوآوری‌های دیگر مقاله است. سهم نویسندگان در ارایه یک الگوریتم زمانبندی کارآمد در اینترنت اشیا در موارد ذیل خلاصه می‌شود:

- پیش بینی منابع آزاد با شبکه عصبی LSTM و زمانبندی وظایف طبق منابع آزاد
- ارایه یک نسخه گسسته از الگوریتم بهینه‌سازی سوسک سرگین برای زمانبندی
- زمانبندی دو سطحی در لایه مه و ابر
- تلفیق هوش گروهی و یادگیری عمیق در زمانبندی بهینه
- ارایه یک تابع هدف جدید برای زمانبندی وظایف در لایه مه و ابر

این مقاله یک روش زمانبندی دو سطحی در لایه مه و ابر و دارای ۵ بخش است. در بخش II کارهای مرتبط در زمینه زمانبندی کارها ارایه می‌شود. در بخش III، سیستم زمانبندی دو سطحی بر اساس یادگیری عمیق و الگوریتم بهینه‌سازی سوسک سرگین توسعه داده شده است. در بخش IV، روش پیشنهادی پیاده‌سازی و با روش‌های مشابه مورد مقایسه قرار می‌گیرد. در بخش V نتایج تحقیق و یافته‌های تحقیق به همراه پیشنهادهای آتی ارایه می‌شود.

۲-کارهای مرتبط

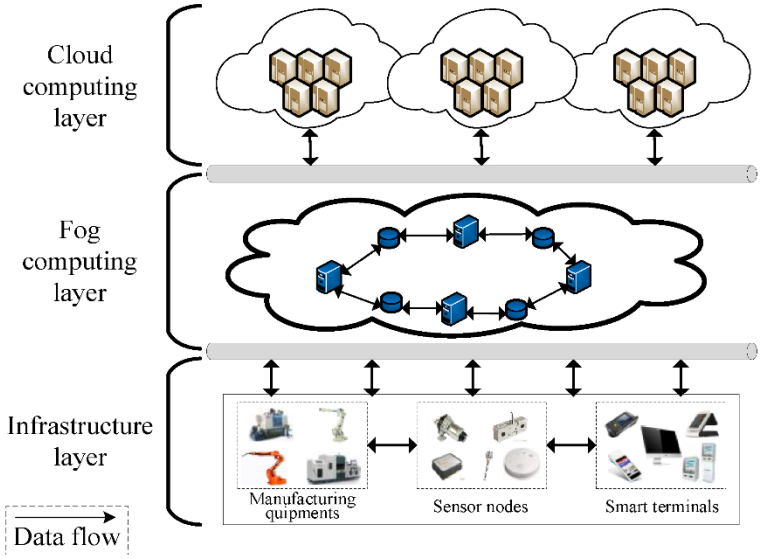
رایانش ابری با استفاده از تکنیک‌های محاسبات موازی و توزیع شده، منابع رایانه مشترک را از طریق اینترنت در دسترس مشتریان قرار می‌دهد. مدل کسب‌وکار «پرداخت به‌حساب» رایانش ابری را تقریباً دموکراتیک کرده است. ارائه دهندگان ابر، ارائه دهندگان خدمات و کاربران نهایی در این مرحله از استقرار نرم افزار شرکت می‌کنند. ارائه دهندگان خدمات ابری قابلیت‌های محاسباتی را از طریق رایانه‌های مجازی به مشتریان خود ارائه می‌دهند. ارائه‌دهندگان خدمات از این ماشین‌های مجازی برای خدمات مشتری در سطح برنامه استفاده می‌کنند.
ارائه‌دهندگان خدمات، الگوریتم‌های زمان‌بندی کار را برای پخش مشاغل مشتری در سراسر ماشین‌های مجازی، کاهش زمان پاسخگویی، اطمینان از کیفیت بالای خدمات و به حداکثر رساندن منابع پیاده‌سازی می‌کنند. به همین دلیل، الگوریتم زمان‌بندی کار، بخش مهمی از هر معماری ابری است. رایانش ابری نیاز به تنظیماتی برای چندین تکنیک زمانبندی مورد استفاده در محیط‌های مختلف رایانه دارد. ممکن است یک روش زمانبندی بهینه شده برای یک خوشه در فضای ابری ضعیف عمل کند. قبل از اینکه الگوریتم بتواند با ساختار محیط ابری مقابله کند، بخش‌های روش باید به فضای مشکل منتقل شوند. هرچه تنوع ماشین‌های مجازی و اندازه بارهای کاری که مدیریت می‌شوند بیشتر باشد، تعداد پیکربندی‌های کار موجود بیشتر است. یافتن کوتاه‌ترین مسیر در میان همه جایگشت‌های بالقوه یکی از چالش برانگیزترین مشکلات در علم کامپیوتر برای زمانبندی است. معماری اینترنت اشیاء به طور معمول مانند شکل۱، از سه لایه شامل لایه زیرساخت، لایه محاسبات مه و لایه محاسبات ابری تشکیل شده است[۱۸]:

[9] Genetic algorithms(GA)
[1] Particle swarm optimization (PSO) algorithm
[1] Whale optimization algorithm(WOA)
[1] Harris hawks optimization(HHO) algorithm
[1] Dung beetle optimizer

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

- لایه زیرساخت: این لایه از دستگاه‌های پایانه با عملکردهای مختلف مانند سنسور(حسگر)های مختلف، دستگاه‌های پردازشی و پایانه‌های مختلف هوشمند تشکیل شده است. پایانه‌های هوشمند، وظایف ساده را به صورت محلی انجام می‌دهند اما قادر به انجام کارهای پیچیده در زمان واقعی نیستند.
- لایه محاسبات مه: این لایه عمدتاً از گره‌های مه تشکیل شده‌است. اینها سرورهایی با قابلیت‌های محاسباتی، ارتباطی و ذخیره‌سازی معین در خطوط تولید هو شمند مانند حـ سگرهای هو شمند، د ستگاه‌های پردازش هو شمند و د ستگاه‌های چندر سانه‌ای هو شمند هســتند. این لایه می‌تواند درخواســت‌های پایانه‌های خط تولید هوشــمند را حس کرده و خدمات مختلفی را بلادرنگ ارائه دهد که می‌تواند تاخیر پردازش وظایف را تا حد زیادی کاهش دهد و کیفیت سرویس اپلیکیشن‌های بلادرنگ را تضمین کند.
- لایه محا سبات ابری: لایه رایانش ابری شامل خو شه‌هایی با ظرفیت محا سبات و ذخیره سازی عظیم ا ست که خدمات از راه دور را برای خطوط تولید هوشمند به‌منظور انجام وظایف محاسباتی پیچیده ارائه می‌دهد.



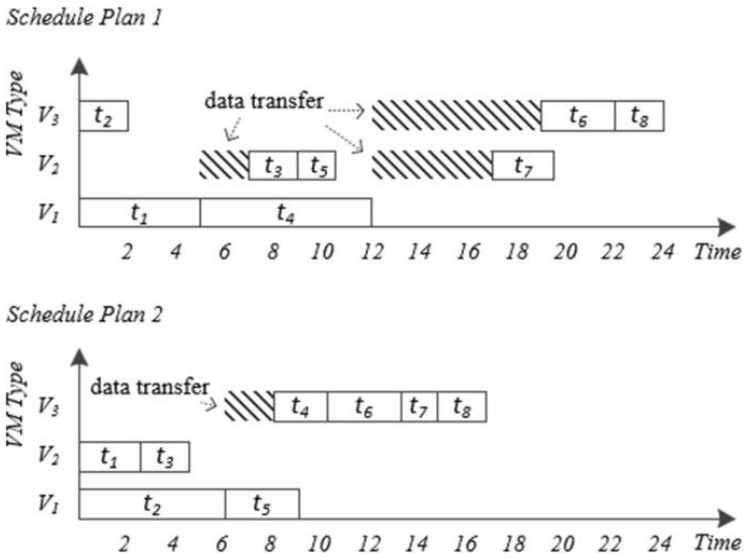
شکل ۱: معماری سه لایه اینترنت اشیاء[۱۸]
Figure 1. The three-layer architecture of the Internet of Things[18]

تعادل بار یکی از مهم‌ترین مشکلات برای اجرای وظایف زمان‌بندی چند منبعی در یک محیط محاسباتی ناهمگن است. اگر رویکرد متعادل کننده بار در برنامه کمک شود، می‌تواند بر مصرف انرژی تأثیر بگذارد. استفاده از CPU نیز برای هر دو قرارداد سطح سرویس و مصرف انرژی اجباری است که مستقیماً بر زمان‌بندی وظایف متعادل‌سازی بار تأثیر می‌گذارد. استفاده از CPU بر پردازش کار تأثیر می‌گذارد و هر گره CPU خود را برای پردازش وظایف دارد. بنابراین، روش پیشنهادی با هدف افزایش زمان‌بندی کار برای بهبود پردازش کار انجام می‌شود. ایجاد تعادل نگرانی بین منابع و وظایف، پاسخ پردازش وظایف و مدیریت منابع را به طور قابل توجهی بهبود می‌بخشد. متعادل‌سازی بار یک تکنیک ضروری در مدیریت منابع است که با تعبیه عوامل مدیریت وظیفه می‌تواند به یک سیستم قوی تبدیل شود. زمان‌بندی شامل یافتن یک نقشه‌برداری بهینه برای تخصیص n وظیفه به پردازنده‌های متعدد است و علاوه بر این، چندین متغیر حیاتی برای افزایش زمان‌بندی در رایانش ابری وجود دارد[۱۹]. زمان‌بندی را می‌توان یک نگاشت از اجرای وظایف روی منابع آزاد در لایه مه یا ابر در نظر گرفت. در شکل۲، یک فرآیند زمانبندی در لایه ابر و روی ماشین‌های مجازی نمایش داده شده است. با توجه به شکل۲، وظایف t1 تا t8 را می‌توان روی منابع مختلف V1، V2 و V3 اجرا کرد. در حالت اول وظیفه اول و دوم در ماشین مجازی اول و سوم اجرا شده اما در سناریوی دوم وظیفه یک و دوم روی ماشین مجازی اول و دوم اجرا شده است. اجرای ترتیبی کارها و وظایف روی ماشین‌های مجازی باعث می‌شود که زمانبندی‌های مختلفی برای اجرای وظایف در لایه مه و ابر

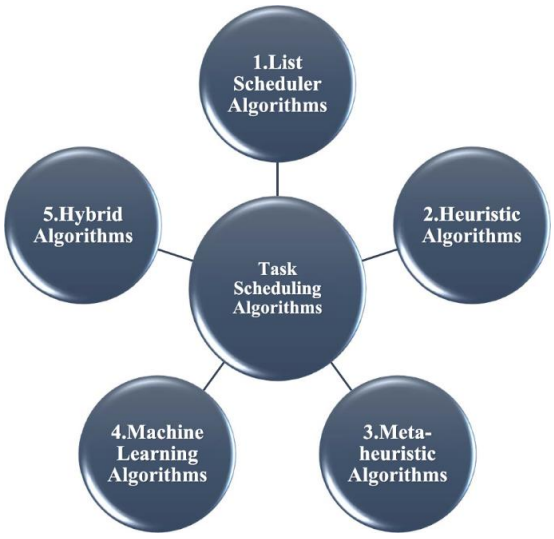
دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

در نظر گرفته شود و از این رو یافتن اجرای وظایف بهینه از اهمیت بالایی برخوردار است[۲۰]. برای انجام زمانبندی، تاکنون روش‌های مختلفی ارایه شده که یک دسته‌بندی از آن در شکل۳، قابل مشاهده است. با توجه به شکل مورد نظر برای زمانبندی وظایف روش‌های مبتنی بر لیست، اکتشافی، فراابتکاری، یادگیری ماشین و ترکیبی ارایه شده است[۲۱].



شکل ۲: اجرای چند وظیفه در منابع اینترنت اشیاء[۲۰]
Figure 2. Multitasking in Internet of Things resources[20]



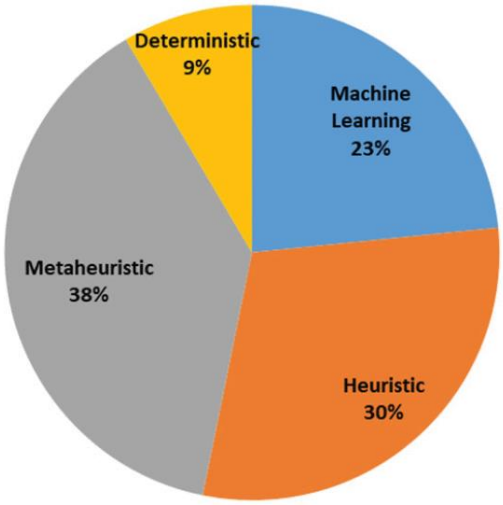
شکل ۳: روشهای مختلف زمانبندی در اینترنت اشیاء[۲۱]
Figure 3. Different scheduling methods in the Internet of Things[21]

مکانیسم‌های یادگیری ماشین را می‌توان در زمان‌بندی وظایف مه برای بهبود کارایی و دقت فرآیند زمان‌بندی استفاده کرد. با تجزیه و تحلیل داده‌های تاریخی و شناسایی الگوها، الگوریتم‌های یادگیری ماشین می‌توانند بارهای ترافیکی آینده را پیش‌بینی و بر این اساس تخصیص وظایف را بهینه کنند. این می‌تواند منجر به کاهش تاخیر و بهبود عملکرد سیستم و همچنین استفاده بهتر از منابع شود.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

علاوه بر این، یادگیری ماشین می‌تواند برای شناسایی ناهنجاری‌ها یا خرابی‌های احتمالی در سیستم و اتخاذ اقدامات پیشگیرانه برای جلوگیری از آنها استفاده شود. به طور کلی، ترکیب مکانیسم‌های یادگیری ماشین در زمان‌بندی وظایف مه می‌تواند اثربخشی فرآیند را افزایش داده و عملکرد کلی سیستم‌های محاسباتی مه را بهبود بخشد[۲۲]. مکانیسم‌های مبتنی بر اکتشاف برای زمان‌بندی وظایف مه روش‌هایی هستند که از قوانین سرانگشتی یا بهترین شیوه‌ها برای تخصیص وظایف به دستگاه‌ها در یک محیط محاسباتی مه استفاده می‌کنند. این مکانیسم‌ها بر تکنیک‌های یادگیری ماشین تکیه نمی‌کنند، بلکه بر قوانین تصمیم‌گیری ساده مبتنی بر تجربه یا شهود هستند. مکانیسم‌های مبتنی بر اکتشاف می‌توانند ساده و آسان برای پیاده‌سازی باشند، اما ممکن است همیشه به تخصیص بهینه کار منجر نشوند. آنها همچنین مانند رویکردهای مبتنی بر یادگیری ماشینی، با شرایط در حال تغییر سازگار نیستند یا از تجربیات گذشته درس نمی آموزند. بخش زیر چندین روش زمان‌بندی مه مبتنی بر اکتشافی را بررسی می‌کند که از عوامل مختلفی برای دستیابی به نتایج زمان‌بندی بهبودیافته استفاده می‌کنند[۲۲]. در الگوریتم‌های فراابتکاری، یک فضای حل تصادفی برای زمان‌بندی کار استفاده می‌شود. با تغییرات اندکی در الگوریتم‌های فراابتکاری، می‌توان از آنها برای حل مسائل مختلف بهینه‌سازی استفاده کرد. الگوریتم‌های فراابتکاری، مستقل از مسئله هستند. به‌کارگیری محاسبات مه با استفاده از الگوریتم‌های فراابتکاری در بهبود استفاده از منابع و کاهش تأخیر در محیط‌های محاسباتی مه توزیع شده امیدوارکننده است[۲۲]. مکانیسم‌های قطعی همیشه خروجی یکسانی را برای یک ورودی معین تولید می‌کنند. مکانیسم‌های قطعی برای زمان‌بندی وظایف مه شامل استفاده از مجموعه‌ای از قوانین و الگوریتم‌های از پیش تعریف شده برای تخصیص وظایف به گره‌های مه است. جست‌وجوی جامع نمونه‌ای از رویکرد قطعی به مسئله زمان‌بندی کار است. الگوریتم‌های جامع کل فضای جست‌وجو را برای یافتن طرح بهینه بر اساس مدل هزینه داده شده برمی‌شمارند. این مکانیسم‌ها شامل هیچ روش تصادفی یا احتمالی نیستند. مکانیسم‌های قطعی برای زمان بندی وظایف مه ساده و آسان برای پیاده‌سازی هستند. با این حال، ممکن است همیشه منجر به استفاده بهینه از منابع یا کاهش تأخیر نشوند[۲۲]. شکل۴، سهم انواع روش‌های زمانبندی و بخصوص الگوریتم‌های فراابتکاری و ابتکاری را نشان می‌دهد[۲۲].



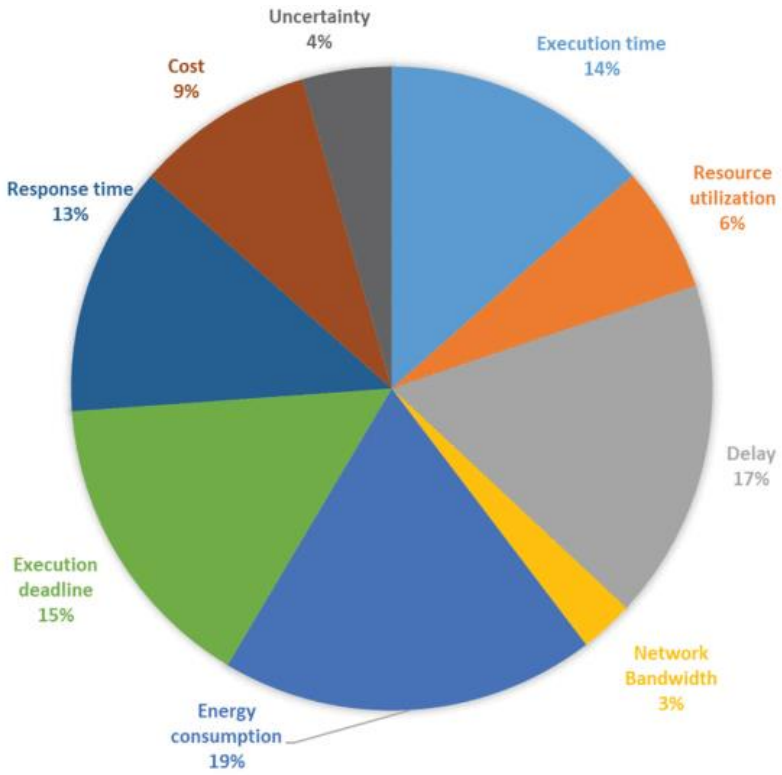
شکل ۴: سهم روش‌های مختلف زمانبندی در اینترنت اشیا[۲۲]
Figure 4. Contribution of different scheduling methods in Internet of Things[22]

با توجه به نمودار فوق، الگوریتم‌های زمانبندی فراابتکاری و ابتکاری در زمانبندی کارها و وظایف دارای سهمی به ترتیب و در حدود ۳۸ درصد و ۳۰ درصد است.

در زمانبندی وظایف شاخص‌های مختلفی به‌کار گرفته می‌شود که تعدادی از آنها در شکل۵، نمایش داده شده است. با توجه به شکل مورد نظر می‌توان گفت که مصرف انرژی، مهلت زمانی در اجرای وظایف، تاخیر اجرا و زمان اجرا از مهمترین شاخص‌های ارزیابی در زمانبندی است[۲۲].

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279



شکل ۵: شاخص های اصلی در زمانبندی[۲۲]
Figure 5. Main indicators in timing[22]

برای زمانبندی وظایف در لایه مه یا ابر تاکنون چندین مطالعه انجام شده است که بیشتر آنها بر محور روش‌های بهینه‌سازی و فراابتکاری است که در ادامه تعدادی از کارهای مرتبط در این زمینه مرور می‌شود.

در[۲۳]، سال ۲۰۲۳، یک روش زمان‌بندی وظایف و متعادل‌سازی منابع محاسبات مه در کارخانه هوشمند ارایه شده‌است. در این پژوهش یک الگوریتم زمان‌بندی ژنتیکی تقسیم بازه‌ای بهبود یافته یا الگوریتم زمان‌بندی تقسیم فاصله زمانی بر اساس الگوریتم ژنتیک برای برنامه‌ریزی و تخصیص وظایف در کارخانه هوشمند پیشنهاد شده‌است. نتایج شبیه‌سازی نشان می‌دهد که در مقایسه با الگوریتم زمان‌بندی پیش‌فرض Kubernetes، IDGSA می‌تواند زمان پردازش داده‌ها را تا ۵۰ درصد کاهش دهد و استفاده از منابع محاسباتی مه را تا ۶۰ درصد بهبود بخشد. در مقایسه با الگوریتم ژنتیک سنتی، با تکرارهای کمتر، روش آنها می‌تواند زمان پردازش داده‌ها را تا ۷ درصد کاهش دهد و استفاده از منابع محاسباتی مه را تا ۹ درصد بهبود بخشد و در مقایسه با روش معمولی Joines&Houck، الگوریتم پیشنهادی می‌تواند بسیار سریع‌تر همگرا شود و نتایج بهینه‌سازی بهتری را ارایه کند. علاوه بر این، شبیه‌سازی نشان می‌دهد که روش آنها در محاسبات مشترک ابری و مه می‌تواند کل تاخیر کار را به ترتیب ۱۸ و ۷ درصد در مقایسه با محاسبات فقط ابری و فقط مه کاهش دهد.

در[۲۴]، سال ۲۰۲۳، مروری بر تکنیک‌های زمان‌بندی وظایف در محاسبات ابری و مه انجام شده است. زمان‌بندی کارآمد بر روی ابر برای استفاده بهینه از منابع در مراکز داده حیاتی است. با ظهور برنامه‌های نسل پنجم و اینترنت اشیاء که تعداد زیادی کار با الزامات تأخیر دقیق ایجاد می‌کنند، چالش‌برانگیزتر شده است. این امر محاسبات مه / لبه را به وجود می‌آورد که یک لایه مکمل برای ابر محاسباتی است. تأخیر وظایف در محاسبات مه را می‌توان کاهش داد زیرا پردازش در شبکه به دستگاه‌های نهایی و کاربران نزدیک تر انجام می‌شود، اما به دلیل در دسترس بودن منابع محدود، نمی‌توان هر کار را در مه برنامه‌ریزی کرد. الگوریتم‌های زمان بندی مرسوم اغلب در بهره‌برداری از منابع ناهمگن شکست می‌خورند. بنابراین، الگوریتم‌های زمان‌بندی به‌خوبی طراحی‌شده و به‌خوبی تنظیم شده برای دستیابی به کیفیت بهتر خدمات مورد نیاز هستند. در این مطالعه، الگوریتم‌های زمان‌بندی کار پیشرفته در محیط‌های ابر و مه در مجموعه‌ای از ابعاد مختلف مورد بررسی قرار می‌گیرند. مطالعات نشان

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

داد ۹۷ درصد از مطالعات بر اهداف چندگانه تمرکز می‌کنند و ۶۸ درصد از تکنیک‌ها غیر قطعی هستند. علاوه بر این، در مجموع بیست هدف زمان‌بندی مختلف با زمان ساخت، استفاده از منابع، تاخیر، متعادل‌سازی بار و مصرف انرژی به‌عنوان مهم‌ترین معیارها شناسایی شده‌اند. روش‌های ارزیابی شامل شبیه‌سازی (۵۱ درصد)، آزمایش‌های واقعی (۴درصد)، معادلات تحلیلی (۲ درصد) و مجموعه داده‌ها (۴۳ درصد) و غیره بررسی می‌شوند. در پایان، مسائل باز، چالش‌ها و جهت‌گیری‌های آتی مورد بحث قرار می‌گیرد.

در[۲۵]، سال ۲۰۲۳، یک روش زمان‌بندی کار بهینه در شبکه مه و ابر با استفاده از الگوریتم پروانه-شعله چندهدفه ارایه دادند. آزمایشات آنها نشان داد روش پیشنهادی از الگوریتم بهینه‌سازی ازدحام ذرات، الگوریتم کرم شب تاب، الگوریتم‌های ازدحام سالپ، بهینه ساز شاهین هریس و کلونی زنبورهای مصنوعی در زمانبندی تاخیر کمتری ارایه می‌کند. بر اساس آزمایش‌ها، راه‌حل پیشنهادی زمان تکمیل وظایف اینترنت اشیا و زمان خروجی را کاهش داده‌است، بنابراین تأخیر ناشی از پردازش وظایف، مصرف انرژی و انتشار CO_2 و افزایش نرخ عملکرد سیستم را کاهش داده است.

در[۲۶]، سال ۲۰۲۳، برای زمان‌بندی کار در معماری لبه-مه-ابر یک رویکرد متعادل کننده بار چند هدفه با استفاده از الگوریتم یادگیری تقویتی ارایه دادند. در این پژوهش، یک الگوریتم زمان‌بندی مه یادگیری تقویتی برای رسیدگی زمانبندی وظایف پیشنهاد شده‌است. نتایج تجربی نشان می‌دهد که الگوریتم پیشنهادی تعادل بار را افزایش داده و زمان پاسخ را در مقایسه با الگوریتم‌های زمان‌بندی موجود کاهش می‌دهد. علاوه بر این، الگوریتم پیشنهادی از نظر تعداد دستگاه‌های مورد استفاده از سایر رویکردها بهتر عمل می‌کند.

در[۲۷]، سال ۲۰۲۳، یک الگوریتم جدید الهام گرفته از طبیعت برای زمان‌بندی وظایف بهینه در سیستم مه-ابر ارایه دادند. در سال های اخیر، به دلیل رشد تصاعدی داده‌های تولید شده توسط دستگاه‌های هوشمند متصل به هم، استفاده از برنامه‌های کاربردی اینترنت اشیاء مبتنی بر ابر مه به طور پیوسته در حال افزایش بوده‌است. با این حال، ارائه‌دهندگان ابری که مسئول این برنامه‌های کاربردی اینترنت اشیا هستند با دو مشکل اساسی نحوه محافظت از سیستم در برابر کاربران غیرقابل اعتماد و نحوه تخصیص واحدهای پردازش برای برآورده کردن خواسته‌ها با هزینه‌های قابل قبول روبرو هستند. در این پژوهش، یک نسخه بهبود یافته از الگوریتم بهینه‌سازی مبتنی بر انتخاب زندگی[۴][۱]برای حل زمان‌بندی کار ارایه شده‌است. زمان‌بندی وظایف یکی از برجسته‌ترین مشکلات در تخصیص منابع است. الگوریتم پیشنهادی آنها نه تنها سرعت همگرایی را افزایش می‌دهد، بلکه تنوع را بهتر حفظ می‌کند و قدرت، تأخیر و هزینه را بهینه می‌کند. آزمایشات نشان می‌دهد الگوریتم بهینه‌سازی مبتنی بر انتخاب زندگی با دستیابی به نتایج بهتر برای تأخیر و مصرف انرژی از روش‌های پیشرفته مشابه بهتر عمل می‌کند.

در[۲۸]، سال ۲۰۲۳، یک الگوریتم جست‌وجوی گرانشی چند هدفه برای زمان‌بندی وظایف در سیستم محاسباتی ابری مبتنی بر اینترنت اشیا ارایه شده است. در این پژوهش، یک تابع چند هدفه جدید برای به حداقل رساندن هزینه ساخت، انرژی و هزینه پولی برای زمان‌بندی کار در سیستم مه-ابر پیشنهاد شده است. در این پژوهش، عملکرد الگوریتم پیشنهادی در دو سناریو مورد ارزیابی قرار گرفته‌است. ابتدا، عملکرد روش پیشنهادی با برخی از روش‌های بهینه‌سازی چندهدفه محبوب در برخی از توابع تست استاندارد مقایسه می‌شود. نتایج نشان می‌دهد که الگوریتم پیشنهادی نسبت به سایر روش‌ها عملکرد بهتری دارد. سپس از این الگوریتم برای یافتن راه حل مناسب برای مسئله زمانبندی کار استفاده می‌شود. رویکرد پیشنهادی مصرف انرژی را تا ۲۲ درصد و هزینه پردازش را تا ۴۰ درصد بهبود بخشیده است. تجزیه و تحلیل آماری نشان می‌دهد که در[۲۹]، سال ۲۰۲۳، یک استراتژی زمان‌بندی کار مبتنی بر الگوریتم بهینه‌سازی کرکس‌های آفریقایی برای محاسبات مه-ابر ارایه دادند. این پژوهش یک استراتژی برنامه‌ریزی وظایف مبتنی بر الگوریتم بهینه‌سازی کرکس‌های آفریقایی پیشرفته برای محاسبات مه-ابر پیشنهاد می‌کند. کمینه‌سازی زمان ساخت، هزینه و مصرف انرژی در الگوریتم پیشنهادی به عنوان تابع هدف در نظر گرفته می‌شود. برای اولویت بندی وظایف، بهترین و بدترین روش به‌منظور رسیدگی به حساسیت تاخیرهای کار استفاده می‌شود. وظایف حساس به تأخیر، به محیط مه فرستاده می‌شوند، در حالی که وظایف تحمل تأخیر به ابر ارسال می‌شوند. نتایج شبیه‌سازی نشان می‌دهد روش آنها شاخص makepan را تا ۵۳ درصد و مصرف انرژی را تا ۴۴ درصد بهبود می‌دهد.

در[۳۰]، سال ۲۰۲۳، یک روش زمان‌بندی بهینه در شبکه اینترنت اشیا-مه-ابر با استفاده از ترکیب الگوریتم بهینه‌ساز عقاب طلایی و الگوریتم بهینه‌سازی کرکس‌های آفریقایی ارایه دادند. در این پژوهش مشکل زمان‌بندی وظایف در محیط مه-ابر برای اجرای درخواست‌های وظایف اینترنت اشیاء به عنوان یک مشکل بهینه‌سازی با توجه به نیازهای بیشتر QoS در نظر گرفته شده است. ترکیبی از الگوریتم بهینه‌ساز عقاب طلایی و الگوریتم بهینه‌سازی کرکس‌های آفریقایی برای حل مشکل زمان بندی کار در محیط مه-ابر استفاده شده است. روش پیشنهادی آنها در زمانبندی

[1] Life-choice-based optimization algorithm (ILCO)

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

برای بهبود فرآیند اکتشاف الگوریتم کرکس استفاده شد و مرحله اکتشاف الگوریتم عقاب طلایی با فاز بهره برداری کرکس ترکیب شد. کمینه‌سازی تابع زمان انجام برای ارزیابی الگوریتم بهینه‌ساز برای بهینه‌سازی مسئله زمان‌بندی و یافتن بهترین ماشین‌های مجازی در محیط ابر-مه به‌کار گرفته شده است. روش پیشنهادی آنها در زمانبندی برای دو مجموعه داده با استفاده از معیارهای زمان ساخت، مقدار تابع تناسب، PIR و زمان توان اعمال شد. آزمایشات نشان داد روش پیشنهادی آنها در زمانبندی در مقایسه با الگوریتم‌های HHO، FA، PSO، AVOA و AO از نظر معیارهای ارزیابی مانند تأخیر بهتر عمل می‌کند.

در[۳۱]، سال ۲۰۲۴، یک روش زمانبندی مبتنی بر یادگیری تقویتی عمیق برای بهینه‌سازی بار سیستم و زمان پاسخگویی در محیط‌های محاسباتی لبه و مه ارایه شده‌است. از آنجایی که تعداد زیادی از برنامه‌های اینترنت اشیاء نیاز به اجرا در منابع لبه/مه دارند، ممکن است سرورها بیش از حد بارگذاری شوند. از این رو، ممکن است سرورهای لبه/مه را مختل کند و همچنین بر زمان پاسخگویی برنامه‌های اینترنت اشیا تأثیر منفی بگذارد. علاوه بر این، بسیاری از برنامه‌های کاربردی اینترنت اشیا از اجزای وابسته تشکیل شده‌اند که محدودیت‌های اضافی را برای اجرای خود متحمل می‌شوند. علاوه بر این، محیط‌های محاسباتی لبه/مه و برنامه‌های اینترنت اشیا ذاتاً پویا و تصادفی هستند. بنابراین، زمان‌بندی کارآمد و سازگار برنامه‌های اینترنت اشیا در محیط‌های محاسباتی لبه/مه ناهمگن، از اهمیت بالایی برخوردار است. با این حال، منابع محاسباتی محدود در سرورهای لبه/مه بار اضافی را برای اعمال تکنیک‌های بهینه اما محاسباتی نیازمند تحمیل می‌کند. برای غلبه بر این چالش‌ها، آنها یک الگوریتم برنامه‌ریزی و زمانبندی مبتنی بر یادگیری تقویتی را پیشنهاد دادند تا زمان پاسخ‌دهی برنامه‌های ناهمگن اینترنت اشیا را به طور سازگار و کارآمد بهینه کند و بار سرورهای لبه/مه را متعادل کند. نتایج به‌دست‌آمده نشان می‌دهد که روش آنها در مقایسه با الگوریتم‌های فراابتکاری و سایر تقویت‌کننده‌ها، به ترتیب تا ۵۵، ۳۷ درصد و ۵۰ درصد هزینه اجرای برنامه‌های اینترنت اشیا را از نظر تعادل بار، زمان پاسخ و هزینه وزنی کاهش می‌دهد.

در[۳۲]، سال ۲۰۲۴، رویکرد مبتنی بر الگوریتم بهینه‌سازی هریس هاکس و سیستم فازی برای بهبود زمان‌بندی وظایف مبتنی بر ابر را ارایه دادند. زمان‌بندی وظایف مستلزم تخصیص وظایف مختلف به ماشین‌های مجازی است. در نتیجه، الگوریتم‌های زمان‌بندی برای دستیابی به مجموعه‌ای از اهداف، از جمله کاهش زمان ساخت، به حداقل رساندن مصرف انرژی، افزایش بهره‌وری منابع، دستیابی به متعادل‌سازی بار و بهینه‌سازی هزینه‌ها، با دقت ساخته می‌شوند. با توجه به اهمیت عمیق این اهداف، الگوریتم‌هایی که برای چنین سناریوهایی طراحی شده اند، همواره اهداف متعددی را در بر می‌گیرند. این پژوهش یک الگوریتم زمان‌بندی کار چندهدفه ابتکاری را برای محاسبات ابری معرفی می‌کند که به طور یکپارچه الگوریتم بهینه‌سازی هریس هاکس را ادغام می‌کند و قدرت منطق فازی را در خود جای می‌دهد. این روش، از الگوریتم شاهین برای کشف فضای راه حل گسترده استفاده می‌کند در حالی که راه حل‌های تولید شده را از طریق منطق فازی در معرض ارزیابی دقیق قرار می‌دهد. آزمایشات نشان داد روش آنها در حدود ۷۳ درصد باعث کاهش در مصرف انرژی و ۱۹ درصد کاهش چشمگیر هزینه می‌شود.

۳- روش پیشنهادی

روش پیشنهادی در این بخش برای زمانبندی وظایف، ارایه و معرفی می‌شود. در روش پیشنهادی در فاز اول با استفاده از شبکه عصبی LSTM میزان بار هر منبع پیش‌بینی می‌شود و در ادامه هر منبع که دارای بار محاسباتی کمتری باشد برای به‌کارگیری زمانبندی تخصیص داده می‌شود. برای تخصیص وظایف به منابع که می‌تواند ماشین‌های فیزیکی و مجازی باشد از الگوریتم بهینه‌سازی سوسک سرگین استفاده می‌شود. روش پیشنهادی برای تخصیص منابع از الگوریتم LSTM و برای زمانبندی از الگوریتم سوسک سرگین استفاده می‌شود. مراحل روش پیشنهادی برای تخصیص منابع و زمانبندی در شکل۶، نمایش داده شده است.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

شکل۶: چارچوب سیستم پیشنهادی برای تخصیص منابع و زمانبندی وظایف
Figure 6. A proposed system framework for resource allocation and task scheduling

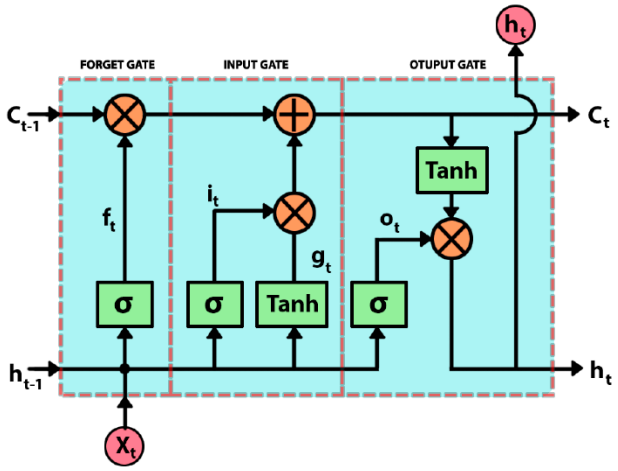
۱-۳- پیش بینی منابع در دسترس با LSTM

یکی از چالش‌های مهم انواع روش‌های زمانبندی بی‌توجهی به تخصیص منابع است. در بیشتر مطالعات مرتبط با زمانبندی، فرض بر آن است که همه منابع در دسترس است اما در عمل، برخی از منابع که می‌توانند ماشین‌های فیزیکی یا مجازی باشند دارای بار محاسباتی زیادی می‌شوند. استفاده از منابعی که مرتبط اشغال می‌شوند باعث می‌شود تا وظایف به کندی روی آنها اجرا شود. یک رویکرد مناسب آن است که وظایف روی منابعی اجرا شود که بار محاسباتی آنها زیاد نشود. به عبارت بهتر یک مکانیزم پیش‌بینی برای اشغال یا آزاد بودن منابع می‌تواند به زمانبندی بهینه کمک نماید. برای پیش‌بینی وضعیت منابع در روش پیشنهادی از شبکه عصبی مصنوعی LSTM استفاده می‌شود زیرا این شبکه، توانایی بالایی برای پیش‌بینی رویدادها دارد. واحدهای LSTM توسعه شبکه‌های تکراری[1]هستند که به آن‌ها اجازه می‌دهد تا ورودی‌های خود را برای مدت طولانی‌تری مانند حافظه رایانه به خاطر بسپارند. می‌تواند اطلاعات را از سلول خود بخواند، بنویسد و حذف کند، اطلاعات را ذخیره کند و از ناپدید شدن آنها در طول زمان جلوگیری کند. سلول بر اساس درجه اهمیتی که به آن اختصاص داده شده است تصمیم می‌گیرد اطلاعاتی را که دریافت می‌کند ذخیره یا حذف کند. این اهمیت از طریق وزن اتصالات تعیین می‌شود، بنابراین LSTM در طول زمان یاد می‌گیرد که تشخیص دهد چه بخشی از اطلاعات مهم است و چه چیزی مهم نیست. اجرای یک حالت داخلی سلول که حافظه بلندمدت ایجاد می‌کند. همانطور که در شکل۷، نشان داده شده است، یک نورون LSTM از سه گیت: ورودی، فراموشی و خروجی تشکیل شده است که تعیین می‌کند آیا ورودی جدید مجاز است، کدام اطلاعات حذف می‌شود یا اینکه اجازه دارد در زمان فعلی بر خروجی تأثیر بگذارد[۳۳].

[1] Recurrent networks

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279



شکل ۷: ساختار یک نورون عصبی در شبکه LSTM
Figure 6. The structure of a neuron in a network

ساختار نورون LSTM (دروازه فراموشی، دروازه ورودی، دروازه خروجی) در شکل فوق نشان داده شده است. σ مخفف تابع سیگموئید، Xt داده ورودی و ht خروجی شبکه جاری است. اولین دروازه‌ای که از چپ به راست در شکل فوق ظاهر می شود، دروازه فراموشی است که تصمیم می‌گیرد چه اطلاعاتی از حالت سلول (Ct) از طریق یک تابع سیگموئید (σ) کنار گذاشته شود و مطابق معادله۱، فرموله می‌شود:

$$f_t = \sigma(W_i * [h_{t-1}, x_t] + b_i] \quad (۱)$$

که در آن (ht-1) خروجی قبلی واحد، (xt) ورودی، (Wf) ماتریس وزن دروازه فراموشی و (bf) بایاس گیت است. خروجی بین مقداری از ۰ تا ۱ خواهد بود که ۰ نشان دهنده حذف کل و ۱ کل اطلاعات را حفظ می‌کند. پس از این مرحله، دروازه ورودی ادامه می‌یابد که نشان می‌دهد چه بخشی از اطلاعات جدید در حالت سلول ذخیره می‌شود. این دروازه از دو تابع یکی سیگموئیدی و دیگری مماس هذلولی تشکیل شده است که میزان به‌روز رسانی حالت را تعیین می‌کند که در معادلات ۲ و ۳ فرموله شده‌اند.

$$t_i = \sigma(W_i * [h_{t-1}, x_t] + b_i] \quad (۲)$$

$$g_t = \tanh(W_g * [h_{t-1}, x_t] + b_g] \quad (۳)$$

ماتریس وزن مطابق معادله۴، با حذف اطلاعات غیر ضروری و به روز رسانی آن با اطلاعات جدید به دست آمده، وضعیت جدید واحد به دست می‌آید.

$$C_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (۴)$$

در نهایت خروجی واحد را تعیین کنید (ht)، که نسخه‌ای از حالت سلولی است که توسط ورودی فعلی و خروجی قبلی فیلتر شده است. جایی که W_o و b_o ماتریس وزن و بایاس برای دروازه خروجی هستند که در معادلات ۵ و ۶ فرموله شده است:

$$\sigma_t = \sigma(W_o[h_{t-1}, x_t] + b_o \quad (۵)$$

$$h_t = o_t * \tanh(C_t) \quad (۶)$$

در روش پیشنهادی شبکه عصبی LSTM با اجراهای قبلی از سیستم آموزش داده می‌شود. ورودی شبکه LSTM منابع و خروجی آن پیش‌بینی اشغال یا آزاد بودن منبع است و هدف از به‌کارگیری LSTM در روش پیشنهادی، اجرای وظایف در منابع آزاد است.

۲-۳- تابع هدف زمانبندی

هنگامی که درخواست‌های لایه حسگر به لایه مه در قالب وظایف ارسال می‌شوند، به وظایف کوچک و مستقل تجزیه می‌شوند تا در زیرساخت محاسباتی مه و ابر پردازش شوند. هر کار دارای ویژگی‌هایی نظیر تعداد دستورالعمل‌ها، حافظه مورد نیاز، اندازه فایل‌های ورودی و خروجی

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

است[۳۴]. با فرض اینکه Tk نشان دهنده k امین وظیفه باشد، در هر بار مجموعه‌ای از n وظیفه مستقل به صورت معادله۷، به سیستم ارسال می‌شود:

$$T = \{T_1, T_2, T_3, ..., T_n\} \quad (۷)$$

زیرساخت محاسباتی مه و ابر، از پردازنده‌ها، یعنی گره‌های ابری و گره‌های مه که ویژگی‌های یکسانی مانند نرخ CPU، هزینه استفاده از CPU، هزینه استفاده از حافظه و هزینه استفاده از پهنای باند را دارند، تشکیل شده است. با این حال، گره‌های ابری معمولاً قدرتمندتر از گره‌های مه هستند اما هزینه استفاده از آنها بیشتر است. در روش پیشنهادی فرض می‌شود که m پردازنده در گره‌های لایه مه و ابر قرار دارد و تعداد گره‌های لابه مه و ابر با f و c نمایش داده می‌شود ($N=N_{cloud} \cup N_{Fog}$) و به صورت معادله۸ بیان می‌شود:

$$N = \{N_1, N_2, N_3, ..., N_m\} \quad (۸)$$

که در آن Ni گره پردازش شماره i را نشان می‌دهد. هر کار Tk به پردازنده Ni اختصاص داده می‌شود که به صورت Tik نمایش داده می‌شود. یک پردازنده را می‌توان برای پردازش مجموعه‌ای از یک یا چند کار اختصاص داد که در معادله۹، ارایه شده است:

$$N_i Tasks = \{T_x^i, T_y^i, ..., T_z^i\} \quad (۹)$$

مسئله زمانبندی کار در محیط محاسباتی مه و ابر را می‌توان به صورت جست‌وجوی مجموعه‌ای معادله ۱۰، فرموله کرد:

$$NodeTasks = \{T_1^a, T_2^b, T_3^c ..., T_n^p\} \quad (۱۰)$$

برای مجموعه‌ای از وظایف NiTasks، زمان اجرا (EXT) آن گره نیاز به تکمیل تمام وظایف محول شده به شرح رابطه۱۱ محاسبه می‌شود:

$$EXT(N_i) = \sum_{T_k^i \in N_i Tasks} Exetime(T_k^i) = \frac{\sum_{T_k^i \in N_i Tasks} length(T_k)}{CPUrate(N_i)} \quad (۱۱)$$

که در آن ExeTime (T_k^i) زمان اجرای Tk پردازش شده در گره Ni است که به صورت معادله۱۲، محاسبه می‌شود:

$$ExeTime(T_k^i) = \frac{lentgth(T_k^i)}{CPUrate(N_i)} \quad (۱۲)$$

با length(Tk) طول زمانی اجرای وظیفه Tk و CPUrate (N_i)، نرخ پردازش در پردازنده گره i ام است. برای تکمیل تمام وظایف، از زمانی که درخواست دریافت می‌شود تا زمانی که آخرین کار تکمیل یا زمانی که آخرین ماشین تمام می‌شود، تعریف شده است که با Makespan نمایش داده شده و با فرمول معادله۱۳، تعیین می‌شود:

$$Makespan = \underset{1 \langle i \langle m}{Max}[EXT(N_i)] \quad (۱۳)$$

فرض کنید MinMakespan حد پایینی Makespan باشد، یعنی سیستم به کوتاه‌ترین زمان نیاز دارد تا تمام آن کار را کامل کند. در حالت ایده‌آل، زمانی که همه گره‌ها همه وظایف محول شده را همزمان به پایان برسانند با MinMakespan نشان داده شده و با معادله۱۴، فرموله می‌شود و می‌توان آن را مانند معادله۱۵، فرموله کرد:

$$MinMakespan = EXT(N_1) = ... = EXT(N_m) \quad (۱۴)$$

$$MinMakespan = \frac{\sum_{1 \le k \le n} length(T_k)}{\sum_{1 \le i \le n} CPUrate(N_i)} \quad (۱۵)$$

MinTotalCost کمترین هزینه‌ای است که برای انجام مجموعه‌ای از وظایف T در سیستم مه و ابر لازم است و زمانی به‌دست می‌آید که هر وظیفه به ارزان ترین گره اختصاص داده شود. با ارائه اطلاعات هر گره، به راحتی می‌توان تعیین کرد که کدام گره وظیفه Tk را با کمترین هزینه پردازش می‌کند که با MinCost(Tk) نمایش داده می‌شود، بنابراین، MinTotalCost یک مجموعه از وظایف T خاص است و می‌تواند به صورت معادله۱۶ تعریف شود:

$$MinTotal\,Cost = \sum_{T_k \in T} Min\,cost(T_k) = \sum_{T_k \in T} Min_{1 \le i \le m}(Cost(T_k^i)) \quad (۱۶)$$

با توجه به مطالب ارایه شده یک تابع هدف مناسب می‌تواند به صورت معادله۱۷، در نظر گرفته شود:

$$\min(F) = \alpha * \frac{MinMakespan}{Makespan} + (1-\alpha) * \frac{MinTotal\,Cost}{Total\,Cost} \quad (۱۷)$$

اگر Alpha برابر ۰/۵ باشد؛ یعنی زمان و هزینه در بهینه‌سازیT اولویت یکسانی دارند. زمانی که Alpha از ۰/۵ بیشتر باشد نشان دهنده آن است که حداقل رساندن زمان ساخت با اولویت بالاتر از هزینه کل تمرکز می‌کند، این مورد، زمانی است که کاربر مایل است پول بیشتری برای

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

دستیابی به عملکرد بهتر بپردازد. برعکس، زمانی که Alpha از ۰/۵ کمتر است، هزینه بیشتر از زمان، اولویت بندی می‌شود یعنی کاربر، بودجه محدودی دارد. هدف، بهینه سازی زمان اجرا و هزینه پردازش است، یعنی یافتن راه حل به گونه ای که TotalCost حداقل و نزدیک به MinTotalCost باشد.

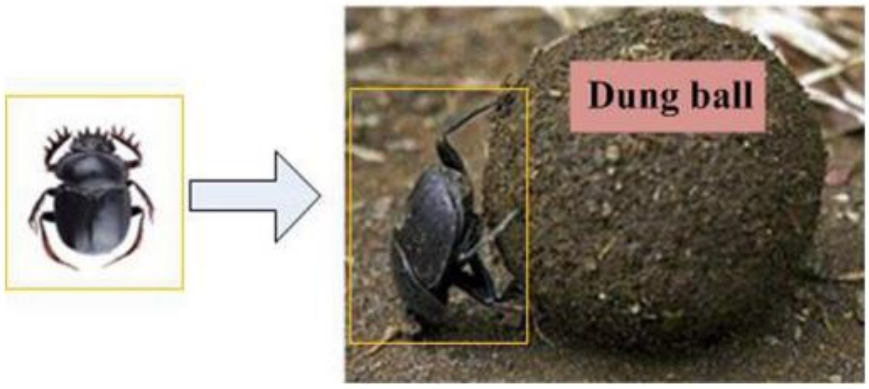
۳-۳- زمانبندی با الگوریتم سوسک سرگین

در روش پیشنهادی برای بهینه‌سازی نگاشت وظایف به منابع آزاد در لایه مه و ابر از الگوریتم بهینه‌سازی سوسک سرگین استفاده می‌شود. در روش پیشنهادی هر نگاشت از مجموعه‌ای از وظایف روی تعدادی منبع به عنوان یک راه‌حل در الگوریتم بهینه‌سازی سوسک سرگین در نظر گرفته می‌شود. در شکل۷، نگاشت مجموعه‌ای از وظایف به منابع در لایه مه و ابر نمایش داده شده است که این ماتریس را می‌توان یک سوسک سرگین در نظر گرفت. در اینجا فرض می‌شود که ۱۰ وظیفه وجود دارد که روی سه منبع آزاد زمانبندی می‌شوند. صفر نشان دهنده آن است که وظیفه در منبع اجراء نمی‌شود و یک نشان دهنده اجرای وظیفه در منبع است[۳۴].

	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}
N_1	0	1	0	0	1	0	0	0	1	0
N_2	0	0	0	1	0	1	1	0	0	0
N_3	1	0	1	0	0	0	0	1	0	1

شکل ۷: ساختار یک راه حل در روش پیشنهادی برای زمانبندی وظایف
Figure 7. Structure of a solution in the proposed method for task scheduling

الگوریتم بهینه‌سازی سوسک سرگین از رفتار سوس‌کهای سرگین مدلسازی و فرموله شده است. گونه های مختلفی برای سوسک‌های سرگین وجود دارد، مانند Copris ochus Motschulsky، Onthophagus gibbulus، Caccobius jessoensis Harold و غیره. معروف است که سوسک سرگین به عنوان یک حشره معمولی در طبیعت از سرگین حیوانات تغذیه می‌کند. سوسک‌های سرگین در اکثر نقاط جهان یافت می‌شوند و به عنوان تجزیه کننده در طبیعت عمل می‌کنند، به این معنا که آنها در اکوسیستم، اهمیت حیاتی دارند. تحقیقات نشان داده است که سوسک‌های سرگین عادت جالبی دارند؛ سرگین را به شکل توپ در می‌آورند سپس آن را حرکت در می‌آورند، همانطور که در شکل۸، نشان داده شده است. شایان ذکر است هدف سوسک‌های سرگین این است که گوی سرگین خود را به سرعت و کارآمدتر حرکت دهند که این کار می‌تواند از تکمیل شدن آنها توسط سایر سوسک‌های سرگین جلوگیری کند[۱۷].



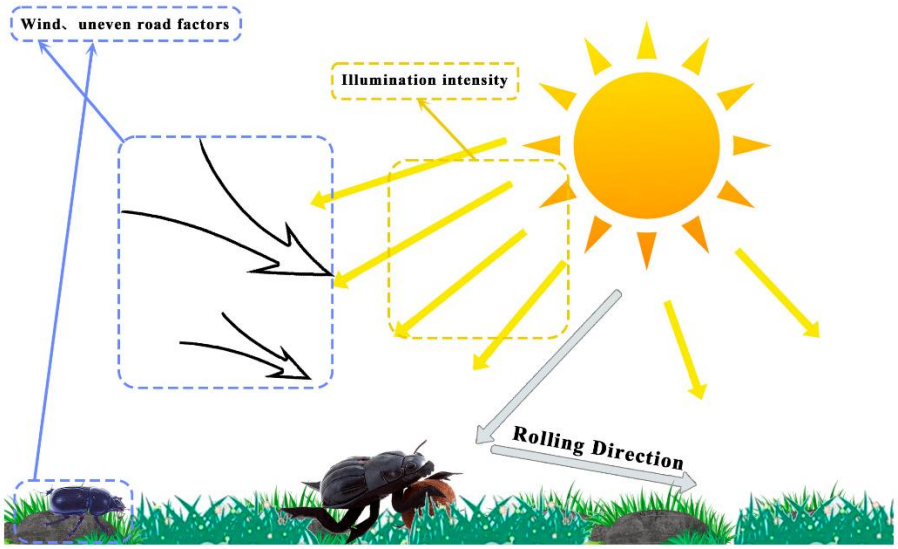
شکل ۸: حرکت دادن سرگین توسط سوسک سرگین
Figure 8. Dung moving by dung beetles

الگوریتم بهینه‌سازی سوسک سرگین یک الگوریتم جدید بهینه‌سازی هوش ازدحامی است که از رفتار اجتماعی جمعیت سوسک‌های سرگین الهام گرفته شده است. این سوسک‌ها سرگین را قبل از اینکه به مکانی امن بغلتانند، به صورت توپ، فشرده می‌کنند. آن‌ها می‌توانند توپ‌های سرگین بسیار بزرگ‌تری را بغلتانند و از نشانه‌های آسمانی برای چرخاندن آن‌ها در یک خط مستقیم در صورت وجود منبع نور استفاده کنند. با این حال، در غیاب منبع نور، مسیرهای آنها منحنی و مستعد اختلالات طبیعی می‌شود. بقای سوسک‌های سرگین به طور پیچیده‌ای با به دست

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

آوردن گلوله‌های سرگین مرتبط است، جایی که برخی برای تولید مثل و پرورش فرزندان‌شان استفاده می‌شوند در حالی که بقیه به عنوان غذا مورد استفاده قرار می‌گیرند. الگوریتم بهینه‌سازی سوسک سرگین پنج رفتار کلیدی را که توسط سوسک‌های سرگین به‌کار گرفته می‌شود، شبیه‌سازی می‌کند و این رفتارها توپ غلتاندن، رقصیدن، جستجوی غذا، دزدی و تولید مثل است. جمعیت سوسک‌های سرگین به چهار زیر گروه تقسیم می‌شوند که عبارتند از: غلتک‌ها، بازتولیدکنندگان، خردسالان و دزدها و هر نوع سوسک استراتژی‌های جستجوی متفاوتی برای خود در نظر می‌گیرند[۱۷]. در طول فرآیند غلتیدن، سوسک‌های سرگین باید با استفاده از نشانه‌های آسمانی، به‌ویژه خورشید و ماه حرکت کنند تا مسیر حرکت مستقیم توپ سرگین را حفظ کنند. در شکل۹، مدل مسیر سوسک سرگین نمایش داده شده است. می‌توان مشاهده کرد که سوسک‌های سرگین از خورشید برای جهت‌یابی استفاده می‌کنند، با فلش نشان دهنده، جهت چرخش آنها مشخص شده است.[۳۵].



شکل ۹: حرکت دادن سرگین توسط سوسک سرگین
Figure 9. Dung moving by dung beetles

با فرض اینکه شدت منبع نور نیز بر مسیر غلتشی سوسک‌های سرگین تأثیر می‌گذارد، موقعیت سوسک به روز می‌شود و می‌توان آن را به صورت معادله۱۸، نشان داد:

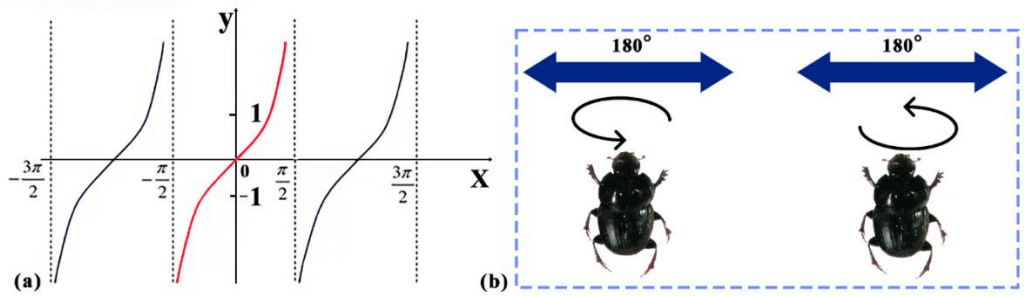
$$X_i(t+1) = X_i(t) + \alpha \times k \times X_i(t+1) + b \times \Delta x \quad (۱۸)$$

$$\Delta x = | X_i(t) - X^w | \quad (۱۹)$$

در این معادلات t نشان دهنده تعداد تکرار فعلی است. Xi(t) نشان دهنده اطلاعات موقعیت سوسک i ام در تکرار t است. پارامتر k یک عدد ثابت بین ۰ و ۰/۲ است که ضریب انحراف را نشان می دهد. b یک مقدار ثابت متعلق به بازه صفر و یک است. α یک ضریب طبیعی است که به صورت ۱ یا ۱- اختصاص داده می‌شود. Xw نشان دهنده بدترین موقعیت است و Δx تغییر شدت نور را شبیه‌سازی می‌کند. پارامتر a عوامل طبیعی (مانند باد و زمین ناهموار) را شبیه‌سازی می‌کند که می‌تواند باعث انحراف سوسک سرگین از جهت اصلی خود شود. به طور خاص، a =1 نشان دهنده عدم انحراف است، در حالی که a= −1 نشان‌دهنده انحراف از جهت اصلی است. Δx بزرگتر به معنی منبع نور ضعیف‌تر است که دو مزیت را به همراه دارد که عبارتند از: کاوش کامل کل فضای مشکل در طول فرآیند بهینه‌سازی و افزایش قابلیت‌های جستجو و کاهش احتمال به دام افتادن در بهینه محلی. عوامل طبیعی مختلف، مانند باد و زمین ناهموار، می‌توانند تأثیری قابل توجه بر مسیر حرکت سوسک‌های سرگین داشته باشند. در چنین شرایطی، سوسک‌های سرگین معمولاً به روی گوی سرگین می‌روند و رفتار رقصانی را انجام می‌دهند که شامل یک سری چرخش و مکث است. آنها از طریق این رفتار رقصی، جهت حرکت خود را با تغییر جهت تعیین می‌کنند و در نتیجه مسیر جدیدی را به دست می‌آورند. برای تقلید از این رفتار رقصیدن، یک تابع مماس برای به دست آوردن جهت چرخش جدید استفاده می‌شود. شکل۱۰، مدل تابع مماس و مدل رقص سوسک‌های سرگین را نشان می‌دهد. توجه به این نکته مهم است که فقط مقادیر تعریف شده در بازه ۰ و ۱ تابع مماس باید در نظر گرفته شوند[۳۵]:

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279



شکل ۱۰: تابع مماس و مدل رقص سوسک های سرگین. (الف) تابع مماس در جهت نور. (ب) مدل رقص سوسک های سرگین.

Figure 10. Tangent function and dance model of dung beetles. (a)Tangent function in light direction. (b) Dance model of dung beetles

هنگامی که جهت صحیح با موفقیت مشخص شد، سوسک غلتکی باید به چرخاندن توپ به جلو ادامه دهد. در این مرحله، به روزرسانی موقعیت سوسک غلتکی به شرح معادله ۲۰، است:

$$X_i(t+1) = X_i(t) + \tan\beta \,|\, X_i(t) - X_i(t-1) \,|$$ (۲۰)

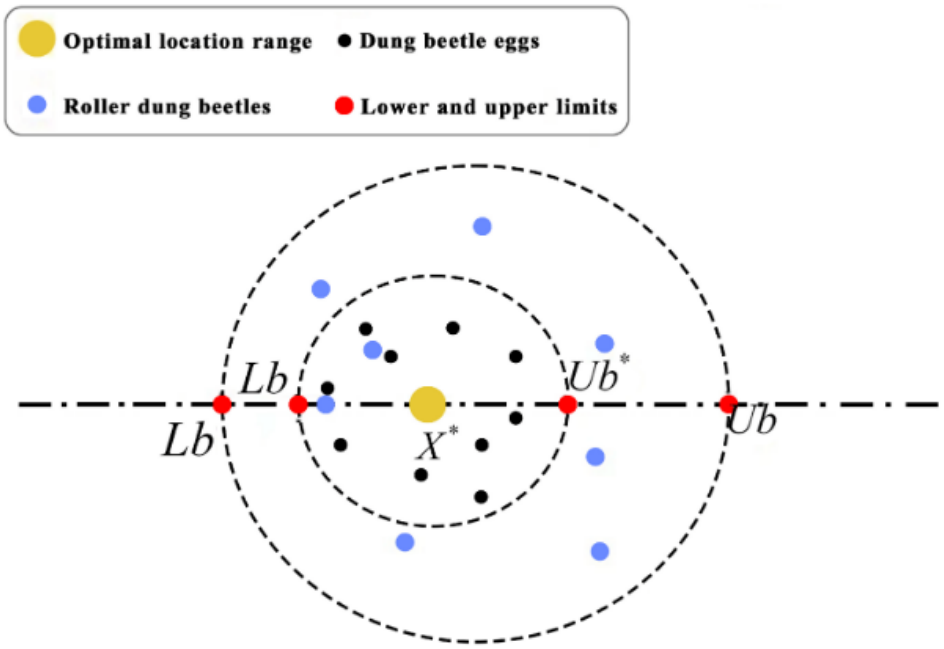
که در آن زاویه انحراف $\beta \in [0, \pi]$ است و در معادله، t تعداد تکرارهای جاری را نشان می‌دهد. Xi(t) اطلاعات موقعیت سوسک غلتکی i در تکرار t را نشان می‌دهد. |(Xi(t) – Xi(t – 1)| تفاوت مطلق بین موقعیت i ام سوسک در تکرار t و موقعیت آن در تکرار قبلی (t - 1) است.

می‌توان مشاهده کرد که به روز رسانی موقعیت سوسک غلتکی ارتباط نزدیکی با اطلاعات موقعیت فعلی و سابق آن دارد. توجه به این نکته ضروری است که اگر زاویه انحراف برابر با ۰، $\pi/2$ یا π باشد، موقعیت سوسک به روز نمی‌شود. برای ایجاد یک محیط امن برای فرزندان خود، انتخاب محل تخمگذاری مناسب برای سوسک‌های سرگین بسیار مهم است. با الهام از بحث فوق الذکر، یک استراتژی انتخاب مرز برای شبیه‌سازی منطقه‌ای که سوسک‌های ماده سرگین تخم می‌گذارند پیشنهاد شده است و به صورت زیر معادله ۲۱ و ۲۲ فرموله می‌شود:

$$\min(X^* * (1-R), Uh)$$ (۲۱)

$$\min(X^* * (1-R), Lh)$$ (۲۲)

که در آن X^*نشان دهنده بهترین موقعیت محلی فعلی است. Lh^* و Uh^* به ترتیب مرزهای بالایی و پایینی ناحیه تخمگذاری را نشان می‌دهند. R = 1 - t/Tmax، که در آن Tmax حداکثر تعداد تکرار را نشان می‌دهد. Lh و Uh به ترتیب کران‌های بالایی و پایینی فضای جستجو هستند، همانطور که در شکل ۱۱، نشان داده شده است[۳۵].



دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

شکل ۱۱: مدل مفهومی استراتژی انتخاب مرز
Figure 11. Conceptual model of boundary selection strategy

شایان ذکر است که هر گلوله سرگین حاوی یک تخم یک سوسک سرگین است. علاوه بر این، نقاط قرمز مرزهای بالایی و پایینی را نشان می‌دهد. در الگوریتم سوسک سرگین فرض بر این است که هر سوسک سرگین ماده در هر تکرار فقط یک تخم می‌گذارد. با توجه به تغییرات دینامیکی در محدوده مرزی در طول تکرار، این به جلوگیری از به دام افتادن الگوریتم در بهینه محلی کمک می‌کند که عمدتاً توسط مقدار وزن اینرسی R تعیین می‌شود. بنابراین، موقعیت توپ‌های تخم‌مرغ نیز در طول فرآیند تکرار پویا است و با معادله ۲۳ می‌شود:

$$Y_i(t+1) = X^* + b_1(Y_1(t) - Lh^*) + b_2(Y_i(t) - Uh^*) \quad (۲۳)$$

که در آن Yi(t) اطلاعات موقعیت i امین توپ تخم مرغ را در تکرار t-امین نشان می‌دهد، X^* نشان دهنده بهترین موقعیت محلی است و از طرفی Lh^* و Uh^* نشان دهنده مرزهای بالایی و پایینی ناحیه تخم‌گذاری است. از طرفی ($Yi(t)-Lh^*$)، ($Yi(t)$- b2 یک بردار تصادفی در محدوده ۰ و ۱ است و D بعد مسئله بهینه سازی را نشان می‌دهد. برخی از سوسک‌های سرگین بالغ در جستجوی غذا در زمین فرو می‌روند و این نوع سوسک‌ها به عنوان سوسک‌های سرگین کوچک شناخته می‌شوند. برای شبیه‌سازی فرآیند جستجوی علوفه سوسک‌های سرگین کوچک، تعیین منطقه بهینه علوفه ضروری است. معادله شبیه‌سازی برای این ناحیه به صورت زیر معادله۲۴ و ۲۵ مشخص می‌شود:

$$Lm = \max(X^h * (1-R), Lh) \quad (۲۴)$$

$$Lm = \min(X^h * (1-R), Lh) \quad (۲۵)$$

به گونه‌ای که Xh نشان دهنده بهترین موقعیت سراسری است. Lm و Um به ترتیب مرزهای پایین و بالایی منطقه جستجوی بهینه هستند. Lh و Uh مرزهای پایین و بالایی فضای جستجو هستند. بنابراین، به روز رسانی موقعیت برای سوسک‌های سرگین کوچک به شرح ۲۶، است:

$$X_i(t+1) = X_i(t) + C_1 * (X_i(t) - Lm) + C_2(X_i(t) - U_m) \quad (۲۶)$$

که در آن Xi (t) نشان دهنده موقعیت i-امین سوسک سرگین در تکرار t-ام است. C1 –(t) عددی است که به طور تصادفی به دنبال توزیع نرمال ایجاد می‌شود و C2 یک بردار تصادفی در محدوده ۰ و ۱ است. با توجه به اینکه برخی از سوسک‌های سرگین که به آنها دزد می‌گویند، توپ‌های سرگین را از سایر سوسک‌ها می‌دزدند، اطلاعات موقعیت سارق در طول فرآیند تکرار به‌صورت زیر به‌روزرسانی می‌شود. بر اساس رابطه ۲۶، می توان مشاهده کرد که Xh منبع غذایی بهینه را نشان می‌دهد. با فرض اینکه منطقه اطراف Xh نشان دهنده مکان اصلی برای غذای رقابتی است، به روز رسانی موقعیت برای دزد مطابق معادله(۲۷)، ارائه می شود:

$$X_i(t+1) = X^h + W * g * |X_i(t) - X^*| + |X_i(t) - X^h| \quad (۲۷)$$

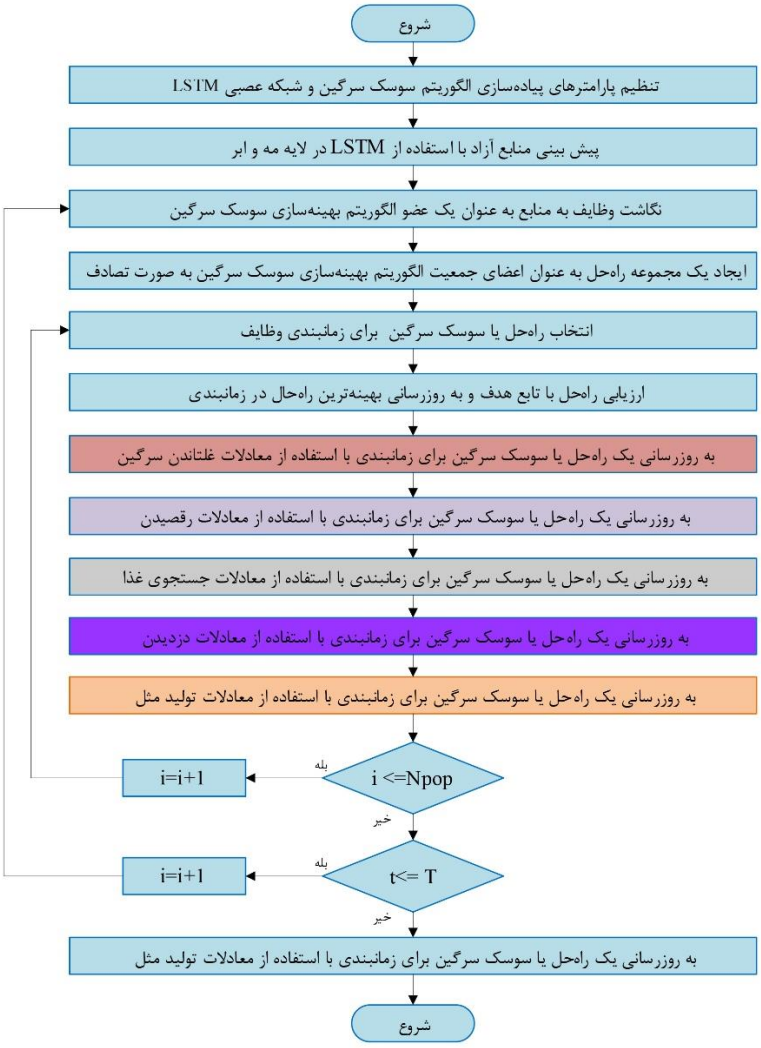
که در آن Xi(t) اطلاعات موقعیت i-امین سارق را در تکرار t-امین نشان می‌دهد، Xh نشان دهنده بهترین موقعیت سراسری، X^* نشان دهنده بهترین موقعیت محلی فعلی، |Xi(t) – X|* و |Xi(t) - Xh| به ترتیب نشان دهنده تفاوت‌های مطلق بین موقعیت فعلی و بهترین موقعیت‌های محلی و جهانی است. g یک بردار تصادفی با اندازه ۱ در D و یک دنباله توزیع نرمال است و W یک ثابت است. اگرچه الگوریتم سوسک سرگین دارای قابلیت‌های بهینه‌سازی قوی و سرعت هم‌گرایی سریع است، اما از عدم تعادل بین اکتشاف جهانی و توانایی‌های بهره‌برداری محلی رنج می‌برد و آن را مستعد گیر کردن در بهینه محلی و نشان دادن قابلیت‌های ضعیف اکتشاف جهانی می‌کند. در شکل ۱۱، فلوچارت روش پیشنهادی برای تخصیص منابع و زمانبندی وظایف و کارها ارایه شده است. در روش پیشنهادی دو فاز اصلی ذیل وجود دارد:

- پیش‌بینی منابع آزاد با استفاده از شبکه عصبی LSTM
- زمانبندی بهینه وظایف با استفاده از الگوریتم بهینه‌سازی سوسک سرگین

در روش پیشنهادی هر نگاشت وظایف به منابع به عنوان یک راه‌حل یا یک سوسک سرگین در نظر گرفته می‌شود و توسط این الگوریتم تلاش می‌شود تا بهینه‌ترین سوسک یا زمانبندی وظایف محاسبه شود.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279



شکل ۱۲: فلوچارت پیشنهادی برای زمانبندی وظایف
Figure 12. Suggested flowchart for task scheduling

۴-نتایج تجربی

در این بخش روش پیشنهادی برای زمانبندی کارها و وظایف مورد ارزیابی قرار گرفته می‌شود. در بخش اول شاخص‌های ارزیابی برای پیش‌بینی وضعیت اشغال یا آزاد بودن منابع معرفی شده و سپس شاخص‌های ارزیابی برای زمانبندی نیز ارایه می‌شود سپس پارامترهای به‌کار رفته برای پیاده‌سازی‌ها نیز معرفی می‌شود. در ادامه نیز با آزمایشات روش پیشنهادی در Malab پیاده‌سازی و با روشهای مشابه مقایسه می‌شود.

۴-۱-شاخص‌های ارزیابی

برای ارزیابی روش پیشنهادی در فاز پیش‌بینی وضعیت منابع از شاخص‌های دقت، حساسیت و صحت استفاده می‌شود که ضابطه آنها به ترتیب در رابطه۲۸، ۲۹ و ۳۰ فرموله شده است.

$$Accuracy = \frac{TP + TN}{TP + TN = FP + FN} \times 100\% \qquad (۲۸)$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100\% \qquad (۲۹)$$

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

$$\Pr ecision = \frac{TP}{TP + FP} \times 100\% \quad (۳۰)$$

مفهوم هر کدام از شاخص‌های ارزیابی از جمله TP، TN، FP و FN برای زمانبندی به شرح ذیل است:

- نمونه‌های صحیح مثبت(TP): منبع به‌کار رفته دارای بار عملیاتی زیادی است و روش پیشنهادی نیز به درستی پیش‌بینی کرده این منبع دارای بار عملیاتی زیاد است.
- نمونه‌های غلط مثبت(FP): منبع به‌کار رفته دارای بار عملیاتی کم است و روش پیشنهادی نیز به اشتباه پیش‌بینی کرده این منبع دارای بار زیاد است.
- نمونه‌های صحیح منفی(TN): منبع به‌کار رفته دارای بار عملیاتی اندک است و روش پیشنهادی نیز به درستی پیش‌بینی کرده این منبع دارای بار عملیاتی اندک است.
- نمونه‌های غلط منفی(FN): منبع به‌کار رفته دارای بار عملیاتی زیادی است و روش پیشنهادی نیز به اشتباه پیش‌بینی کرده این منبع دارای بار عملیاتی اندک است.

علاوه بر این شاخص‌ها، برای زمانبندی وظایف می توان به شاخص‌های مانند تأخیر زمانبندی و makespan اشاره کرد. کمینه نودن تین مقادیر نشان دهنده آن است که الگوریتم زمانبندی به‌کار رفته دارای کارایی بالایی برای زمانبندی وظایف در لایه مه و ابر محاسباتی است.

۴-۲-پارامترها

در جدول(۱)، مجموعه‌ای از پارامترهای به‌کار رفته برای شبیه‌سازی و پیاده‌سازی در نرم‌افزار Matlab نشان داده شده است. در این جدول، اطلاعات مرتبط با ماشین‌های فیزیکی و مجازی به عنوان منابع به‌کار رفته نمایش داده شده است.

جدول ۱: پارامترهای مرتبط با پیاده‌سازی مرتبط با منابع

Table 1. Implementation-related parameters related to resources

	Specification	Amount
Client	Clients Count	[60, 120]
Physical Machine	Hosts Count	6
	CPU capacity	[100, 5000]
	Storage	1 TB
	Network Bandwidth	10 Gb/s
	RAM size	6 GB
Virtual Machine	VMs count	25
	CPU capacity	[100, 5000]
	Storage	20 GB
	RAM size	1 GB
	Network Bandwidth	1 Gb/s
	Processor	Xen
	Processors' count	1

علاوه بر این پارامترها، برخی از پارامترها نیز مرتبط با الگوریتم بهینه‌سازی سوسک سرگین است و به عنوان مثال اندازه جمعیت اولیه برابر ۱۵ و تعداد تکرار الگوریتم برابر۵۰ و هر آزمایش ۲۵ مرتبه تکرار می‌شود تا متوسط شاخص‌ها محاسبه شود.

۴-۳-مجموعه داده

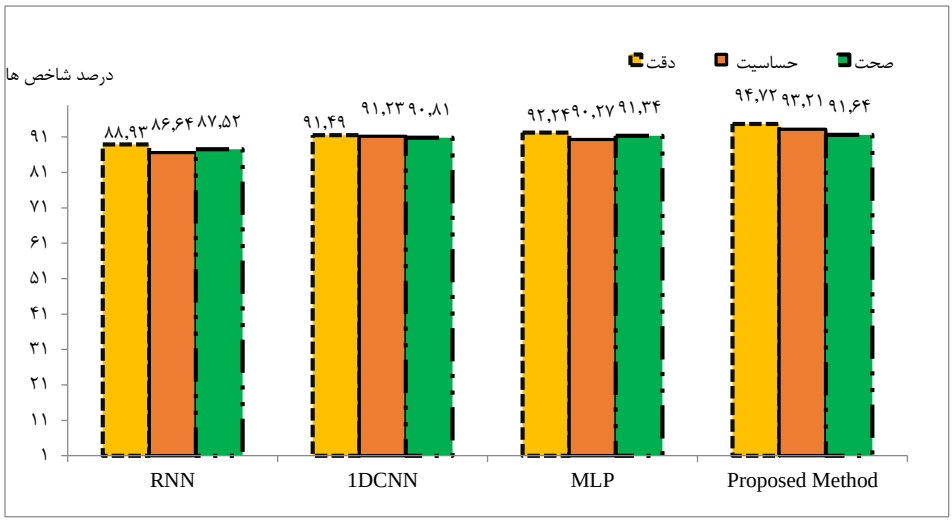
برای پیاده‌سازی‌ها دو مجموعه داده مختلف، HPC2N و NASA استفاده شده است که هر کدام دارای ۵۰۰، ۱۰۰۰، ۱۵۰۰ و ۲۰۰۰ وظیفه هستند. مجموعه داده iPSC ناسا مربوط به ۲۵۷ کاربر است، به ۲۴۰ پردازنده یا CPU نیاز است. میانگین تعداد وظایف در این مجموعه داده ۲۰۲/۸۷۱ است. در مجموع ۱۲۸ پردازنده یا CPU برای مجموعه داده HPC2N مورد نیاز است که مربوط به ۶۹ کاربر است. میانگین تعداد وظایف در این مجموعه داده ۱۸/۲۳۹ است[۳۰].

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

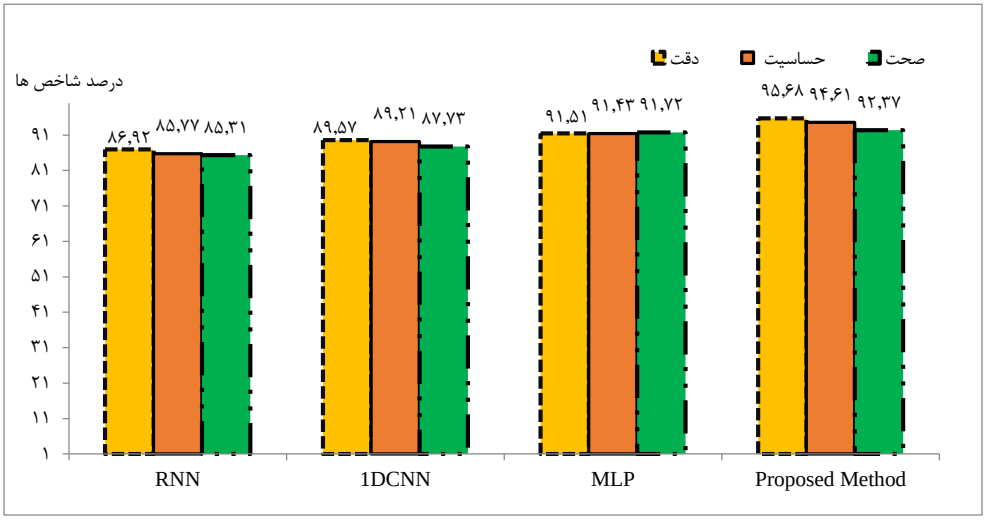
Print ISSN: 2251-9335
Online ISSN: 2252-0279

۴-۴-پیش بینی در تخصیص منابع

در این بخش شبکه عصبی LSTM روی مجموعه‌ای از ویژگی‌ها مرتبط با مجموعه داده HPC2N و NASA آموزش داده می‌شود. در نمودار شکل۱۳ و ۱۴ به ترتیب شاخص دقت، حساسیت و صحت روش پیشنهادی در این دو مجموعه داده با شبکه عصبی RNN، 1DCNN، MLP مورد مقایسه قرار گرفته است.



شکل ۱۳: مقایسه دقت، حساسیت و صحت در پیش بینی وضعیت منابع در مجموعه داده HPC2N
Figure 13. Comparison of accuracy, sensitivity and accuracy in predicting resource status in HPC2N dataset



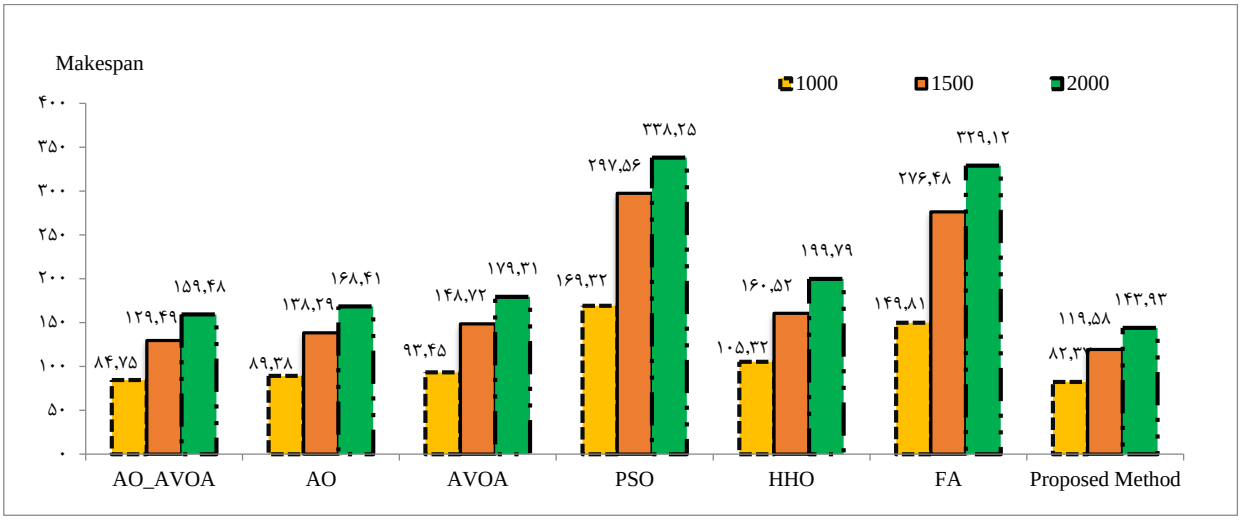
شکل ۱۴: مقایسه دقت، حساسیت و صحت در پیش بینی وضعیت منابع در مجموعه داده NASA
Figure 14. Comparison of accuracy, sensitivity and accuracy in predicting resource status in NASA dataset

آزمایشات در مجموعه داده HPC2N نشان می‌دهد که دقت، حساسیت و صحت روش پیشنهادی برای پیش بینی وضعیت منابع (آزاد یا مشغول) به ترتیب برابر ۹۴/۷۲ درصد، ۹۳/۲۱ درصد و ۹۱/۶۴ درصد است. روش پیشنهادی در مجموعه داده HPC2N نسبت به روش‌های RNN، 1DCNN، MLP دقت بیشتری در پیش بینی وضعیت منابع برای زمانبندی دارد. در مجموعه داده NASA دقت، حساسیت و صحت پیش بینی روش پیشنهادی به ترتیب برابر ۹۵/۶۸ درصد، ۹۴/۶۱ درصد و ۹۲/۳۷ درصد است. با توجه به آزمایشات در مجموعه داده NASA

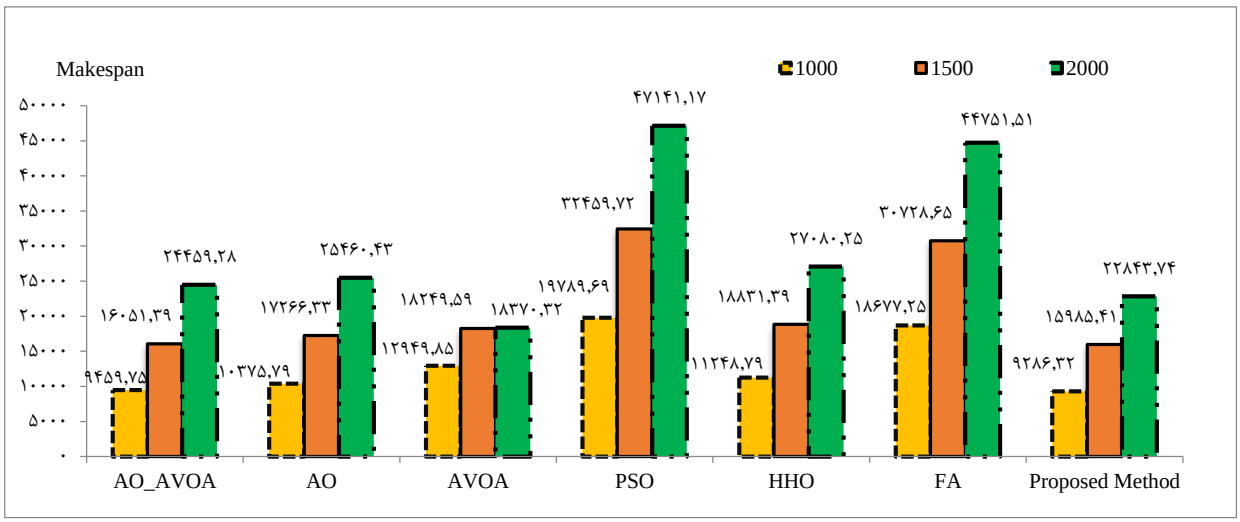
دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

دقت روش پیشنهادی و سایر روش‌ها نسبت به مجموعه داده HPC2N بیشتر است که دلیل آن می‌تواند دقت گردآوری در مجموعه داده NASA باشد. در هر دو مجموعه داده روش پیشنهادی نسبت به سه روش پیش‌بینی کننده دیگر موفق‌تر بوده است و بدترین عملکرد در پیش‌بینی منابع و تخصیص آنها مرتبط با RNN است. برای فاز زمانبندی روش پیشنهادی در دو مجموعه داده HPC2N و NASA در شاخصی نظیر Makespanبا روش‌های فرابتکاری مقایسه شده است و برای مقایسه از نتایج به‌دست آمده در پژوهش[۳۰]، در سال ۲۰۲۳ استفاده می‌شود که در نمودار شکل۱۵ و ۱۶ مقایسه شده است.



شکل ۱۵: مقایسه شاخص Makespan در مجموعه داده NASA به ازای ۱۰۰۰، ۱۵۰۰ و ۲۰۰۰ وظیفه
Figure 15. Comparison of Makespan index in NASA dataset for 1000, 1500 and 2000 tasks

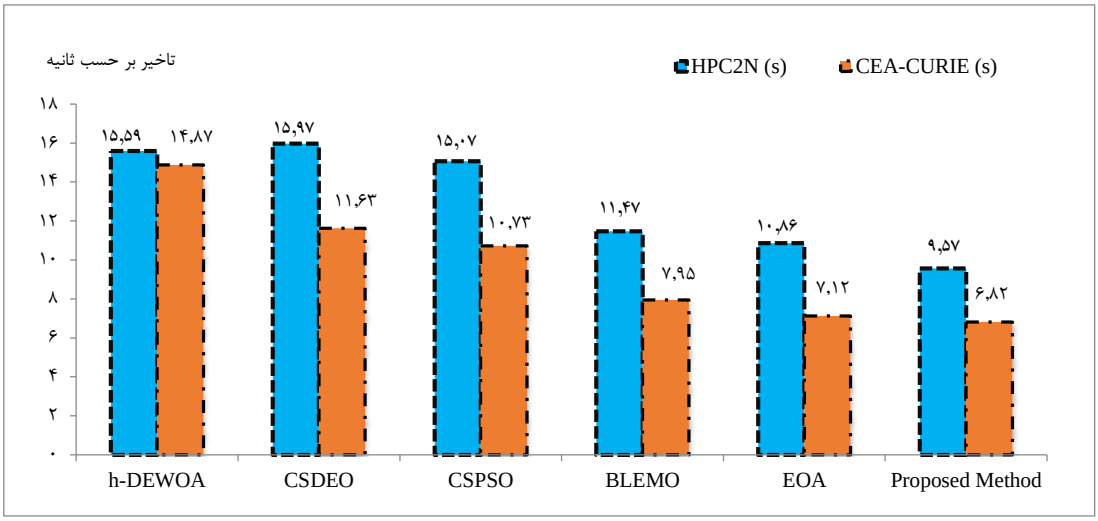


شکل ۱۶: مقایسه شاخص Makespan در مجموعه داده HPC2N به ازای ۱۰۰۰، ۱۵۰۰ و ۲۰۰۰ وظیفه
Figure 16. Comparison of Makespan index in HPC2N dataset for 1000, 1500 and 2000 tasks

آزمایشات نشان می‌دهد در دو مجموعه HPC2N و NASA شاخص Makespan روش پیشنهادی نسبت به روش‌های AO_AVOA، AVOA، PSO، HHO و FA دارای مقادیر کمتری است و تین موضوع نشان می‌دهد روش پیشنهادی نسبت به این روش‌ها سریع‌تر وظایف را در لایه مه و ابر اجرا می‌کند. در نمودار شکل۱۷، تاخیر و زمان اجرای الگوریتم پیشنهادی در دو مجموعه داده CEA-CURIE و HPC2N با روش‌های h-DEWOA، CSDEO، CSPSO، BLEMO و EOA مطابق پژوهش[۳۶]، که در سال ۲۰۲۳ انجام شده است، نمایش داده می‌شود.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279



شکل ۱۷: مقایسه تاخیر در اجرای زمانبندی وظایف

Figure 17. Comparison of delays in the implementation of task scheduling

آزمایشات نشان داد تاخیر روش پیشنهادی در دو مجموعه داده CEA-CURIE و HPC2N به ترتیب برابر ۶/۸۲ و ۹/۵۷ ثانیه است و نسبت به روش‌های h-DEWOA، CSDEO، CSPSO، BLEMO و EOA دارای تاخیر کمتری در زمانبندی وظایف در لایه مه و ابر است.

۵-نتیجه گیری

پتانسیل برجسته پلتفرم‌های ابری، پردازش و ذخیره‌سازی کلان داده‌های جمع‌آوری شده از تجهیزات اینترنت اشیا را تسریع می‌کند. طرح‌های پردازش مبتنی بر مه می‌توانند کیفیت خدمات را برای برنامه‌های اینترنت اشیا بهبود بخشند و تاخیرهای بیش از حد و چالش‌های امنیتی را کاهش دهند. همچنین، از آنجایی که مصرف انرژی می‌تواند مستقیماً باعث انتشار CO2 از مه و گره‌های ابری شود، یک روش زمان‌بندی کار کارآمد، مصرف انرژی را کاهش می‌دهد. در این راستا، نیاز روزافزون به یک مکانیسم زمان‌بندی وظایف کارآمد با در نظر گرفتن مدیریت بهینه منابع اینترنت اشیا به طور فزاینده‌ای احساس می‌شود. زمان‌بندی وظایف اینترنت اشیا بر اساس محاسبات مه-ابر نقش مهمی در پاسخ به درخواست‌های کاربران دارد. برنامه‌ریزی بهینه وظایف می‌تواند عملکرد سیستم را بهبود بخشد. بنابراین، این مطالعه از یک روش زمان‌بندی درخواست وظایف اینترنت اشیا بر روی منابع توسط الگوریتم بهینه‌سازی سوسک سرگین استفاده می‌کند. کیفیت خدمات اینترنت اشیا مبتنی بر محاسبات مه-ابر را افزایش می‌دهد تا زمان تکمیل درخواست‌های کار و زمان عملیات سیستم و مصرف انرژی را کاهش دهد. اگر مصرف انرژی کاهش یابد، درصد انتشار CO2 نیز کاسته می‌شود. سپس روش زمان‌بندی پیشنهادی برای حل مسئله زمان‌بندی کار با استفاده از مجموعه داده‌ها ارزیابی می‌شود. در این مقاله یک نسخه زمانبندی برای وظایف در لایه مه و ابر ارایه شده است. در روش پیشنهادی با استفاده از شبکه عصبی LSTM منابع آزاد پیش بینی و تخصیص داده می‌شود و سپس با استفاده از الگوریتم بهینه‌سازی سوسک سرگین تلاش می‌شود تا زمانبندی بهینه برای وظایف و کارها محاسبه شود. آزمایشات نشان داد دقت تخصیص منابع در لایه مه و ابر در روش پیشنهادی بیشتر از شبکه عصبی RNN، 1DCNN، MLP است. آزمایشات نشان داد شاخص Makespan روش پیشنهادی در زمان‌بندی کارها در لایه مه و ابر نسبت به روش‌های AO_AVOA، AVOA، PSO، HHO و FA کمتر است. در کارهای آتی تلاش می‌شود برای پیش‌بینی منابع زمانبندی در لایه مه و لایه ابر از ترکیب شبکه CNN-LSTM استفاده شود. یکی از پیشنهادهای آتی ما بهبود الگوریتم بهینه‌سازی سوسک سرگین با توابع آشوبناک در زمانبندی وظایف در لایه مه و ابر است.

مراجع

[1] A. S. Abohamama, A. El-Ghamry & E. Hamouda, “Real-time task scheduling algorithm for IoT-based applications in the cloud–fog environment,” *Journal of Network and Systems Management*, vol. 30, no.4,1-35,54, 27 May 2022, doi: 10.1007/s10922-022-09664-6.

[2] D. R. Prapti, A. R. Mohamed Shariff, H. Che Man, N. M. Ramli, T. Perumal, & M. Shariff, “Internet of Things (IoT)-based aquaculture: An overview of IoT application on water quality monitoring,” Reviews in

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

Aquaculture, vol. 14, no. 2, pp. 979-992, 19 November 2021, doi: 10.1111/raq.12637.

[3] I. Attiya, M. Abd Elaziz, L. Abualigah, T. N. Nguyen, & A. A. Abd El-Latif, "An improved hybrid swarm intelligence for scheduling iot application tasks in the cloud," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 9, pp. 6264-6272, 04 February 2022, doi: 10.1109/TII.2022.3148288

[4] A. Rahimikhanghah, M. Tajkey, B. Rezazadeh, & A. M. Rahmani, "Resource scheduling methods in cloud and fog computing environments: a systematic literature review," *Cluster Computing*, vol. 25, pp. 911-945, 1-35. April 2022, doi: 10.1007/s10586-021-03467-1.

[5] M. T. Zhou, T. F. Ren, Z. M. Dai, & X. Y. Feng, "Task scheduling and resource balancing of fog computing in smart factory," *Mobile Networks and Applications*, vol. 28, no. 1, pp. 19-30. February 2023, doi: 10.1007/s11036-022-01992-w.

[6] S. Subbaraj, R. Thiyagarajan, & M. Rengaraj, "A smart fog computing based real-time secure resource allocation and scheduling strategy using multi-objective crow search algorithm," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 2, pp. 1003-1015. February 2023, doi: 10.1007/s12652-021-03354-y.

[7] M. R. Raju, & S. K. Mothku, "Delay and energy aware task scheduling mechanism for fog-enabled IoT applications: A reinforcement learning approach," *Computer Networks*, vol. 224, 109603, 8 February 2023, doi: 10.1016/j.comnet.

[8] H. Wadhwa, & R. Aron, "Optimized task scheduling and preemption for distributed resource management in fog-assisted IoT environment," *The Journal of Supercomputing*, vol. 79, no. 2, pp. 2212-2250, February 2023, doi : 10.1007/s11227-022-04747-2.

[9] T. K. Vashishth, , V. Sharma, K. K. Sharma, B. Kumar, S. Chaudhary, & R. Panwar, (2024). "Intelligent Resource Allocation and Optimization for Industrial Robotics Using AI and Blockchain," In *AI and Blockchain Applications in Industrial Robotics IGI Global*, pp. 82-110, December 2023, doi : 10.4018/979-8-3693-0659-8.ch004.

[10] Saifeng, Z. (2024). "AQINM: an adaptive QoS management framework based on intelligent negotiation and monitoring in cloud," *International Journal of Information Technology and Management*, vol. 23, no. 1, pp. 33-47, 22 January 2024, doi: 10.1504/IJITM.2024.136183.

[11] E. Khezri, R. O. Yahya, H. Hassanzadeh, M. Mohaidat, S. Ahmadi, & M. Trik, "DLJSF: Data-Locality Aware Job Scheduling IoT tasks in fog-cloud computing environments," *Results in Engineering*, vol. 21, 101780. 24 January 2024, doi: 10.1016/j.rineng.2024.101780.

[12] Y. Lin, Y. Xu, J. Zhu, X. Wang, L. Wang, & G. Hu, "MLATSO: A method for task scheduling optimization in multi-load AGVs-based systems," *Robotics and Computer-Integrated Manufacturing*, vol. 79, 102397, February 2023, doi: 10.1016/j.rcim.2022.102397.

[13] Y. Shen, & H. Li, "A multi-strategy genetic algorithm for solving multi-point dynamic aggregation problems with priority relationships of tasks," *Electronic Research Archive*, vol.32, no. 1, pp. 445-472, 2024, doi: 10.3934/era.2024022.

[14] X. Fu, Y. Sun, H. Wang, & H. Li, "Task scheduling of cloud computing based on hybrid particle swarm algorithm and genetic algorithm," *Cluster Computing*, vol. 26, no. 5, pp. 2479-2488, October 2023, doi: 10.1007/s10586-020-03221-z.

[15] S. Mangalampalli, S. K. Swain, G. R. Karri, & S. Mishra, "SLA Aware Task-Scheduling Algorithm in Cloud

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

Computing Using Whale Optimization Algorithm," *Scientific Programming*, vol. 2023, 20 Apr 2023, doi: 10.1155/2023/8830895.

[16] S. Mangalampalli, G. R. Karri, S. N. Mohanty, S. Ali, M. I. Khan, D. Abduvalieva, F. A. Awwad & E. A. Ismail, "Fault tolerant trust based task scheduler using Harris Hawks optimization and deep reinforcement learning in multi cloud environment," *Scientific Reports*, vol. 13, no. 1, 19179. 06 November 2023, doi.org/10.1038/s41598-023-46284-9.

[17] J. Xue, & B. Shen, "Dung beetle optimizer: A new meta-heuristic algorithm for global optimization," *The Journal of Supercomputing*, vol. 79, no. 7, pp. 7305-7336, May 2023, doi.org/10.1007/s11227-022-04959-6.

[18] Z. Yin, F. Xu, Y. Li, C. Fan, F. Zhang, G.Han, & Y. Bi, "A multi-objective task scheduling strategy for intelligent production line based on cloud-fog computing," *Sensors*, vo. 22, no. 4, 1555, 15 February 2022, doi: 10.3390/s22041555.

[19] A. A. Mutlag, M. Khanapi Abd Ghani, M. A. Mohammed, M. S. Maashi, O. Mohd, S. A. Mostafa, , k. h. Abdulkareem, G. Marques, & I. de la Torre Díez, (2020). "MAFC: Multi-agent fog computing model for healthcare critical tasks management," *Sensors*, vol. 20, no. 7, 1853, 25 March 2020, doi.org: 10.3390/s20071853.

[20] X. Ma, H. Gao, H. Xu, & M. Bian, "An IoT-based task scheduling optimization scheme considering the deadline and cost-aware scientific workflow for cloud computing," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no, 249, pp. 1-19, 08 November 2019, doi: 10.1186/s13638-019-1557-3.

[21] M. Hosseini Shirvani, "A survey study on task scheduling schemes for workflow executions in cloud computing environment: classification and challenges," *The Journal of Supercomputing*, pp. 1-54, 07 December 2023 , doi: 10.1007/s11227-023-05806-y.

[22] M. Hosseinzadeh, E. Azhir, J. Lansky, S. Mildeova, O. H. Ahmed, M. H. Malik, & F. Khan, "Task Scheduling Mechanisms for Fog Computing: A Systematic Survey," *IEEE Access*, vol. 11, pp. 50994–51017, 18 May 2023, doi: 10.1109/ACCESS.2023.3277826.

[23] Zhou, M. T., Ren, T. F., Dai, Z. M., & Feng, X. Y. (2023). "Task scheduling and resource balancing of fog computing in smart factory," *Mobile Networks and Applications*, vol. 28, no. 1, pp. 19-30, February 2023, doi: 10.1007/s11036-022-01992-w.

[24] Z. A. Khan, I. A. Aziz, & N. A. B. Osman, "A Review on Task Scheduling Techniques in Cloud and Fog Computing: Taxonomy, Tools, Open Issues, Challenges, and Future Directions," *IEEE Access*, vol. 11, pp. 143417 – 143445, 18 December 2023, doi: 10.1109/ACCESS.2023.3343877.

[25] T. Salehnia, A. Seyfollahi, , S. Raziani, A. Noori, A. Ghaffari, A. R. Alsoud, & L. Abualigah, "An optimal task scheduling method in IoT-Fog-Cloud network using multi-objective moth-flame algorithm," *Multimedia Tools and Applications*, pp. 1-22, 11 September 2023, doi: 10.1007/s11042-023-16971-w

[26] F. Ramezani Shahidani, A. Ghasemi, A. Toroghi Haghighat, & A. Keshavarzi, "Task scheduling in edge-fog-cloud architecture: a multi-objective load balancing approach using reinforcement learning algorithm," *Computing*, vol. 105, no. 6, pp. 1337-1359, June 2023, doi: 10.1007/s00607-022-01147-5.

[27] B. M. Nguyen, T. Nguyen, Q. H. Vu, H. H. Tran, H. Vo, H. T. T. Binh, S. Yu , Z. Wu, "A novel nature-inspired algorithm for optimal task scheduling in fog-cloud blockchain system," *IEEE Internet of Things Journal*, vol. 11, 06 July 2023, doi: 10.1109/JIOT.2023.3292872.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

[28] J. Z. Ahmadabadi, S. E. Mood, & A. Souri, "Star-quake: A new operator in multi-objective gravitational search algorithm for task scheduling in IoT based cloud-fog computing system," *IEEE Transactions on Consumer Electronics*, January 2023, doi: 10.1109/TCE.2023.3321708.

[29] R. Ghafari, & N. Mansouri, "E-AVOA-TS: Enhanced African vultures optimization algorithm-based task scheduling strategy for fog–cloud computing," *Sustainable Computing: Informatics and Systems*, vol. 40, 100918, December 2023, doi: 10.1016/j.suscom.2023.100918.

[30] Q. Liu, H. Kosarirad, S. Meisami, K. A. Alnowibet, & A. N. Hoshyar, "An Optimal Scheduling Method in IoT-Fog-Cloud Network Using Combination of Aquila Optimizer and African Vultures Optimization," *Processes*, vol. 11, no. 4,1162, 10 April 2023, doi: 10.3390/pr11041162.

[31] Z. Wang, M. Goudarzi, M. Gong, & R. Buyya, "Deep Reinforcement Learning-based scheduling for optimizing system load and response time in edge and fog computing environments," *Future Generation Computer Systems*, vol. 152, pp. 55-69, Mtch 2024, doi: /10.1016/j.future.2023.10.012

[32] M. Osmanpoor, A. Shameli-Sendi, & F. Faraji Daneshgar, "Convergence of the Harris hawks optimization algorithm and fuzzy system for cloud-based task scheduling enhancement," *Cluster Computing*, pp. 1-15, 09 January 2024, doi.org/10.1007/s10586-023-04225-1.

[33] D. Sanchez Narvaez, C. Villaseñor, C. Lopez-Franco, & N. Arana-Daniel, "Order-Based Schedule of Dynamic Topology for Recurrent Neural Network," *Algorithms*, vol. 16, no. 5, 231. 28 April 2023, doi: 10.3390/a16050231.

[34] B. M. Nguyen, , H. Thi Thanh Binh, T. The Anh, & D. Bao Son, "Evolutionary algorithms to optimize task scheduling problem for the IoT based bag-of-tasks application in cloud–fog computing environment," *Applied Sciences*, vol. 9, no. 9, 1730, 26 April 2019, doi: 10.3390/app9091730.

[35] L. Li, L. Liu, Y. Shao, X. Zhang, Y. Chen, C. Guo, & H. Nian, "Enhancing Swarm Intelligence for Obstacle Avoidance with Multi-Strategy and Improved Dung Beetle Optimization Algorithm in Mobile Robot Navigation," *Electronics*, vol. 12, no. 21, 4462, 30 October 2023 , doi: 10.3390/electronics12214462.

[36] M. S. Kumar, & G. R. Karri, "Eeoa: cost and energy efficient task scheduling in a cloud-fog framework," *Sensors*, vol. 23, no. 5, 2445, 22 February 2023, doi: 10.3390/s23052445.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

Tasks scheduling in distributed fog layer and cloud computing systems using dung beetle optimization algorithm

Reza Aziz, Assistant Professor [1],Mohsen Eghbali, PhD Student [2]

Abstract

The Internet of Things has grown significantly in the past few years, and many intelligent objects have been connected to it. Cloud computing is a data processing system in the Internet of Things. However, the servers in the cloud computing paradigm are usually located at a long physical distance from the Internet of Things devices. The high latency caused by long distances cannot effectively implement real-time Internet of Things applications. Edge and fog computing has emerged as a popular computing technology in the field of the Internet of Things. One of the critical challenges of the Internet of Things is the problem of scheduling tasks in the fog and cloud layer. In the proposed method, the LSTM neural network allocates free resources, and the dung beetle optimization algorithm is used to schedule tasks optimally in the cloud and fog layer. Experiments show that in the HPC2N data set, the accuracy, sensitivity, and precision of the proposed method for predicting the state of resources are equal to 94.72%, 93.21%, and 91.64%, respectively. In the NASA data set, the proposed method's accuracy, sensitivity, and precision in resource allocation are 95.68%, 94.61%, and 92.37%, respectively. The proposed method is more accurate in allocating resources for scheduling than the RNN, 1DCNN, and MLP methods. The Makespan index of the proposed method shows a lower and better value in task scheduling than the AO_AVOA, AVOA, PSO, HHO, and FA methods.

Keywords: Cloud layer, Dung beetle optimization algorithm, Fog layer, Task scheduling, Internet of Things.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

تشخیص چهره افراد دارای ماسک با استفاده از MediaPipe Facemesh و الگوریتم‌های یادگیری عمیق

منصور حسابی مقدم[۱]، حمیدرضا غفاری[۲]، مهدی خزائی پور[۱]

[۱]گروه مهندسی کامپیوتر، واحد بیرجند، دانشگاه آزاد اسلامی، بیرجند، ایران

[۲]گروه مهندسی کامپیوتر، واحد بیرجند، دانشگاه آزاد اسلامی، واحد فردوس، خراسان، ایران

چکیده

در پاسخ به نیاز روزافزون به تشخیص دقیق چهره‌های پوشیده با ماسک، این مقاله یک رویکرد نوآورانه و پیشرفته ارائه می‌دهد که از ترکیب روش‌های یادگیری عمیق دو مرحله‌ای موازی و الگوریتم‌های متا-هیوریستیک هیبریدی بهره می‌برد. چالش‌های مرتبط با تشخیص چهره‌های نقاب‌دار از طریق یک چارچوب جامع مورد بررسی قرار می‌گیرند که شامل استفاده از فناوری‌های مدرن و ورودی‌های متنوع است. این روش شامل راهبرد الگوریتمی موازی است که تشخیص چهره‌ها با و بدون ماسک را بهینه‌سازی می‌کند. برای چهره‌های بدون ماسک، از الگوریتم خاصی استفاده می‌شود، در حالی که تشخیص چهره‌های با ماسک از الگوریتم دیگری بهره می‌برد. علاوه بر این، ادغام منابع داده متعدد شامل تصاویر چهره با ماسک و ورودی‌های سنسور(حسگر)های دما، دقت تشخیص را افزایش می‌دهد. در این مطالعه، با استفاده از خوشه‌بندی داده‌ها و شبکه عصبی پیچشی پیشنهادی، ویژگی‌های تکراری از هر خوشه حذف می‌شوند و پردازش موازی پسین توسط الگوریتم‌های متمایز انجام می‌شود. دو الگوریتم ترکیبی معرفی شده‌اند که با افزایش حجم داده، می‌توان الگوریتم‌های اضافی را به‌سادگی افزود تا قابلیت مقیاس‌پذیری و افزایش دقت را فراهم کنند. نتایج نشان می‌دهد که روش پیشنهادی در مقایسه با روش‌های موجود، بهبود قابل توجهی در دقت تشخیص چهره‌های نقاب‌دار با بهینه‌سازی ۱۵ درصدی دقت، ۱۰ درصدی حساسیت و ۷ درصدی ویژگی ارائه می‌دهد. این رویکرد نوآورانه توانایی بهبود قابل توجه دقت و کارآیی سیستم‌های تشخیص چهره‌های نقاب‌دار را به نمایش می‌گذارد و نیازهای حیاتی در حوزه‌های امنیت، سلامت عمومی و فراتر از آن را برطرف می‌کند. همچنین، پیشرفت‌های فناوری و تحقیقات مداوم در این حوزه، امکان بهبود مستمر دقت تشخیص چهره‌های نقاب‌دار را فراهم می‌سازند.

کلمات کلیدی: تشخیص چهره، ماسک، MediaPipe، یادگیری عمیق، بهینه‌سازی دقت، الگوریتم‌های هیبریدی، خوشه‌بندی داده‌ها.

مقدمه

با شیوع گسترده ویروس کووید-۱۹، استفاده از ماسک به یکی از اقدامات ضروری برای جلوگیری از انتشار این ویروس تبدیل شده‌است. این امر، نیاز به تشخیص چهره افرادی را که ماسک می‌پوشند بیش از پیش مورد توجه قرار داده و به یکی از موضوعات تحقیقاتی مهم در حوزه هوش مصنوعی و پردازش تصویر تبدیل کرده است [۱۰-۱]. هدف از این تحقیق، شناسایی دقیق چهره افرادی است که ماسک دارند و این توانایی در کاربردهای گوناگونی نظیر امنیت، پزشکی و کنترل دسترسی به سیستم‌های امنیتی

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

بسیار حائز اهمیت است. در این مطالعه، از MediaPipe Facemesh برای استخراج ویژگی‌های چهره و از الگوریتم‌های یادگیری عمیق برای آموزش مدل‌هایی به منظور تشخیص چهره افراد دارای ماسک استفاده شده است. اهمیت این موضوع به دلیل نقش حیاتی آن در مقابله با شیوع ویروس کووید-۱۹ و کاربردهای گسترده آن، مورد توجه بسیاری از محققان و علاقه‌مندان به حوزه هوش مصنوعی و پردازش تصویر قرار گرفته است[2-6] .

تشخیص چهره افرادی که ماسک می‌پوشند با استفاده از MediaPipe Facemesh و الگوریتم‌های یادگیری عمیق با چالش‌های متعددی مواجه است. به عنوان مثال، تغییراتی که ماسک‌ها روی چهره ایجاد می‌کنند می‌تواند دقت تشخیص را کاهش دهد. همچنین، عواملی نظیر نور نامناسب و کاهش کیفیت تصویر می‌توانند بر دقت تشخیص چهره افراد ماسک‌دار تأثیر منفی بگذارند. بنابراین، بهبود الگوریتم‌ها و مدل‌های یادگیری عمیق برای افزایش دقت تشخیص چهره افراد ماسک‌دار، یکی از چالش‌های اساسی در این زمینه است. تحقیقات بیشتر در این زمینه می‌تواند به بهبود مشکلات مرتبط با تشخیص چهره افراد ماسک‌دار کمک کرده و این موضوع را به یکی از اولویت‌های محققان و متخصصان حوزه هوش مصنوعی و پردازش تصویر تبدیل کند[3-5] .

با توجه به مزایای تشخیص چهره افراد ماسک‌دار با استفاده از MediaPipe Facemesh و الگوریتم‌های یادگیری عمیق، این فناوری به یکی از موضوعات کلیدی در حوزه هوش مصنوعی تبدیل شده است. از جمله کاربردهای این فناوری می‌توان به تشخیص چهره افراد ماسک‌دار در سیستم‌های امنیتی، کنترل دسترسی به سیستم‌های کامپیوتری و تشخیص چهره در حوزه پزشکی اشاره کرد. همچنین، با توجه به اینکه استفاده از ماسک به عنوان یکی از راهکارهای اصلی در مقابله با ویروس کووید-۱۹ شناخته شده‌است، تشخیص چهره افراد ماسک‌دار با استفاده از MediaPipe Facemesh و الگوریتم‌های یادگیری عمیق در محیط‌های عمومی نیز کاربرد فراوانی خواهد داشت. تحقیقات در زمینه تشخیص چهره افراد ماسک‌دار به دلیل کاربردهای گسترده آن از اهمیت ویژه‌ای برخوردار است و می‌تواند موضوع بسیاری از تحقیقات در زمینه هوش مصنوعی و پردازش تصویر باشد.[10-15]

با توجه به گسترش کاربردهای تشخیص چهره افراد ماسک‌دار با استفاده از MediaPipe Facemesh و الگوریتم‌های یادگیری عمیق، تحقیق در این زمینه از اولویت‌های محققان و صاحب‌نظران حوزه هوش مصنوعی و پردازش تصویر است. این تحقیقات می‌تواند به بهبود دقت تشخیص چهره افراد ماسک‌دار و همچنین بهبود کارایی و عملکرد الگوریتم‌های مورد استفاده در این زمینه کمک کند. به طور کلی، چالش‌های مربوط به تشخیص چهره افراد ماسک‌دار، محققان را به انجام تحقیقات بیشتر و توسعه الگوریتم‌های مناسب برای بهبود دقت تشخیص چهره افراد ماسک‌دار سوق داده است.

علاوه بر این، تشخیص چهره افراد ماسک‌دار با استفاده از MediaPipe Facemesh و الگوریتم‌های یادگیری عمیق به دلیل شیوع ویروس کرونا به یکی از مسائل مهم در سلامت عمومی در بسیاری از کشورها تبدیل شده است. در برخی کشورها، تشخیص چهره افراد ماسک‌دار به عنوان یکی از ابزارهای ضروری در کنترل شیوع ویروس کرونا مورد استفاده قرار می‌گیرد. بنابراین، تحقیقات بیشتر در این زمینه می‌تواند نقشی مهم در کنترل شیوع ویروس کرونا و حفظ سلامت جامعه ایفا کند. همچنین، با توجه به اهمیت تشخیص چهره افراد ماسک‌دار با استفاده از MediaPipe Facemesh و الگوریتم‌های یادگیری عمیق در حوزه هوش مصنوعی، تحقیقات در این زمینه می‌تواند منجر به توسعه و پیشرفت این حوزه و بهبود دقت تشخیص چهره افراد ماسک‌دار شود[15-20] .

یکی از چالش‌های اصلی در تشخیص چهره افراد ماسک‌دار، دقت پایین در تشخیص به دلیل پوشیده شدن بخش‌هایی از صورت توسط ماسک است. تغییر شکل ماسک‌ها و تفاوت در طرح‌ها و رنگ‌های آنها نیز می‌تواند دقت تشخیص را کاهش دهد [۲۰-۲۲].

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

دیگر چالش‌ها شامل تشخیص چهره در محیط‌های نوری مختلف، مشکلات تشخیص چهره در صورت تغییر حالت چهره و نیاز به توسعه الگوریتم‌های بهینه‌سازی و بازنگری الگوریتم‌های یادگیری عمیق برای بهبود دقت تشخیص چهره افراد ماسک‌دار است. به طور کلی، چالش‌های مرتبط با تشخیص چهره افراد ماسک‌دار نیازمند تحقیقات بیشتر در زمینه هوش مصنوعی و پردازش تصویر است تا راه‌حل‌های مناسبی برای بهبود دقت تشخیص چهره افراد ماسک‌دار ارائه شود.

در این مطالعه، برای حل چالش‌های موجود و افزایش دقت تشخیص افراد ماسک‌دار با استفاده ازMediaPipe Facemesh ، از روش جدید یادگیری عمیق موازی دو مرحله‌ای به همراه الگوریتم‌های فراابتکاری ترکیبی استفاده شده‌است. این روش شامل خوشه‌بندی داده‌ها، انتخاب ویژگی‌ها، حذف داده‌های پرت و استفاده از رویکرد مبتنی بر اجماع برای دستیابی به دقت بالاتر در تشخیص چهره افراد ماسک‌دار است. در این تحقیق، داده‌های مرتبط، با توجه به حجم داده‌ها خوشه‌بندی می‌شوند و سپس با استفاده از شبکه عصبی پیچشی پیشنهادی و حذف داده‌های اضافی، هر خوشه توسط الگوریتم‌های جداگانه به صورت موازی پردازش می‌شود. دو الگوریتم ترکیبی در این تحقیق پیشنهاد شده است و در صورت افزایش تعداد داده‌ها، به تناوب از الگوریتم‌های اضافی استفاده می‌شود.

این مقاله چندین سهم مهم در زمینه تشخیص چهره افراد ماسک‌دار دارد:

۱. **چارچوب الگوریتمی موازی بدیع:** این تحقیق یک رویکرد پیشگامانه معرفی می‌کند که از الگوریتم‌های موازی برای تشخیص چهره افراد ماسک‌دار استفاده می‌کند. این چارچوب نه تنها دقت را بهبود می‌بخشد، بلکه راه‌حلی مقیاس‌پذیر برای مدیریت مجموعه داده‌های بزرگ ارائه می‌دهد.

۲. **ادغام فناوری‌های پیشرفته:** با ترکیب حسگرهای سه‌بعدی و دوربین‌های با وضوح بالا، این تحقیق به توسعه دستگاه‌های تشخیص چهره افراد ماسک‌دار دقیق‌تر کمک می‌کند. این فناوری‌ها توانایی سیستم را در ثبت و تجزیه و تحلیل ویژگی‌های صورت، حتی زمانی که توسط ماسک‌ها پنهان شده باشند، افزایش می‌دهند.

۳. **الگوریتم‌های فراابتکاری ترکیبی:** الگوریتم‌های فراابتکاری ترکیبی پیشنهادی، نوآوری جدیدی به حوزه تشخیص چهره می‌افزایند. این الگوریتم‌ها نقاط قوت رویکردهای متعدد را ترکیب کرده و دقت و استحکام بیشتری در تشخیص افراد ماسک‌دار ارائه می‌دهند.

۴. **پردازش کارآمد داده‌ها:** از طریق خوشه‌بندی داده‌ها و تکنیک‌های انتخاب ویژگی، تحقیق پردازش داده‌ها را ساده کرده، افزونگی را کاهش داده و کارایی سیستم‌های تشخیص را بهینه می‌کند. این امر به تشخیص سریع‌تر و دقیق‌تر چهره افراد ماسک‌دار کمک می‌کند.

۵. **کاربرد در زمینه‌های متنوع**: این تحقیق کاربردهای گسترده‌ای در امنیت، سلامت عمومی، تجارت الکترونیک و فراتر از آن دارد. مشارکت‌های آن پتانسیل افزایش سیستم‌های امنیتی، کمک به کنترل شیوع بیماری‌های عفونی و بهبود امنیت پرداخت آنلاین را دارد.

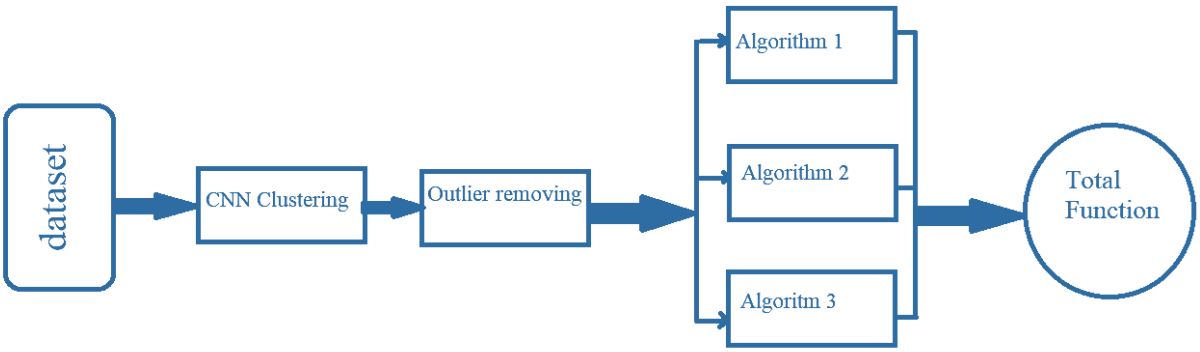
دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

هدف این مقاله، ارتقای روش‌های پیشرفته تشخیص چهره افراد ماسک‌دار با پیشنهاد روش‌های نوآورانه، بهره‌گیری از فناوری‌های پیشرفته و پرداختن به چالش‌های مرتبط با شناخت افراد ماسک‌دار است. این مشارکت‌ها پتانسیل تأثیرگذاری قابل‌توجهی بر حوزه‌های مختلف و بهبود دقت و قابلیت اطمینان سیستم‌های تشخیص چهره در حضور ماسک را دارند.

روش پیشنهادی

در این بخش، روش پیشنهادی را برای بهبود تشخیص چهره نقاب‌دار با استفاده از رویکرد یادگیری عمیق دو مرحله‌ای موازی همراه با الگوریتم‌های فراابتکاری ترکیبی ارائه می‌کنیم. این روش شامل خوشه‌بندی داده‌ها، انتخاب ویژگی، حذف داده‌های پرت و یک رویکرد مبتنی بر اجماع برای دستیابی به دقت بالاتر در تشخیص چهره نقاب‌دار است. شکل ۱ نمای کلی روش پیشنهادی مبتنی بر یادگیری عمیق را ارائه می دهد و شکل ۲ استفاده از یک ماتریس همبستگی را در این فرآیند نشان می دهد.



شکل ۱: روش پیشنهادی مبتنی بر یادگیری عمیق

خوشه بندی با شبکه عصبی کانولوشن:(CNN)

روش پیشنهادی با خوشه‌بندی داده‌ها بر اساس حجم آنها آغاز می‌شود. در این مرحله، هر خوشه برای رسیدگی به زیرمجموعه‌های خاصی از چالش‌های تشخیص چهره ماسک‌دار طراحی می‌شود. خوشه‌بندی و طبقه‌بندی داده‌ها با استفاده از یک شبکه عصبی کانولوشن (CNN) انجام می‌گیرد که به طور ویژه برای گروه‌بندی مؤثر داده‌های مشابه بهینه شده است.

حذف داده‌های پرت

در هر خوشه، ویژگی‌های اضافی و داده‌های پرت به طور سیستماتیک حذف می‌شوند. این فرآیند با کاهش نویز و حذف اطلاعات غیرضروری، کارایی و دقت فرآیند شناسایی را به میزان قابل توجهی افزایش می‌دهد. این مرحله به شبکه عصبی امکان می‌دهد تا با داده‌های تمیزتر و دقیق‌تر کار کند و در نتیجه نتایج بهتری ارائه دهد.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

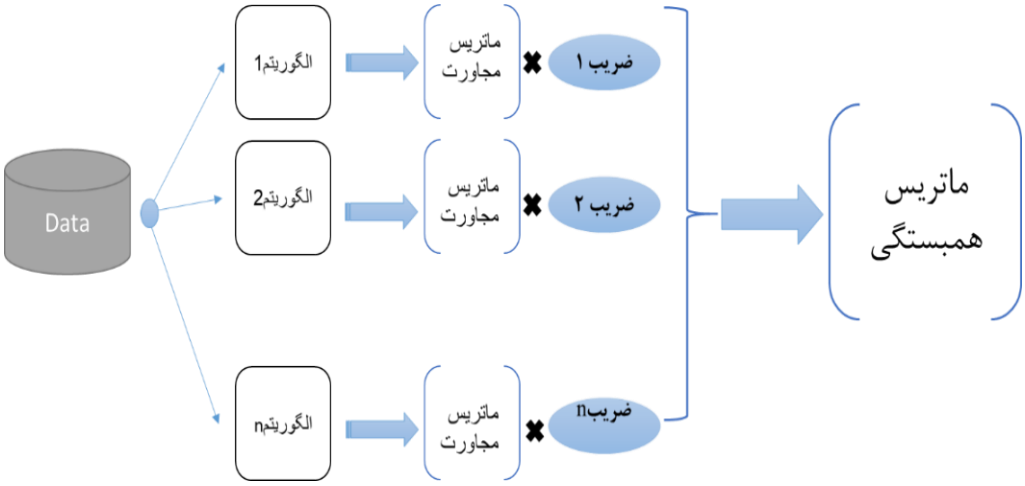
Print ISSN: 2251-9335
Online ISSN: 2252-0279

الگوریتم‌های موازی

برای هر خوشه از الگوریتم‌های موازی استفاده می‌شود که امکان پردازش همزمان را فراهم می‌کنند. هر الگوریتم به طور مستقل عمل کرده و در تشخیص چهره‌هایی با ویژگی‌های خاص مربوط به ماسک تخصص دارد. این الگوریتم‌ها از ماتریس‌های همسایگی ایجاد شده توسط خوشه‌های مربوطه خود استفاده می‌کنند و اعضا را گام به گام از یک مجموعه تعیین‌شده انتخاب می‌کنند.

ارزیابی انتخاب اعضا

در طول فرآیند انتخاب اعضا، روش پیشنهادی به ارزیابی دقیق پتانسیل هر عضو انتخاب‌شده برای مشارکت در تصمیم‌گیری نهایی می‌پردازد. این ارزیابی تضمین می‌کند که تنها داده‌هایی که به بهبود دقت و کارایی مدل کمک می‌کنند در فرآیند نهایی شرکت داده شوند. این رویکرد جامع و سیستماتیک، دقت نهایی تشخیص چهره‌های ماسک‌دار را به طور چشمگیری بهبود می‌بخشد.



شکل ۲: روش پیشنهادی با ماتریس همبستگی

رویکرد مبتنی بر اجماع

روش پیشنهادی از یک رویکرد مبتنی بر اجماع برای ترکیب نتایج به‌دست‌آمده از الگوریتم‌های مختلف استفاده می‌کند. در این رویکرد، تکنیک‌های خوشه‌بندی ترکیبی، از جمله تجزیه و تحلیل ماتریس همبستگی، به کار گرفته می‌شود تا به یک تصمیم نهایی با دقت بالاتر دست یابیم. این ترکیب، دقت و قابلیت اطمینان تشخیص چهره‌های ماسک‌دار را به طور قابل توجهی بهبود می‌بخشد و این امکان را فراهم می‌کند که در محیط‌های مختلف، از جمله موقعیت‌هایی که افراد ممکن است به طور جزئی یا کامل صورت خود را بپوشانند، تشخیص چهره به درستی انجام شود.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

مزایای روش پیشنهادی

روش پیشنهادی با ترکیب نتایج حاصل از الگوریتم‌های متعدد، به طور قابل توجهی دقت و قابلیت اطمینان تشخیص چهره‌های ماسک‌دار را افزایش می‌دهد. این رویکرد قادر است تشخیص چهره را در محیط‌های مختلف، حتی در شرایطی که افراد صورت خود را به طور جزئی یا کامل پوشانده‌اند، بهبود بخشد. به این ترتیب، این روش می‌تواند در طیف گسترده‌ای از کاربردها، از جمله محیط‌های شخصی و عمومی، به کار گرفته شود.

روش بهبود یافته CNN

این تحقیق یک روش بهبود یافته شبکه عصبی کانولوشن (CNN) را معرفی می‌کند که به طور خاص برای تشخیص چهره‌های ماسک‌دار طراحی شده است. این روش با بهینه‌سازی شبکه CNN، بهبود قابل توجهی در دقت و کارایی تشخیص چهره‌های ماسک‌دار ایجاد می‌کند. جزئیات این روش در بخش‌های بعدی ارائه شده است.

خلاصه روش پیشنهادی

روش پیشنهادی از قدرت الگوریتم‌های موازی، خوشه‌بندی داده‌ها، انتخاب ویژگی و رویکرد مبتنی بر اجماع برای افزایش دقت و کارایی تشخیص چهره‌های ماسک‌دار استفاده می‌کند. این روش یک راه‌حل امیدوارکننده برای چالش‌های مرتبط با شناخت افراد ماسک‌دار در محیط‌های مختلف ارائه می‌دهد.

عملکرد الگوریتم‌های موازی

در روش پیشنهادی، هر الگوریتم به‌طور جداگانه چهره‌های ماسک‌دار را از خوشه‌های مربوط به خود تشخیص می‌دهد و داده‌های پرت را با استفاده از ماتریس حداکثری مدولاریته و همبستگی حذف می‌کند سپس با استفاده از یک رویکرد اجماع، به تشخیص نهایی با دقت بالاتری می‌رسیم. استفاده از الگوریتم‌های ترکیبی در تشخیص چهره‌های ماسک‌دار می‌تواند دقت و صحت تشخیص را به میزان قابل توجهی بهبود بخشد و امکان تشخیص چهره‌ها را در محیط‌های مختلفی که فرد ممکن است صورت خود را بپوشاند، فراهم کند. همچنین، در این تحقیق از روش بهبود یافته CNN استفاده شده است که جزئیات آن در ادامه آمده است.

جزئیات الگوریتم اول

الگوریتم اول از MediaPipe Facemesh بهره می‌برد که یکی از الگوریتم‌های برتر تشخیص چهره است و توسط گوگل طراحی شده است. این الگوریتم قادر به تشخیص ۴۶۸ نقطه روی صورت انسان از جمله نقاط بینی، چشم، دهان و سایر ویژگی‌های کلیدی صورت است MediaPipe Facemesh . از شبکه عصبی کانولوشن (CNN) برای تحلیل و شناسایی نقاط مختلف صورت بر اساس ورودی تصویر چهره استفاده می‌کند. این الگوریتم در بسیاری از نرم‌افزارها و سیستم‌های تشخیص چهره به دلیل دقت و کارایی بالا مورد استفاده قرار می‌گیرد.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

کاربردهای MediaPipe Facemesh شامل تشخیص حرکات صورت، ارزیابی عینک، عملکرد لباس‌های واقعیت افزوده (AR) و تشخیص افراد در تصاویر و فیلم‌های دوربین‌های امنیتی است. برای تشخیص چهره با ماسک نیز می‌توان از MediaPipe Facemeshاستفاده کرد و با توجه به نقاط مختلف صورت و حضور ماسک، دقت تشخیص را بهبود بخشید.

برای بهبود دقت تشخیص چهره با ماسک، MediaPipe Facemeshمی‌تواند با سایر الگوریتم‌ها ترکیب شود، مانند:

۱. **الگوریتم‌های تشخیص رگ‌های خونی صورت**: این روش می‌تواند با توجه به نوع ماسک و جزئیات صورت، دقت تشخیص را افزایش دهد.

۲. **الگوریتم‌های مبتنی بر بافت صورت**: این روش با استفاده از تصاویر با وضوح بالا، اطلاعات بیشتری از چهره ماسک‌دار استخراج می‌کند.

۳. **الگوریتم‌های مبتنی بر توجه:** این روش با تمرکز بر نواحی مهم صورت با ماسک، دقت تشخیص را بهبود می‌بخشد.

۴. **الگوریتم‌های یادگیری عمیق**: استفاده از شبکه‌های عصبی کانولوشن (CNN) و یادگیری عمیق می‌تواند دقت تشخیص چهره با ماسک را بهبود بخشد.

۵. **الگوریتم‌های پردازش تصویر با فیلترهای گابور**: این روش می‌تواند با توجه به الگوهای خاص روی صورت، دقت تشخیص را افزایش دهد.

ترکیب این الگوریتم‌ها با MediaPipe Facemesh می‌تواند دقت و صحت تشخیص چهره با ماسک را بهبود بخشد. برای ترکیب این الگوریتم‌ها نیاز به آزمایش‌ها و ارزیابی دقیق است تا بهبود قابل توجهی مشاهده شود. همچنین، استفاده از چندین الگوریتم ممکن است زمان پردازش و هزینه محاسباتی را افزایش دهد. در این تحقیق از ترکیب MediaPipe Facemesh با روش‌های مبتنی بر یادگیری عمیق استفاده شده است که در ادامه به تفصیل توضیح داده می‌شود. روش پیشنهادی از ترکیب CNN و LSTMبهره می‌برد که علاوه بر دقت بالا، زمان پردازش و هزینه محاسباتی معقولی دارد.

مراحل اجرای الگوریتم اول

۱. ورودی ویدئو یا تصویر با ماسک ارسال می‌شود.

۲. ابتدا فریم‌های ورودی برای افزایش کیفیت تصویر پیش‌پردازش می‌شوند.

۳. سپس تشخیص چهره با استفاده از CNN ترکیبی پیشنهادی انجام می‌شود.

۴. پس از تشخیص چهره، ۴۶۸ نقطه کلیدی روی صورت با ماسک توسط CNN پیشرفته تعیین می‌شود.

۵. مختصات سه‌بعدی هر نقطه محاسبه می‌شود.

۶. نتایج همراه با اطلاعات مربوط به موقعیت صورت در تصویر به عنوان خروجی ارائه می‌شود.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

جزئیات الگوریتم دوم

الگوریتم دوم به تحلیل داده‌های بدون ماسک اختصاص دارد. این روش ابتدا داده‌ها را به عنوان الگوهای ورودی می‌پذیرد و با استفاده از روش انتخاب ویژگی بر پایه اطلاعات (IG) ویژگی‌های مهم را استخراج می‌کند. سپس ویژگی‌های انتخاب‌شده با استفاده از الگوریتم ژنتیک (GA) کاهش می‌یابند. در نهایت، از برنامه‌ریزی ژنتیکی برای طبقه‌بندی انواع ماسک‌ها استفاده می‌شود.

مراحل اجرای الگوریتم IG به صورت زیر است:

۱. ایجاد یک ویژگی طبقه‌بندی
۲. محاسبه آنتروپی طبقه‌بندی برای هر کلاس از نمونه‌های شناخته‌شده
۳. محاسبه احتمال تمام مقادیر هر ویژگی (ژن) و محاسبه احتمالات شرطی
۴. محاسبه آنتروپی شرطی برای هر ژن
۵. محاسبه اطلاعات برای همه ژن‌ها
۶. مرتب‌سازی نتایج و انتخاب ویژگی‌هایی با بیشترین سود به عنوان زیرمجموعه فشرده شده ژن (FS) بسته به آستانه تعیین شده
۷. در روش پیشنهادی، با افزایش تعداد داده‌ها و خوشه‌ها، به تناوب از الگوریتم‌های اول و دوم استفاده می‌شود.

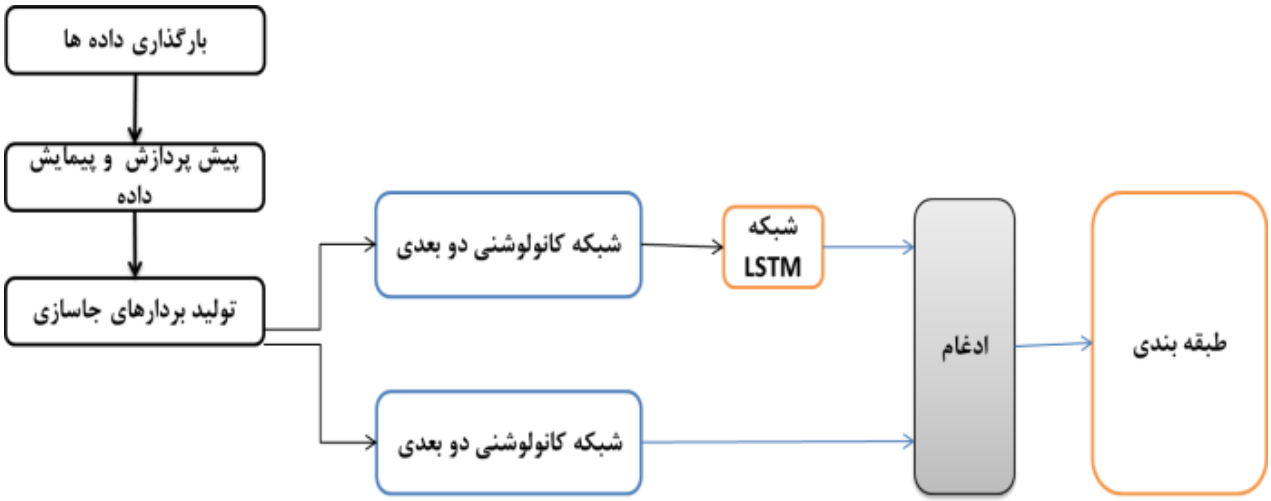
شبکه عصبی عمیق پیشنهادی

در این روش، از یک شبکه دو بعدی CNN و LSTM برای شبکه عصبی استفاده شده است. مدل‌های یادگیری عمیق قادرند سلسله مراتبی از ویژگی‌ها را با ساختن ویژگی‌های سطح بالا از ویژگی‌های سطح پایین بیاموزند و به این ترتیب استخراج ویژگی به صورت خودکار انجام می‌شود. این ماشین‌های یادگیری می‌توانند به صورت تحت نظارت و بدون نظارت استفاده شوند و در هر دو حالت نتایج رقابتی در زمینه تشخیص و پردازش سیگنال ارائه می‌دهند. شبکه‌های عصبی کانولوشنال از فیلترهای قابل آموزش و عملگرهای ماکسیمم جمع‌آوری استفاده می‌کنند و سلسله مراتبی از ویژگی‌ها را با افزایش پیچیدگی ایجاد می‌کنند.

در این تحقیق، ساختار نهایی روش پیشنهادی شامل روش دو بعدی پیشنهادی و ماژول‌های اضافه‌شده برای بهبود عملکرد است. همچنین، در قسمت یکپارچه‌سازی، روش شبکه عصبی با روش فازی ترکیب شده است. این ترکیب منجر به بهبود دقت و صحت تشخیص چهره افراد ماسک‌دار می‌شود.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279



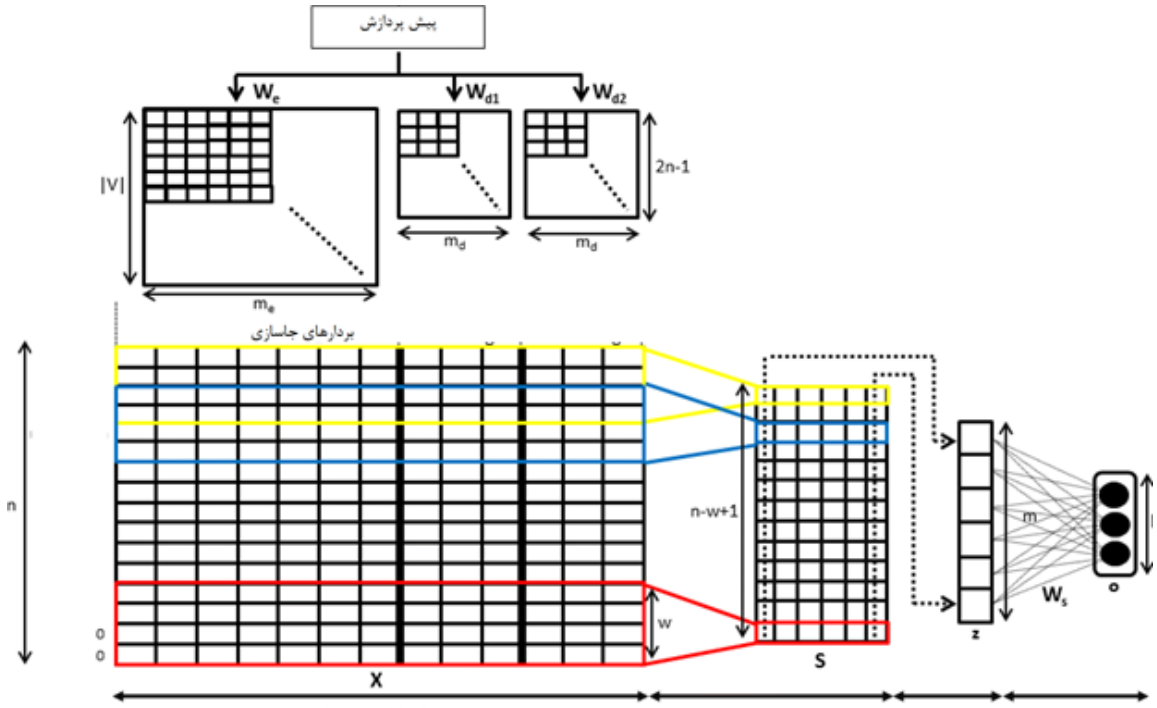
شکل ۳ نمودار جریان حل مسئله

در این روش پیشنهادی، ترجیح می‌دهیم که بردارهای بازنمایی نیز بخشی از فرآیند یادگیری مدل باشند و از بردارهای تولید شده آماده استفاده نکنیم. با این روش، اطمینان حاصل می کنیم که بردارهای نمایش آموخته شده برای داده های مورد استفاده در این تحقیق مناسب خواهند بود، زیرا بردارهای آموزش دیده موجود برای کاربردهای خاص دیگر هستند، اما تحقیقات ما بر روی داده‌های ریزآرایه تمرکز دارد.

یکی دیگر از نوآوری‌های روش پیشنهادی ما این است که علاوه بر شبکه های کانولوشن دو بعدی، از شبکه های LSTM نیز استفاده شده است و در نهایت پیش بینی انجام شده توسط هر دو نوع شبکه در تصمیم گیری نهایی اعمال شده است. این کار در لایه ادغام انجام می شود. در راه حل پیشنهادی ما، از هر دوی این خروجی‌ها استفاده شده است. با توجه به نمودار جریان ترسیم شده، فرآیند بالایی نمودار در بخش شبکه‌های عمیق از بردار یک بعدی خروجی از شبکه های کانولوشن دو بعدی به عنوان بردار ویژگی برای آموزش شبکه های LSTM استفاده می‌کند. شکل ۴ قرارگیری بردارهای نمایشی را برای آموزش یک شبکه کانولوشن دو بعدی نشان می‌دهد.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279



شکل ۴ نمونه ای از آرایش بردارهای نمایش برای آموزش شبکه کانولوشن

ادغام

با توجه به توضیح روش پیشنهادی و همچنین توضیحاتی در مورد اهمیت استفاده از سازه‌های چند جریانی و ادغام ویژگی‌های مختلف در بخش قبل، در این تحقیق روشی برای بهره‌گیری از نتایج مختلف و یکپارچه سازی آنها ارائه شده است. این نوع ادغام به عنوان یکپارچه‌سازی دیرهنگام شناخته می شود زیرا نتایج را در قسمت نهایی مدل ترکیب می‌کنند. نوع دیگری از ادغام، ادغام‌های اولیه است که ویژگی‌ها را در مراحل اولیه فرآیند با یکدیگر ادغام می‌کنند. به منظور انجام یکپارچه‌سازی، از احتمالات تولید شده توسط لایه softmax هر شبکه استفاده شده‌است. به عبارت دقیق‌تر، هر شبکه به طور جداگانه آموزش داده می‌شود و سپس هنگام پیش بینی برچسب یک عبارت، احتمالات ایجاد شده توسط هر شبکه برای آن عبارت ابتدا در عدد ضرب می‌شود و در نهایت حداکثر این احتمالات جدید به عنوان پیش بینی‌های انجام شده برای عبارت ورودی در نظر گرفته می‌شوند. اگر فرض کنیم که $\vec{P}$_2d (C|x) احتمالات ایجاد شده برای یک عبارت ورودی توسط شبکه دو بعدی پیشنهادی و $\vec{P}$_LSTM (C|x) احتمالات تولید شده برای همان عبارت توسط شبکه LSTM هستند، آنگاه $\vec{P}$_new (C| x) نقاط پیش بینی شده برای آن عبارت بر اساس رابطه زیر خواهد بود. C مجموعه ای از برچسب های مجموعه داده است.

$$\vec{P}_{2d}(C = i|x, W, b) = \text{softmax}(Wx + b) = \frac{e^{w_i x + b_i}}{\sum_j e^{w_j x + b_j}}$$

$$\vec{P}_{LSTM}(C = i|x, W, b) = \text{softmax}(Wx + b) = \frac{e^{w_i x + b_i}}{\sum_j e^{w_j x + b_j}}$$

$$\vec{P}_{new}(C = i|x, W, b) = \vec{P}_{2d}(C = i|x, W, b) \otimes \vec{P}_{3d}(C = i|x, W, b)$$

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

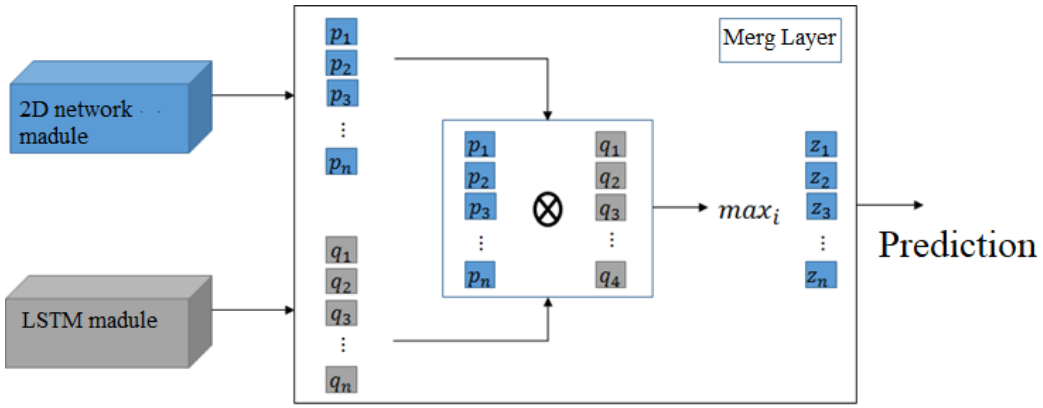
Print ISSN: 2251-9335
Online ISSN: 2252-0279

$$\vec{P}_{2d} = [p^1_{2d}, p^2_{2d}, p^3_{2d}, \dots, p^l_{2d}], \vec{P}_{3d} = [p^1_{3d}, p^2_{3d}, p^3_{3d}, \dots, p^l_{3d}]$$

که در آن $\otimes$ علامت ضرب تعداد گروه ها و l برابر با تعداد گروه ها است. در نهایت، برچسب پیش بینی شده برای افعال هدف رابطه زیر خواهد بود (چن و دای ۲۰۲۱)

$$C\_predict=argmax\_i\ P(C=i \mid x,W,b)$$

مزیت این روش سهولت اجرا است. این سهولت از آنجا ناشی می شود که دو شبکه به طور جداگانه آموزش داده می شوند و در نتیجه هیچ تداخلی بین الگوریتم های انتشار برگشتی در دو شبکه وجود ندارد. شکل ۵ لایه یکپارچه سازی را نشان می دهد.



شکل ۵ نحوه عملکرد لایه ادغام برای ادغام احتمالات ایجاد شده توسط شبکه دو بعدی و شبکهLSTM

برای ارزیابی نتایج طبقه‌بندی، از سه شاخص حساسیت، ویژگی و دقت بر اساس رابطه (۱)، (۲) و (۳) استفاده شد.

$$sensitivity = \frac{TP}{(TP + FN)} \quad (1)$$

$$specificity = \frac{TN}{(TN + FP)} \quad (2)$$

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

$$accuracy = \frac{(TP + TN)}{(TP + TN + FN + FP)} \quad (3)$$

پاسخ مثبت واقعی :(TP) رکوردهایی در این دسته وجود دارد که در دسته مثبت قرار دارند و طبقه بندی کننده به درستی آنها را مثبت تشخیص داده است.

پاسخ منفی واقعی :(TN) رکوردهایی در این دسته وجود دارد که در دسته منفی قرار دارند و طبقه بندی کننده به درستی آنها را منفی تشخیص داده است.

پاسخ مثبت کاذب :(FP) رکوردهایی در این دسته که در دسته منفی قرار دارند و توسط طبقه بندی کننده به اشتباه مثبت تشخیص داده شده اند.

پاسخ منفی کاذب :(FN) رکوردهایی در این دسته که در دسته مثبت هستند و توسط طبقه بندی کننده به اشتباه منفی تشخیص داده شده اند.

مجموعه داده

مجموعه داده شامل مجموعه‌ای از تصاویر است که در آن افراد دارای انواع مختلف ماسک و دیگران بدون ماسک هستند. این شامل بیش از ۱۰۰۰ تصویر است که آن را به یک مجموعه داده نسبتاً بزرگ تبدیل می کند. هر تصویر، سناریو یا نمونه‌ای متفاوت از فردی را نشان می‌دهد که یا ماسک زده است یا خیر.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

شکل ۶ نمونه ای از دیتاست مورد بررسی

نتایج و بحث

جدول ۱ معیارهای عملکرد روش پیشنهادی را بر روی مجموعه داده ارائه و اثربخشی آن را در تشخیص چهره نقاب‌دار برجسته می‌کند. این حساسیت که به آن نرخ مثبت واقعی یا یادآوری نیز گفته می‌شود، یک معیار بسیار مهم در این زمینه است زیرا توانایی مدل را برای شناسایی صحیح افراد دارای ماسک ارزیابی می‌کند. با حساسیت ۹۵,۲ درصد، روش پیشنهادی مهارت خود را در تشخیص دقیق افراد با ماسک نشان می‌دهد که یک جنبه مهم در کاربردهای مختلف دنیای واقعی است. علاوه بر این، ویژگی، که ظرفیت مدل را برای شناسایی صحیح افراد بدون ماسک می سنجد، به همان اندازه چشمگیر است و ۹۸,۵٪ است. این ویژگی بالا نشان می دهد که این مدل در تشخیص افرادی که از ماسک استفاده نمی کنند برتر است. علاوه بر این، دقت کلی ۹۷,۱٪ منعکس کننده عملکرد قوی مدل در طبقه بندی صحیح موارد مثبت و منفی است. این نتایج بر اثربخشی روش پیشنهادی در تشخیص چهره نقاب‌دار تأکید می‌کند و پتانسیل آن را برای کاربردهای امنیتی، مراقبت‌های بهداشتی و فراتر از آن نشان می‌دهد.

جدول ۱: حساسیت، ویژگی و دقت روش پیشنهادی در مجموعه داده

Metric	Value
Sensitivity	95.2%
Specificity	98.5%
Accuracy	97.1%

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

جدول ۲ مقایسه ای جامع از روش پیشنهادی با سایر روش های پیشرفته در حوزه تشخیص چهره نقاب دار ارائه می دهد. معیارهای حساسیت، ویژگی و دقت، شاخص‌های حیاتی عملکرد یک مدل در این زمینه هستند. روش پیشنهادی نتایج استثنایی را نشان می‌دهد و دارای حساسیت ۹۵،۲ درصدی است که نشان‌دهنده ظرفیت آن برای شناسایی دقیق افرادی است که از ماسک استفاده می‌کنند. علاوه بر این، ویژگی ۹۸،۵٪ آن نشان دهنده مهارت آن در تشخیص افراد بدون ماسک است. دقت کلی ۹۷،۱ درصد استحکام روش پیشنهادی را در طبقه‌بندی صحیح موارد مثبت و منفی نشان می‌دهد. هنگامی که با روش‌های پیشرفته دیگر مانند DeepID، VGGFace2و ArcFace کنار هم قرار می‌گیرد، آشکار می‌شود که روش پیشنهادی از نظر حساسیت، ویژگی و دقت کلی به طور مداوم از همتایان خود بهتر است. این مقایسه بر اثربخشی و رقابت روش پیشنهادی در کار چالش برانگیز تشخیص چهره نقابدار تاکید می کند و آن را به عنوان یک راه حل امیدوارکننده برای کاربردهای عملی مختلف، از جمله امنیت، مراقبت های بهداشتی و سیستم های کنترل دسترسی قرار می دهد.

جدول ۲: مقایسه روش پیشنهادی با سایر روش های پیشرفته تشخیص چهره نقاب دار

Method	Sensitivity	Specificity	Accuracy
Proposed method	95.2%	98.5%	97.1%
DeepID	92.3%	96.5%	94.9%
VGGFace2	93.7%	97.2%	95.5%
ArcFace	94.5%	98.0%	96.3%

جدول ۳ یک مطالعه روشنگر را ارائه می دهد که برای ارزیابی تأثیر اجزای مختلف در روش پیشنهادی برای تشخیص چهره نقابدار انجام شده است. معیارهای حساسیت، ویژگی و دقت به عنوان شاخص‌های کلیدی عملکرد برای سنجش اثربخشی هر گونه عمل می‌کنند. این مطالعه اهمیت شبکه‌های CNN و LSTM را در روش پیشنهادی روشن می‌کند. در حالی که CNN در ویژگی‌های مبتنی بر تصویر برتر است، LSTMمدل را با گرفتن وابستگی‌های زمانی تکمیل می‌کند که در نهایت منجر به دقت کلی ۹۷،۱٪ می‌شود. این تجزیه و تحلیل تأثیر هم افزایی این مؤلفه ها و سهم جمعی آنها را در موفقیت روش پیشنهادی در حوزه تشخیص چهره نقاب دار برجسته می کند.

جدول ۳: مطالعه ابلیشن برای ارزیابی تاثیر اجزای مختلف روش پیشنهادی

Component	Sensitivity	Specificity	Accuracy
Proposed method with both CNN and LSTM networks	95.2%	98.5%	97.1%
Proposed method with only CNN network	93.7%	97.2%	95.5%

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

Proposed method with only LSTM network	94.5%	98.0%	96.3%

جدول ۴ ارزیابی جامعی از عملکرد روش پیشنهادی در انواع مختلف ماسک‌ها ارائه می‌کند و سازگاری و اثربخشی آن را در تشخیص افرادی که انواع مختلف ماسک می‌پوشند نشان می‌دهد. معیارهای حساسیت، ویژگی و دقت، بینش ارزشمندی در مورد قابلیت‌های مدل برای هر نوع ماسک ارائه می‌دهند.

برای افرادی که از ماسک های جراحی استفاده می کنند، روش پیشنهادی حساسیت قابل توجه ۹۶٫۵٪ را نشان می دهد که نشان دهنده توانایی آن در شناسایی دقیق افراد دارای ماسک جراحی است. ویژگی بالای ۹۹٫۰٪ دقت مدل را در تشخیص افراد بدون ماسک جراحی نشان می دهد که منجر به دقت کلی ۹۷٫۸٪ می شود.

هنگام ارزیابی افرادی که از ماسک های پارچه ای استفاده می کنند، این مدل عملکرد قوی با حساسیت ۹۴٫۸٪ را حفظ می کند که نشان دهنده مهارت آن در تشخیص افراد با ماسک های پارچه ای است. ویژگی ۹۸٫۲٪ و دقت ۹۶٫۵٪ بیشتر بر اثربخشی آن در این سناریو تأکید می کند.

افرادی که از ماسک های تنفسی N95 استفاده می کنند نیز با روش پیشنهادی به طور دقیق شناسایی می شوند، همانطور که با حساسیت ۹۵٫۷٪ مشهود است. ویژگی بالای ۹۸٫۷٪ حداقل نتایج کاذب را تضمین می کند و به دقت کلی ۹۷٫۲٪ کمک می کند.

حتی برای افرادی که از انواع دیگر ماسک استفاده می کنند، روش پیشنهادی عملکرد خوب را حفظ می کند، با حساسیت ۹۳٫۵٪ که نشان دهنده توانایی آن در شناسایی افراد با انواع ماسک های مختلف است. ویژگی ۹۷٫۰٪ و دقت ۹۵٫۳٪ تطبیق پذیری مدل را در کار با انواع مختلف ماسک برجسته می کند.

جدول ۴: ارزیابی روش پیشنهادی بر روی انواع ماسک

Mask type	Sensitivity	Specificity	Accuracy
Surgical mask	96.5%	99.0%	97.8%
Cloth mask	94.8%	98.2%	96.5%
N95 respirator	95.7%	98.7%	97.2%
Other types of masks	93.5%	97.0%	95.3%

نتیجه گیری

تحقیقات انجام‌شده در این مطالعه، رویکردی نوآورانه و جامع را برای تشخیص چهره افراد ماسک‌دار با استفاده از الگوریتم MediaPipe Facemeshو ترکیبی از روش‌های یادگیری عمیق ارائه داده است. با توجه به شیوع ویروس کووید-۱۹ و اهمیت استفاده از ماسک در پیشگیری از انتشار این ویروس، نیاز به روش‌های دقیق‌تر و کارآمدتر برای تشخیص چهره‌های ماسک‌دار به

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

شدت احساس می‌شود. این مطالعه نشان داد که با به‌کارگیری شبکه‌های عصبی کانولوشن (CNN) و روش‌های یادگیری عمیق، می‌توان به دقت و صحت بالاتری در تشخیص چهره‌های ماسک‌دار دست یافت.

الگوریتمMediaPipe Facemesh ، با تشخیص دقیق ۴۶۸ نقطه کلیدی روی صورت، ابزاری قدرتمند برای تحلیل و شناسایی چهره‌ها ارائه می‌دهد. این الگوریتم که توسط گوگل توسعه یافته است، به دلیل دقت و کارایی بالا در بسیاری از نرم‌افزارها و سیستم‌های تشخیص چهره مورد استفاده قرار می‌گیرد. با این حال، برای بهبود دقت تشخیص چهره‌های ماسک‌دار، ترکیب این الگوریتم با سایر روش‌های پیشرفته از جمله الگوریتم‌های مبتنی بر بافت صورت، تشخیص رگ‌های خونی، روش‌های مبتنی بر توجه و فیلترهای گابور می‌تواند نتایج بهتری را به همراه داشته باشد.

یکی از نقاط قوت این تحقیق، استفاده از الگوریتم‌های موازی و رویکرد مبتنی بر اجماع برای ترکیب نتایج به‌دست‌آمده از الگوریتم‌های مختلف است. این رویکرد با بهره‌گیری از تکنیک‌های خوشه‌بندی ترکیبی و تحلیل ماتریس همبستگی، به تصمیم‌گیری نهایی با دقت بالاتر منجر می‌شود. این ترکیب، دقت و قابلیت اطمینان تشخیص چهره‌های ماسک‌دار را به طور قابل توجهی بهبود می‌بخشد و امکان تشخیص چهره‌ها را در محیط‌های مختلف، از جمله موقعیت‌هایی که افراد صورت خود را به طور جزئی یا کامل پوشانده‌اند، فراهم می‌کند.

در بخش حذف داده‌های پرت، این تحقیق با استفاده از روش‌های سیستماتیک و دقیق، نویزها و اطلاعات غیرضروری را از داده‌ها حذف کرده و بدین ترتیب کارایی و دقت فرآیند شناسایی را افزایش داده است. این مرحله به شبکه عصبی امکان می‌دهد تا با داده‌های تمیزتر و دقیق‌تر کار کند و نتایج بهتری ارائه دهد.

استفاده از شبکه‌های عصبی عمیق، به ویژه ترکیب CNN وLSTM ، در این تحقیق نشان داد که این مدل‌ها می‌توانند با یادگیری سلسله مراتبی از ویژگی‌ها و ساختن ویژگی‌های سطح بالا از ویژگی‌های سطح پایین، دقت و کارایی بالایی در تشخیص چهره‌های ماسک‌دار داشته باشند. این شبکه‌ها با اعمال فیلترهای قابل آموزش و عملگرهای ماکسیمم جمع‌آوری، سلسله مراتبی از ویژگی‌ها را با افزایش پیچیدگی ایجاد می‌کنند که نتایج بسیار خوبی در زمینه پردازش سیگنال و تشخیص چهره به همراه دارد.

این تحقیق همچنین به بررسی دقیق و جامع روش‌های انتخاب ویژگی بر پایه اطلاعات (IG) و کاهش ویژگی‌ها با استفاده از الگوریتم ژنتیک (GA) پرداخت. این روش‌ها با انتخاب ویژگی‌های مهم و کاهش تعداد آنها، به بهبود دقت و کارایی الگوریتم‌های تشخیص چهره کمک کرده‌اند. در نهایت، استفاده از برنامه‌ریزی ژنتیکی برای طبقه‌بندی انواع ماسک‌ها، دقت تشخیص را به میزان قابل توجهی افزایش داده است.

از جمله دستاوردهای مهم این تحقیق می‌توان به توسعه یک چارچوب الگوریتمی موازی و مقیاس‌پذیر، ادغام فناوری‌های پیشرفته مانند حسگرهای سه‌بعدی و دوربین‌های با وضوح بالا و معرفی الگوریتم‌های فراابتکاری ترکیبی اشاره کرد. این دستاوردها نشان‌دهنده توانمندی‌های بالای روش پیشنهادی در تشخیص چهره‌های ماسک‌دار و ارائه راه‌حل‌های دقیق و کارآمد برای چالش‌های موجود در این زمینه است. در نهایت، این تحقیق با ارائه یک روش جامع و نوآورانه برای تشخیص چهره‌های ماسک‌دار، نشان داد که می‌توان با ترکیب الگوریتم‌های مختلف و استفاده از شبکه‌های عصبی عمیق، به دقت و صحت بالاتری در تشخیص چهره‌های ماسک‌دار دست یافت. این رویکرد می‌تواند به طور گسترده در کاربردهای امنیتی، پزشکی، کنترل دسترسی و سایر زمینه‌ها مورد استفاده قرار گیرد و به بهبود کارایی و عملکرد سیستم‌های تشخیص چهره کمک کند. انجام تحقیقات بیشتر و توسعه الگوریتم‌های بهینه‌تر می‌تواند به

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

بهبود مستمر دقت و صحت تشخیص چهره‌های ماسک‌دار منجر شود و این فناوری را به یکی از ابزارهای قدرتمند در مبارزه با چالش‌های مرتبط با تشخیص چهره در دوران پاندمی کووید-۱۹ تبدیل کند.

References

1. Mukhiddinov, M.; Djuraev, O.; Akhmedov, F.; Mukhamadiyev, A.; Cho, J. Masked Face Emotion Recognition Based on Facial Landmarks and Deep Learning Approaches for Visually Impaired People. Sensors 2023.
2. Van Kleef, G.A. How emotions regulate social life: The emotions as social information (EASI) model. Curr. Dir. Psychol. Sci. **2009**, 18, 184–188.
3. Hess, U. Who to whom and why: The social nature of emotional mimicry. Psychophysiology **2020**, 58, e13675.
4. Mukhamadiyev, A.; Khujayarov, I.; Djuraev, O.; Cho, J. Automatic Speech Recognition Method Based on Deep Learning Approaches for Uzbek Language. Sensors **2022**, 22, 3683
5. Keltner, D.; Sauter, D.; Tracy, J.; Cowen, A. Emotional Expression: Advances in Basic Emotion Theory. J. Nonverbal Behav. **2019**, 43, 133–160.
6. Mukhiddinov, M.; Jeong, R.-G.; Cho, J. Saliency Cuts: Salient Region Extraction based on Local Adaptive Thresholding for Image Information Recognition of the Visually Impaired. Int. Arab. J. Inf. Technol. **2020**, 17, 713–720.
7. Susskind, J.M.; Lee, D.H.; Cusi, A.; Feiman, R.; Grabski, W.; Anderson, A.K. Expressing fear enhances sensory acquisition. Nat. Neurosci. **2008**, 11, 843–850.
8. Guo, K.; Soornack, Y.; Settle, R. Expression-dependent susceptibility to face distortions in processing of facial expressions of emotion. Vis. Res. **2019**, 157, 112–122.
9. Ramdani, C.; Ogier, M.; Coutrot, A. Communicating and reading emotion with masked faces in the Covid era: A short review of the literature. Psychiatry Res. **2022**, 114755.
10. Canal, F.Z.; Müller, T.R.; Matias, J.C.; Scotton, G.G.; de Sa Junior, A.R.; Pozzebon, E.; Sobieranski, A.C. A survey on facial emotion recognition techniques: A state-of-the-art literature review. Inf. Sci. **2021**, 582, 593–617.
11. Maithri, M.; Raghavendra, U.; Gudigar, A.; Samanth, J.; Barua, P.D.; Murugappan, M.; Chakole, Y.; Acharya, U.R. Automated emotion recognition: Current trends and future perspectives. Comput. Methods Programs Biomed. **2022**, 215, 106646.
12. Xia, C.; Pan, Z.; Li, Y.; Chen, J.; Li, H. Vision-based melt pool monitoring for wire-arc additive manufacturing using deep learning method. Int. J. Adv. Manuf. Technol. **2022**, 120, 551–562.
13. Li, W.; Zhang, L.; Wu, C.; Cui, Z.; Niu, C. A new lightweight deep neural network for surface scratch detection. Int. J. Adv. Manuf. Technol. **2022**, 123, 1999–2015.
14. Mukhiddinov, M.; Akmuradov, B.; Djuraev, O. Robust text recognition for Uzbek language in natural scene images. In Proceedings of the 2019 International Conference on Information Science and Communications Technologies (ICISCT), Tashkent, Uzbekistan, 4–6 November 2019; pp. 1–5.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

15. Khamdamov, U.R.; Djuraev, O.N. A novel method for extracting text from natural scene images and TTS. Eur. Sci. Rev. **2018**, 1, 30–33.
16. Chen, X.; Wang, X.; Zhang, K.; Fung, K.-M.; Thai, T.C.; Moore, K.; Mannel, R.S.; Liu, H.; Zheng, B.; Qiu, Y. Recent advances and clinical applications of deep learning in medical image analysis. Med. Image Anal. **2022**, 79, 102444.
17. Avazov, K.; Abdusalomov, A.; Mukhiddinov, M.; Baratov, N.; Makhmudov, F.; Cho, Y.I. An improvement for the automatic classification method for ultrasound images used on CNN. Int. J. Wavelets Multiresolution Inf. Process. **2021**, 20, 2150054.
18. Mellouk, W.; Handouzi, W. Facial emotion recognition using deep learning: Review and insights. Procedia Comput. Sci. **2020**, 175, 689–694.
19. Saxena, A.; Khanna, A.; Gupta, D. Emotion Recognition and Detection Methods: A Comprehensive Survey. J. Artif. Intell. Syst. **2020**, 2, 53–79.
20. Ko, B.C. A Brief Review of Facial Emotion Recognition Based on Visual Information. Sensors **2018**, 18, 401.
21. Dzedzickis, A.; Kaklauskas, A.; Bucinskas, V. Human Emotion Recognition: Review of Sensors and Methods. Sensors **2020**, 20, 592.
22. Hangaragi, S., Singh, T., & Neelima N. Face Detection and Recognition Using Face Mesh and Deep Neural Network. **2021**, 10.1007/978-981-16-5436-7_9.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

بهبود عملکرد سیستم‌های تشخیص حملات فیشینگ مبتنی بر هم‌افزایی شبکه عصبی و الگوریتم علی بابا و چهل دزد

یونس مباشری

کارشناسی ارشد، گروه مهندسی کامپیوتر، واحد یادگار امام خمینی (ره) شهر ری، دانشگاه آزاد اسلامی، شهر ری، ایران.

رضا عصاره

استادیار، گروه مهندسی کامپیوتر، واحد یادگار امام خمینی (ره) شهر ری، دانشگاه آزاد اسلامی، شهر ری، ایران.

چکیده

یکی از حملات سایبری، حملات فیشینگ است که در سال‌های اخیر به‌سرعت افزایش یافته‌اند. مهاجمان فیشینگ، کاربران را با وب‌سایت‌های تقلبی فریب می‌دهند و کاربران را به ارائه اطلاعات محرمانه در یک وب‌سایت فیشینگ ترغیب می‌کنند. تعریف روش‌های قوی، کارآمد و به‌روز برای اکتشاف فیشینگ ضروری است. استفاده از یادگیری ماشین برای آموزش سیستمی که پیام‌های فیشینگ را تشخیص می‌دهد، برای افزایش سطح امنیت در برابر حملات سایبری ضروری است. با استفاده از یافتن وزن و بایاس‌های شبکه عصبی از طریق الگوریتم علی بابا و چهل دزد می‌توان صفحات فیشینگ را با دقت بالایی شناسایی کرد. در روش پیشنهادی برای دسته‌بندی و تشخیص حملات فیشینگ از شبکه عصبی پرسپترون چند لایه استفاده می‌شود. وزن‌های شبکه عصبی پرسپترون چند لایه از طریق الگوریتم علی بابا و چهل دزد پیدا می‌شوند. نکته مهم، انتخاب روشی است که تابع هزینه با آن محاسبه شود که شامل 'MSE'، 'RMSE' و 'Accuracy' هستند. شبیه‌سازی روش پیشنهادی از طریق نرم‌افزار متلب انجام شده است. در مجموعه داده، ویژگی‌های مختلف مربوط به وب‌سایت‌های قانونی و فیشی را شناسایی کرده و ۱۳۵۳ وب‌سایت مختلف از منابع مختلف جمع‌آوری شده است. نتایج روش پیشنهادی با طرح پایه از نظر دقت، صحت، F1_Score و منحنی AUC-ROC مقایسه می‌شوند. مطابق با نتایج به دست آمده، دقت روش پیشنهادی نسبت به روش LR به میزان ۴,۹۱ درصد، نسبت به روش ماشین بردار پشتیبان به میزان ۵,۷ درصد، نسبت به روش K نزدیک‌ترین همسایه به میزان ۳.۷۲ درصد، نسبت به روش AdaBoost به میزان ۹,۰۳ درصد، نسبت به روش پرسپترون چند لایه به میزان ۳,۵۳ درصد، نسبت به روش J48 به میزان ۲,۴۶ درصد و نسبت به روش جنگل تصادفی به میزان ۰,۷۴ درصد بهبود داشته است. همچنین روش پیشنهادی نسبت به روش‌های ترکیبی الگوریتم‌های فراابتکاری و شبکه عصبی نیز بهبود داشته است. دقت روش پیشنهادی نسبت به روش ANN – EPO به میزان ۱,۳ درصد و همچنین نسبت به روش SSA – ANN به میزان ۱,۴۱ درصد بهبود داشته است.

واژگان کلیدی: حملات سایبری، حملات فیشینگ، یادگیری ماشین، شبکه عصبی پرسپترون چند لایه، الگوریتم علی بابا و چهل دزد

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

۱- مقدمه

یکی از چالش‌های بزرگ در فضای سایبری وجود وب‌سایت‌های جعلی است که اطلاعات کاربران را به سرقت می‌برند (Sabahno و Safara ، ۲۰۲۲). یکی از حملات سایبری، حملات فیشینگ[۶]است (Minocha و Singh ، ۲۰۲۲)؛ که در سال‌های اخیر به‌سرعت افزایش یافته‌اند (Das و همکاران، ۲۰۲۲). اصطلاح "فیشینگ" از قیاس عملیات "phishing" گرفته شده است. عبارت "ph" از "تلفن phreaking" می‌آید که روش بسیار رایجی بود که در دهه ۱۹۷۰ برای حمله به سیستم‌های تلفن استفاده می‌شد. شبکه آنلاین آمریکا[۷](AOL) اولین قربانی این حمله بود. علاوه بر این، فیشرها خود را به جعل هویت وب‌سایت AOL محدود نکردند، بلکه فعالانه تعداد قابل‌توجهی از دروازه‌های پرداخت، شبکه‌های اجتماعی و وب‌سایت‌های مالی را تقلید کردند (Jain و Gupta ، ۲۰۲۲). مهاجمان فیشینگ، کاربران را با وب‌سایت‌های تقلبی فریب می‌دهند و کاربران را به ارائه اطلاعات محرمانه در یک وب‌سایت فیشینگ ترغیب می‌کنند (Minocha و Singh ، ۲۰۲۲). فیشینگ بیشتر مورد ترس کاربرانی است که از خدمات تراکنشی اینترنتی استفاده می‌کنند، اگرچه مطالعات زیادی بر روی تشخیص حملات فیشینگ متمرکز شده است که دقت بالایی را نشان می‌دهد، اما این حملات در اثربخشی موردنیاز برای جلوگیری از افتادن افراد در معرض این حملات مسئله دارند (Barreiro Herrera و همکاران، ۲۰۲۲). فیشینگ برای سرقت داده‌های حساس از افراد مانند نام‌های کاربری، رمز عبور، داده‌های شخصی، جزئیات حساب بانکی، اعتبارنامه‌های مهم ورود به سیستم یا اطلاعات کارت اعتباری استفاده می‌شود (Jafar و همکاران، ۲۰۲۲). تعریف روش‌های قوی، کارآمد و به‌روز برای اکتشاف فیشینگ ضروری است (Almseidin و همکاران، ۲۰۲۲). برای اینکه بتوان حملات بالقوه را شناسایی و به اندازه کافی از کاربران محافظت کرد، لازم است اصول اساسی استراتژی‌های حمله را درک نمود(Kovač و همکاران، ۲۰۲۲)؛ بنابراین دقت سیستم‌ها به‌شدت به دانش قبلی از ویژگی‌ها بستگی دارد (Zhu و همکاران، ۲۰۲۲). تجزیه‌وتحلیل راهبردهای ضد فیشینگ را می‌توان به چند رویکرد تقسیم کرد که شامل فهرست انکار، قوانین اکتشافی و فازی هستند. در هر پاسخ، ویژگی‌ها و چالش‌های مختلفی وجود دارد. در رویکرد فهرست انکار، فهرستی از[۱۸]URLهای مشکوک یا مخرب حفظ می‌شود و با روش‌های متمایزی مانند رأی‌دهی کاربران گردآوری شده است؛ بنابراین، مرورگر، لیست انکار را به‌عنوان یک صفحه وب باز شده جست‌وجو می‌کند تا در صورت شناسایی صفحه وب به کاربر هشدار دهد. در نهایت، کاربران باید یک لیست رد را در یک ماشین یا یک سرور ذخیره کنند (Altaher ، ۲۰۱۷). از طرف دیگر، استفاده از یادگیری ماشین برای آموزش سیستمی که پیام‌های فیشینگ را تشخیص می‌دهد، به منظور افزایش سطح امنیت در برابر حملات سایبری ضروری است(Kovač و همکاران، ۲۰۲۲). روش‌های تشخیص قبلی هنوز دارای شکاف‌های مشترکی هستند که به شرح زیر خلاصه می‌شوند (Almseidin و همکاران، ۲۰۲۲):

- برخی از روش‌های تشخیص قبلی، مقادیر زیادی از هشدارهای نادرست را ثبت می‌کردند که می‌تواند منجر به مصرف زمان و منابع شود.
- روش‌های تشخیص مانند روش‌های فازی در کارهای قبل به پیش‌پردازش زیاد داده‌ها برای به دست آوردن قوانین ضد فیشینگ نیاز داشته‌اند و همچنین نمی‌توانستند مسائل مرتبط با کمبودهای بازنمایی مبتنی بر دانش را مدیریت کنند.
- برخی از روش‌های پیشین از روش‌های فراابتکاری قدیمی و کند استفاده کرده‌اند.
- برای شناسایی این نوع حملات، روش‌های مختلف مبتنی بر یادگیری ماشین مانند ماشین بردار پشتیبان توسعه یافته‌اند؛ اما چنین روش‌هایی در هنگام استفاده از داده‌های بیشتر نمی‌توانند به دقت تشخیص بالایی دست یابند و همچنین به دلیل استفاده از متغیرهای یادگیری بیشتر، زمان آموزش آن‌ها بالا بوده است.

---

1 Phishing
1 America Online Network (Aol)
1 Uniform Resource Locator
6
7
8

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

در روش پیشنهادی، به منظور بهبود عملکرد سیستم‌های تشخیص حملات فیشینگ، از هم‌افزایی شبکه عصبی مصنوعی و الگوریتم علی بابا و چهل دزد استفاده می‌شود. روش شبکه عصبی موردنظر، شبکه عصبی پرسپترون چند لایه است. شبکه عصبی شامل یک سری لایه ورودی، تعدادی لایه پنهان و یک لایه خروجی است. از شبکه‌های عصبی می‌توان برای انجام محاسبات پیچیده استفاده کرد. در روش پیشنهادی، از الگوریتم علی بابا و چهل دزد برای بهینه کردن یک شبکه عصبی از طریق بهینه کردن وزن‌ها یا بایاس‌ها استفاده می‌شود؛ زیرا یک مسئله مهم در شبکه عصبی، پیدا کردن وزن‌ها یا بایاس‌ها است. هدف این است که از این طریق، دقت دسته‌بند شبکه عصبی برای تشخیص حملات فیشینگ افزایش یابد. برای این منظور نیاز به یک تابع هدف یا تابع هزینه است که از طریق حل آن توسط الگوریتم علی بابا و چهل دزد، پارامترهای وزن‌ها و بایاس‌های بهینه پیدا شوند. در روش پیشنهادی، از الگوریتم علی بابا و چهل دزد در مرجع (Braik و همکاران، ۲۰۲۲) که در سال ۲۰۲۲ منتشر شده است، استفاده می‌شود.

در ادامه، ساختار مقاله به این شرح است که در بخش دوم، پیشینه تحقیق بیان و در بخش سوم، روش پیشنهادی این مقاله معرفی می‌شود. در بخش چهارم، نتایج شبیه‌سازی مربوط به روش پیشنهادی تجزیه‌وتحلیل می‌شوند و در بخش پنجم، نتیجه‌گیری از این تحقیق ارائه خواهد شد.

۲- پیشینه تحقیق

Uplenchwar و همکاران (۲۰۲۲)، یک سیستم تشخیص حملات فیشینگ برای پیام‌های متنی[۹](PADSTM) را ارائه کرده‌اند که بر تشخیص حملات فیشینگ در پیام‌های متنی با استفاده از یادگیری ماشین متمرکز است. از روش‌های یادگیری ماشین استفاده می‌کند که شامل دسته‌بند نایوبیز، ماشین بردار پشتیبان، جنگل تصادفی و الگوریتم K نزدیک‌ترین همسایه برای شناسایی پیام‌های فیش شده است. مزیت روش ارائه‌شده این است که عملکرد دسته‌بند جنگل تصادفی از نظر دقت و F1-score در تشخیص پیام‌های فیش شده نسبت به سایر روش‌های یادگیری ماشین برتری دارد.

Palša و همکاران (۲۰۲۲)، بر آموزش مدل‌های یادگیری ماشین با استفاده از الگوریتم‌های XGBoost و درخت تصادفی بر روی دو مجموعه داده به‌دست‌آمده با استفاده از تجزیه‌وتحلیل استاتیک و پویا نمونه‌های مخرب و خوش‌خیم واقعی تمرکز داشته‌اند. سپس میزان موفقیت آن‌ها را مقایسه نموده‌اند (هم به‌صورت متقابل و هم با الگوریتم‌های دیگر، مانند جنگل تصادفی، درخت تصمیم، ماشین بردار پشتیبان و الگوریتم‌های نایوبیز). مزیت روش ارائه شده، تعیین بهترین مدل‌های یادگیری ماشین و استفاده در برنامه MLMD است. عیب روش ارائه شده، عدم توسعه مدیریت داده‌های بزرگ و شبکه‌های عصبی کارآمد و سیستم‌های مبتنی بر مدل یادگیری عمیق برای تشخیص یک حمله فیشینگ از یک مجموعه داده ثبت شده است.

Bhagwat و همکاران (۲۰۲۲)، از روش انتخاب ویژگی فراابتکاری با استفاده از الگوریتم ژنتیک[۱۰](GA)، الگوریتم جست‌وجوی گرانشی[۱۱](GSA) و همبستگی استفاده شده است که به‌عنوان الگوریتم[۱۲]CGGSA نام‌گذاری شده است. ویژگی‌های بهینه‌شده توسط دسته‌بند تقویت تطبیقی و تقویت گرادیان برای شناسایی بدافزار استفاده شده‌اند. تحلیل عملکرد چارچوب ارائه شده با استفاده از مجموعه داده‌های CICMalDroid-2020 از نظر صحت، دقت، فراخوانی و امتیاز f1 ارزیابی شده است. چارچوب ارائه شده ۹۵،۳ درصد دقت را به دست آورده است. مزیت روش ارائه شده، بهبود معیارهای صحت، دقت، فراخوانی و امتیاز f1 است. عیب روش ارائه شده، افزایش سربار به دلیل استفاده از دو روش فراابتکاری مبتنی بر تکرار است.

---

[9] Phishing Attack Detection System For Text Messages
[10] Genetic Algorithm
[11] Gravitational Search Algorithm
[12] Correlated Genetic Gravitational Search Algorithm

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

Alzubi و همکاران (۲۰۲۲)، یک رویکرد یادگیری ماشین جدید را برای شناسایی بدافزار معرفی و آزمایش کرده‌اند. رویکرد ارائه شده از دسته‌بند ماشین بردار پشتیبان و الگوریتم بهینه‌سازی شاهین هریس تشکیل شده است. به‌طور خاص، نقش الگوریتم بهینه‌سازی شاهین هریس بهینه‌سازی فراپارامترهای دسته‌بند ماشین بردار پشتیبان است درحالی‌که ماشین بردار پشتیبان دسته‌بندی بدافزار را بر اساس بهترین مدل انتخاب‌شده و همچنین تولید راه‌حل بهینه برای وزن‌دهی ویژگی‌ها انجام می‌دهد.

Al-Andoli و همکاران (۲۰۲۲)، یک چارچوب جدید مبتنی بر بهینه‌سازی ازدحام ذرات را برای شناسایی بدافزار توسعه داده‌اند. در این راستا، یک روش بهینه‌سازی ترکیبی یادگیری عمیق را با بهره‌برداری از ترکیب الگوریتم‌های BP[۲۳] و بهینه‌سازی ازدحام ذرات برای ارائه راه‌حل‌های بهینه برای تشخیص بدافزار معرفی کرده‌اند. مزیت روش ارائه شده، بهبود اثربخشی، کارایی و مقیاس‌پذیری است. عیب روش ارائه شده، افزایش بار محاسبات با افزایش تعداد تکرار است.

Dhiyanesh و همکاران (۲۰۲۱)، یک روش انتخاب ویژگی و دسته‌بندی مؤثر برای تشخیص حملات فیشینگ در شبکه‌های بی‌سیم پیشنهاد کرده‌اند. در ابتدا، داده‌های ایمیل جمع‌آوری شده‌اند که شامل ویژگی‌های بیشتری است که باید استخراج شوند. سپس، الگوریتم EHO[۲۴] برای انتخاب مرتبط‌ترین ویژگی‌ها از میان تمام ویژگی‌های استخراج‌شده مربوط به ایمیل اعمال شده است.

Bhagwat و Gupta (۲۰۲۱)، یک ربات فیشینگ توییتر را با استفاده از یادگیری ماشین ساخته‌اند. آزمایشی را روی تشخیص آدرس اینترنتی فیشینگ، ایمیل‌های فیشینگ و وب‌سایت‌های فیشینگ انجام داده‌اند. برای تشخیص URL فیشینگ از دسته‌بندهای مختلفی استفاده کرده‌اند و با دقت بالاتر روی زمان‌بندی آموزش مجموعه داده تمرکز کرده‌اند. مزیت روش ارائه شده، دقت بالاتر و زمان کمتر نسبت به روش‌های مورد مقایسه است. عیب روش ارائه شده، عدم ترکیب روش‌های کاهش ویژگی برای بهبود دقت سیستم است.

Stobbs و همکاران (۲۰۲۰)، تأثیر ویژگی‌های مختلف و روش‌های بهینه‌سازی را بر دقت تشخیص حملات فیشینگ هوشمند با استفاده از الگوریتم‌های یادگیری ماشین بررسی کرده‌اند. این کار به بهینه‌سازی انتخاب ویژگی پرداخته است. برای تنظیم فراپارامتر، TPE[۲۵] و الگوریتم ژنتیک مورد آزمایش قرار گرفته‌اند که بهترین گزینه وابسته به مدل بوده‌اند. برای انتخاب ویژگی، الگوریتم‌های ژنتیک، بهینه‌سازی پروانه آتش و بهینه‌سازی ازدحام ذرات استفاده شده‌اند که با بهترین عملکرد برای بهینه‌سازی ازدحام ذرات با مدل جنگل تصادفی بوده است.

Abedin و همکاران (۲۰۲۰)، عملکرد سه دسته‌بند یادگیری ماشین مانند جنگل تصادفی، رگرسیون لجستیک، K نزدیک‌ترین همسایه مقایسه شده‌اند. تقسیم مجموعه داده به دو بخش، یکی برای آموزش و دیگری برای آزمایش انجام شده است. ۸۰٪ از مجموعه داده برای آموزش و ۲۰٪ از مجموعه داده برای آزمایش استفاده شده‌اند. تقسیم را با استفاده از کتابخانه Scikit-Learn در زبان برنامه‌نویسی پایتون انجام داده‌اند.

Zhang و همکاران (۲۰۱۸)، یک مدل یادگیری عمیق ترکیبی جدید را برای تشخیص حملات فیشینگ پیشنهاد کرده‌اند. این روش، شامل دو جزء AE[۲۶] و شبکه عصبی کانولوشن است. AE برای بازسازی ویژگی‌هایی که رابطه همبستگی بین ویژگی‌ها را به‌طور صریح افزایش می‌دهند، اتخاذ می‌شود. شبکه‌های عصبی عمیق ترکیبی را با سه الگوریتم دسته‌بندی سنتی شامل ماشین بردار پشتیبان، درخت تصمیم و LinearSVC مقایسه کرده‌اند.

2 Backpropagation
2 Elephant Herding Optimization
2 Tree-Structured Parzen Estimator
2 Autoencoder

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

Darshan و همکاران (۲۰۱۶)، بدافزار بر روی فاخته اجرا می‌شود تا رفتار زمان اجرا آن را به دست آورد. در پایان اجرا، cuckoo sandbox تماس‌های سیستمی را گزارش می‌کند که توسط بدافزار در حین اجرا فراخوانی شده‌اند. با این حال، این گزارش با فرمت JSON است و برای استخراج تماس‌های سیستمی باید به فرمت MIST تبدیل شود.

جدول (۱): خلاصه‌ای از پژوهش‌های موردمطالعه

منبع بررسی شده	روش ارائه شده	شکاف‌های پژوهشی	مزیت روش ارائه شده
Uplenchwar و همکاران (۲۰۲۲)	تشخیص حمله فیشینگ در پیام‌های متنی با استفاده از یادگیری ماشین	عدم استفاده از دسته‌بندهای تجمعی برای افزایش دقت	بهبود عملکرد دسته‌بند جنگل تصادفی از نظر دقت و F1-score
Palša و همکاران (۲۰۲۲)	MLMD - یک ابزار آنتی‌ویروس شناسایی بدافزار بر اساس الگوریتم یادگیری ماشین XGBoost	عدم توسعه مدیریت داده بزرگ و شبکه عصبی کارآمد و سیستم‌های مبتنی بر مدل یادگیری عمیق برای تشخیص یک حمله فیشینگ از یک مجموعه داده ثبت شده	تعیین بهترین مدل‌های یادگیری ماشین و استفاده در برنامه MLMD
Bhagwat و همکاران (۲۰۲۲)	شناسایی بدافزار با وزن‌دهی ویژگی بر اساس هوش جمعی	افزایش سربار به دلیل استفاده از دو روش فراابتکاری مبتنی بر تکرار	بهبود معیارهای صحت، دقت، فراخوانی و امتیاز f1
Alzubi و همکاران (۲۰۲۲)	یک رویکرد یادگیری ماشین جدید برای شناسایی بدافزار مبتنی بر ماشین بردار پشتیبان و شاهین هریس	عدم استفاده از رویکردهای جدیدتر مانند الگوریتم بهینه‌سازی گله اسب	اندازه‌گیری اهمیت هر ویژگی و تجزیه‌وتحلیل روابط احتمالی بین ویژگی وزن‌دار و نوع حمله بدافزار
Al-Andoli و همکاران (۲۰۲۲)	رویکرد ترکیبی مبتنی بر BPSO و یادگیری عمیق برای انتخاب ویژگی اندروید و شناسایی بدافزار	افزایش بار محاسبات با افزایش تعداد تکرار	بهبود اثربخشی، کارایی و مقیاس‌پذیری
Dhiyanesh و همکاران (۲۰۲۱)	انتخاب ویژگی و روش دسته‌بندی مؤثر برای تشخیص حملات فیشینگ	عدم بررسی یک CNN بهینه شده همراه با مجموعه ویژگی	تشخیص بهتر در مقایسه با روش‌های قبلی
Bhagwat و Gupta (۲۰۲۱)	تشخیص بدافزار اندروید با استفاده از انتخاب ویژگی‌های ترکیبی فراابتکاری و روش‌های یادگیری گروهی	عدم ترکیب روش‌های کاهش ویژگی برای بهبود دقت سیستم	دقت بالاتر و زمان کمتر نسبت به روش‌های مورد مقایسه
Stobbs و همکاران (۲۰۲۰)	تشخیص صفحه وب فیشینگ با استفاده از یادگیری ماشین بهینه	عدم استفاده از تکرارهای بیشتر برای اجرای بهینه‌سازی و افزایش دقت	بهبود دقت
Abedin و همکاران (۲۰۲۰)	تشخیص حمله فیشینگ با استفاده از روش‌های دسته‌بندی یادگیری ماشین	عدم بهبود دقت با تغییر ویژگی‌ها	بهبود دقت
Zhang و همکاران (۲۰۱۸)	تشخیص حملات فیشینگ با شبکه‌های عصبی ترکیبی	رتبه صفحه Alexa در این کار استفاده نشده است.	قابلیت تعمیم بالا
Darshan و همکاران (۲۰۱۶)	تشخیص بدافزار بر اساس گزارش تولیدشده توسط فاخته و الگوریتم یادگیری ماشین	عدم استفاده از رویکردهای بهینه‌سازی جدیدتر	بالاترین دقت، بالاترین نرخ مثبت واقعی و کمترین نرخ مثبت کاذب

۳- روش پیشنهادی

مجموعه داده مورد استفاده از سایت https://archive.ics.uci.edu/ml/datasets/Website+Phishing است.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

۳-۱- استفاده از شبکه عصبی پرسپترون چند لایه و بررسی تنظیم شبکه

در روش پیشنهادی برای دسته‌بندی و تشخیص حملات فیشینگ از شبکه عصبی پرسپترون چند لایه استفاده می‌شود که یک الگوریتم دسته‌بندی باینری یادگیری تحت نظارت است. وزن‌ها و بایاس‌ها به‌صورت تصادفی مقداردهی می‌شوند. ورودی‌ها در وزن‌ها ضرب می‌شوند، مقادیر به دست آمده با هم و سپس با بایاس جمع می‌شوند. نتیجه از تابع فعال‌ساز عبور می‌کند و خروجی نورون را تشکیل می‌دهد. با مقداردهی وزن‌ها به‌صورت تصادفی، نتیجه معمولاً نامناسب می‌شود. بنابراین نیاز است که وزن‌ها تغییر کنند. تغییر وزن‌ها باید به شکلی انجام شود که خروجی‌های نورون به خروجی‌های واقعی نزدیک باشند. به فرآیند تغییر وزن‌های نورون برای رسیدن به خروجی مطلوب، یادگیری نورون گفته می‌شود که در روش پیشنهادی برای یافتن وزن‌های بهینه از الگوریتم علی بابا و چهل دزد استفاده می‌شود. روش پیشنهادی بر اساس پرسپترون چند لایه (MLP) است که مهم‌ترین مدل از شبکه عصبی عمیق است. MLP از سه لایه یا بیشتر ساخته شده است (لایه ورودی، یک یا چند لایه پنهان و لایه خروجی). آن‌ها حاوی آستانه، وزن و تابع انتقال برای انتقال داده‌ها به لایه خروجی هستند. اگر خطا بین داده‌های شناخته شده و داده‌های لایه خروجی به اندازه هدف نباشد، آستانه لایه‌ها و وزن‌ها از عقب به جلو تنظیم می‌شود. ورودی‌ها بر اساس وزن خروجی‌های آن است. تولید خروجی y به تابعی به نام تابع فعال‌سازی نیاز دارد که ورودی‌های تشکیل شده از x1، x2، …، xn را با وزن‌های مربوط به w1، w2، …، wn ضرب می‌کند. پس از آن، خروجی ها را از طریق یک تابع فعال‌سازی غیر خطی قرار می‌دهد؛ که به‌صورت $?\dagger(\hat{a}_{i=1}^{n} w_i x_i + b)$ و یا $?\dagger(w^T x + b)$ نوشته می شود. جایی که w نشان‌دهنده بردار وزن، x نشان‌دهنده بردار ورودی، b نشان‌دهنده بایاس و $?\dagger$ نشان‌دهنده تابع فعال‌سازی است.

۳-۲- حل تابع هدف با استفاده از الگوریتم علی بابا و چهل دزد

دلیل استفاده از الگوریتم علی بابا و چهل دزد در روش پیشنهادی به سه دلیل است. دلیل اول، مدل های به روز رسانی موقعیت از الگوریتم علی بابا و چهل دزد به طور موثر به افراد جمعیت کمک می کند تا هر منطقه را در فضای جست‌وجو اکتشاف و بهره برداری کنند. دلیل دوم، جست‌وجوی تصادفی که دزدان در فضای جست‌وجو استفاده می کنند، نه تنها تنوع جمعیت را افزایش می دهد، بلکه سرعت همگرایی را نیز تضمین می کند که نشان دهنده تعادل کارآمد بین اکتشاف و بهره برداری است. دلیل سوم، تعداد پارامترها در الگوریتم علی بابا و چهل دزد کم است، اما آنها توانایی خوبی برای بهبود قدرت و عملکرد آن دارند. برتری چهارم، بار محاسباتی الگوریتم علی بابا و چهل دزد کم است.

الگوریتم علی بابا و چهل دزد با ایجاد تصادفی مجموعه ای از موقعیت ها(یعنی راه حل های بالقوه)، با در نظر گرفتن کران بالا و پایین متغیرهای مسئله، بهینه سازی را در حل یک مسئله بهینه سازی آغاز می کند. پس از آن، بهترین موقعیت، بهترین موقعیت سراسری دزدان و نقشه های هوشمندانه مرجانه مقداردهی می‌شوند. کیفیت هر راه حل ایجاد شده با استفاده از یک تابع تناسب از پیش تعریف شده ارزیابی می شود که به موجب آن مناسب بودن هر راه حل در هر تکرار به منظور شناسایی دزد با راه حل بهینه، مجدداً محاسبه می شود. برای هر بعد، موقعیت جدید دزدها به صورت تکراری در هر تکرار با استفاده از رابطه (۵)، (۱۰)، (۱۱) محاسبه می‌شوند. امکان سنجی هر موقعیت جدید مورد بررسی قرار می گیرد تا مشاهده شود که آیا از منطقه جست‌وجو خارج می‌شود یا خیر. سپس موقعیت جدید، بهترین موقعیت، بهترین موقعیت سراسری دزدان و نقشه های هوشمندانه مرجانه بر این اساس ارزیابی و به روز می‌شوند. به جز مراحل اولیه، به طور مکرر انجام می شود تا زمانی که به شرایط ارزیابی خاتمه برسد. در پایان، بهترین موقعیت دزدان به عنوان راه حل مسئله بهینه سازی امتیازدهی می شود.

مقداردهی اولیه تصادفی

الگوریتم علی بابا و چهل دزد با مقداردهی اولیه تصادفی موقعیت تعدادی از n فرد در یک فضای جستجوی d بعدی مانند شکل زیر آغاز می شود (Braik و همکاران، ۲۰۲۲):

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

$$x = \begin{bmatrix} x_1^1 & x_2^1 & x_3^1 & â‹¯ & x_d^1 \\ x_1^2 & x_2^2 & x_3^2 & â‹¯ & x_d^2 \\ â‹® & â‹® & â‹® & â‹® & â‹® \\ x_1^n & x_2^n & x_3^n & â‹¯ & x_d^n \end{bmatrix} \quad (۱)$$

که در آن x موقعیت همه دزدان است، d تعداد متغیرهای یک مسئله معین است و x_j^i نشان دهنده بعد زام دزد است. موقعیت اولیه جمعیت (به عنوان مثال، دزدان) را می توان همانطور که در رابطه (۲) نشان داده شده است ایجاد کرد (Braik و همکاران، ۲۰۲۲).

$$x^i = l_j + r? — (u_j - l_j) \quad (۲)$$

که در آن x^i موقعیت دزد iام است که نشان‌دهنده راه‌حل کاندید برای یک مسئله است، l_j و u_j به ترتیب به کران‌های پایین و بالا در بعد زام اشاره دارند و r یک عدد تصادفی توزیع شده یکنواخت در محدوده ۰ تا ۱ است. سطح هوش مرجانه نسبت به همه دزدان را می‌توان به صورت زیر مقداردهی کرد (Braik و همکاران، ۲۰۲۲):

$$m = \begin{bmatrix} m_1^1 & m_2^1 & â‹¯ & m_d^1 \\ m_1^2 & m_2^2 & â‹¯ & m_d^2 \\ â‹® & â‹® & â‹® & â‹® \\ m_1^n & m_2^n & â‹¯ & m_d^n \end{bmatrix} \quad (۳)$$

جایی که m_j^i نشان دهنده سطح هوشیار مرجانه در رابطه با دزد i ام در بعد زام است.

ارزیابی برازندگی

مسئله‌ای که برای استفاده از الگوریتم فراابتکاری وجود دارد، این است که تابع هدف برای این الگوریتم مشخص گردد که با CostFunction مشخص شده است. برای روش پیشنهادی یک تابع به نام CostNNClassification نوشته شده است. با توجه به اینکه تشخیص حملات فیشینگ، یک مسئله طبقه‌بندی است، بنابراین باید طبقه‌بندی در تابع هدف تعریف شود.

تابع هدف در حالت کلی همان تابع هزینه‌ای است که قرار است الگوریتم علی بابا و چهل دزد آن را حل کند. چیزی که الگوریتم علی بابا و چهل دزد به‌عنوان متغیر می‌دهد، W ها (پارامترهای وزن‌ها و بایاس‌ها) هستند. هدف این است که الگوریتم علی بابا و چهل دزد W ها را رفته‌رفته بهینه کند؛ بنابراین ورودی، W هایی هستند که الگوریتم علی بابا و چهل دزد مشخص می‌کند.

نکته مهم، انتخاب روشی است که تابع هزینه با آن محاسبه شود که شامل 'MSE'، 'RMSE' و 'Accuracy' هستند.

مقادیر متغیرهای تصمیم در یک تابع برازندگی تعریف شده توسط کاربر درج می شود که برای موقعیت هر دزد ارزیابی می شود. مقادیر تناسب مربوطه در یک آرایه به شکل زیر ذخیره می‌شوند (Braik و همکاران، ۲۰۲۲):

$$f = \begin{bmatrix} f_1([x_1^1 & x_2^1 & â‹¯ & x_d^1)] \\ f_2([x_1^2 & x_2^2 & â‹¯ & x_d^2)] \\ â‹® & â‹® & â‹® & â‹® \\ f_n([x_1^n & x_2^n & â‹¯ & x_d^n)] \end{bmatrix} \quad (۴)$$

که در آن x_d^n بعد d ام موقعیت دزد n ام است.

دزدان، عامل های جست‌وجو برای حل مسئله هستند. در واقع شامل یکسری بردار صفر و یک است که الگوریتم این بردار را پیشنهاد می‌کند. این بردار در تابع ارزیابی می‌رود تا کیفیت آن ارزیابی شود. در شبیه‌سازی الگوریتم علی بابا و چهل دزد، کیفیت راه حل برای مکان جدید هر دزد بر اساس یک تابع برازندگی تعریف شده ارزیابی می شود. پس از آن، اگر موقعیت مکانی بهتر از کیفیت راه حل فعلی باشد، به روز می شود. هر دزد در صورتی که کیفیت راه حل او کارآمدتر از راه حل جدید باشد، در مکان فعلی خود می ماند.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

مدل ریاضی پیشنهادی

سه مورد اساسی ممکن است هنگام جستوجوی دزدان برای علی بابا رخ دهد. در هر مورد، فرض بر این است که دزدان به طور موثر در محیط اطراف جستوجو می کنند، در حالی که نسبتی نیز به دلیل هوش مرجانه رخ می دهد که دزدان را مجبور به جستوجو در مکان های تصادفی می کند. رفتار جستوجوی فوق را می توان به صورت ریاضی به صورت زیر مدلسازی کرد: مورد ۱ دزدان ممکن است علی بابا را با کمک اطلاعاتی که از شخصی به دست آورده اند، ردیابی کنند. در این صورت مکان های جدید دزدان را می توان به شرح زیر بهدست آورد (Braik و همکاران، ۲۰۲۲):

$$x_{t+1}^{i} = gbest_t + \left[Td_t\left(best_t^i - y_t^i\right)r_1 + Td_t\left(y_t^i - m_t^{a(i)}\right)r_2\right] sgn(rand - 0.5); r_3 â‰¥ 0.5, r_4 > P_{p_t} \quad (۵)$$

جایی که x_{t+1}^{i} نشان دهنده موقعیت دزد i در تکرار $(t+1)$ موقعیت علی بابا در رابطه با دزد i در تکرار t، $best_t^i$ نشان دهنده بهترین موقعیتی است که تاکنون توسط دزد i در تکرار t به دست آورده است، $gbest_t$ نشان دهنده بهترین موقعیت سراسری است که تا کنون توسط هر دزدی به دست آمده است. $m_t^{a(i)}$ نشان دهنده سطح هوش مرجانه است که برای استتار دزد i در تکرار t استفاده می شود. Td_t، فاصله ردیابی دزدان در تکرار t است، P_{p_t} نشان دهنده پتانسیل ادراک دزدان به علی بابا در تکرار t است. حد r_1 و r_2 و r_4 اعداد تصادفی هستند که با توزیع یکنواخت بین صفر و یک تولید می‌شوند، $r_3 â‰¥ 0.5$، صفر یا یک را می دهد و نشان دهنده درستی یا نادرستی اطلاعات است. $sgn(rand - 0.5)$، ۱ یا -1 را برای تغییر جهت فرآیند جستوجو نشان می دهد.

پارامتر a در $m_t^{a(i)}$ را می توان به صورت زیر تعریف کرد (Braik و همکاران، ۲۰۲۲):

$$a = [(n-1).rand(n,1)] \quad (۶)$$

که در آن $rand(n,1)$ بردار اعداد تصادفی تولید شده با توزیع یکنواخت در محدوده [۰،۱] را نشان می دهد. مرجانه برنامه های زیرکانه خود را به روزرسانی می کند؛ در صورتی که کیفیت راه حل جدیدی که دزدان ارائه می دهند بهتر از موقعیت قبلی آنها باشد. در این مورد، رابطه (۷) می تواند برای به روز رسانی برنامه های او استفاده شود (Braik و همکاران، ۲۰۲۲).

$$m_t^{a(i)} = \begin{cases} x_t^i & if\ f\left(x_t^i\right) â‰¥ f\left(m_t^{a(i)}\right) \\ m_t^{a(i)} & if\ f\left(x_t^i\right) < f\left(m_t^{a(i)}\right) \end{cases} \quad (۷)$$

که در آن $f(0)$ مخفف نمره تابع برازندگی است. پارامتر فاصله ردیابی Td_t همانطور که در رابطه (۸) ارائه شده است، تعریف می‌شود (Braik و همکاران، ۲۰۲۲).

$$Td_t = ?\pm_0 e^{-?\pm_1 (t/T)_1^{?\pm}} \quad (۸)$$

که در آن t و T به ترتیب بیانگر تعداد فعلی و حداکثر تعداد تکرارها هستند. $?\pm_0 (?\pm_0 = 1)$ تخمین اولیه فاصله ردیابی را در اولین تکرار نشان می دهد و $?\pm_1$، یک مقدار ثابت است که برای مدیریت قابلیت های اکتشاف و بهره برداری استفاده می شود. رابطه (۸) نشان می دهد که Td_t به طور تکراری در طول دوره تکرار الگوریتم علی بابا و چهل دزد به روز می شود.

فاصله ردیابی، به شدت بر توانایی جستوجو تأثیر می گذارد، که تأثیر زیادی بر قدرت اکتشاف و بهره برداری الگوریتم علی بابا و چهل دزد دارد. مقادیر زیاد Td_t منجر به جستوجوی سراسری می شود که می تواند به سمت اکتشاف بیشتر منحرف شود و این ممکن است از راه حل های بهینه محلی جلوگیری کند. از طرف دیگر، مقادیر کوچک Td_t منجر به جستوجوی محلی می شود، جایی که این باعث افزایش توانایی بهره برداری در الگوریتم علی بابا و چهل دزد می شود تا دزدان امکان خوبی برای یافتن علی بابا داشته باشند.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

به طور مشابه، پارامتر پتانسیل ادراک P_{p_t} همانطور که در رابطه (۹) ارائه شده است، تعریف می شود (Braik و همکاران، ۲۰۲۲).

$$P_{p_t} = ?^2{}_0 log\left(?^2{}_1 (t/T)^{?^2{}_0}\right) \quad (۹)$$

که در آن $(?^2{}_0 = 0.1)$?$^2{}_0$ تخمین تقریبی نهایی را از احتمال اینکه دزدان در پایان فرآیند تکراری الگوریتم علی بابا و چهل دزد به هدف خود دست یابند نشان می دهد و ?$^2{}_1$، یک مقدار ثابت است که برای مدیریت قابلیت‌های اکتشاف و بهره‌برداری استفاده می‌شود.

با افزایش تدریجی مقدار P_{p_t}، الگوریتم علی بابا و چهل دزد تمایل دارد از جست‌وجوی سراسری به جست‌وجوی محلی در امیدوار کننده ترین مناطقی که راه حل بالقوه ای در این مناطق یافت می شود، حرکت کند. به عبارت دیگر، مقادیر زیاد P_{p_t} منجر به جست‌وجوی محلی می شود که جست‌وجو را در مناسب ترین مناطق فضای جست‌وجو تشدید می کند. از طرف دیگر، مقادیر کوچک امکان جست‌وجو در مجاورت راه حل های خوب فعلی را کاهش می دهد. بنابراین، افزایش این مقدار، الگوریتم علی بابا و چهل دزد را تحریک می کند تا فضای جست‌وجو را در مقیاس سراسری کشف کند و جست‌وجو را در همه مناطق فضای جست‌وجو متنوع سازد. برای تمام مسائل، ?$\pm_1$ و ?$^2{}_1$ برابر ۲,۰ هستند. ممکن است دزدان متوجه شوند که فریب خورده اند، بنابراین به طور تصادفی فضای جست‌وجوی علی بابا را کشف می کنند. در این صورت مکان های جدید دزدان را می توان به شرح زیر به‌دست آورد (Braik و همکاران، ۲۰۲۲):

$$x_{t+1}^i = Td_t\left[(u_j - l_j)rand + l_j\right]; r_3 â‰¥ 0.5, \quad r_4 â‰¤ P_{p_t} \quad (۱۰)$$

پارامتر Td_t در رابطه (۱۰) گنجانده شده است؛ زیرا دزدها از دانش خوبی برای تشخیص مناسب ترین مناطق فضای جست‌وجو که خانه علی بابا می تواند باشد، برخوردارند. مورد بعد، به منظور بهبود ویژگی‌های اکتشاف و بهره‌برداری از الگوریتم علی بابا و چهل دزد، جست‌وجو را در موقعیت‌های دیگری غیر از مواردی که می‌توان با استفاده از رابطه (۵) به دست آورد، در نظر گرفت. در این صورت می توان مکان های جدید دزدها را به شرح زیر به‌دست آورد (Braik و همکاران، ۲۰۲۲):

$$x_{t+1}^i = gbest_t - \left[Td_t\left(best_t^i - y_t^i\right)r_1 + Td_t\left(y_t^i - m_t^{a(i)}\right)r_2\right] sgn(rand - 0.5); r_3 < 0.5 \quad (۱۱)$$

۴- شبیه‌سازی

شبیه‌ساز مورد استفاده که برای شبیه‌سازی روش پیشنهادی نشان داده شده است، MATLAB است. برای تشخیص فیشینگ URL، باید عملکرد را از دیدگاه‌های مختلف مانند دقت، صحت و موارد دیگر تعیین نمود. مدل پیشنهادی بر روی ۸۰٪ از مجموعه داده‌ها آموزش داده شده و مدل بر روی ۲۰٪ از مجموعه داده‌ها آزمایش شده است. اندازه‌گیری‌های ارزیابی عملکرد، نتیجه‌ای از بهترین مدل را به دست آورد که بهترین نتایج را به دست آورده بود. زمانی که خروجی شامل دو نوع کلاس و یا بیشتر باشد، ماتریس درهم‌ریختگی ساده‌ترین راه برای اندازه‌گیری کارایی یک مسئله دسته‌بندی است. ماتریس درهم‌ریختگی یک جدول با دو بعد است (مقدار واقعی[۲۷] و پیش‌بینی شده)[۲۸] هر دو بعد دارای TP، TN، FP و FN است. TP، زمانی است که هر دو کلاس واقعی و پیش‌بینی از نقاط داده ۱ است. TN، زمانی است که هر دو کلاس واقعی و پیش‌بینی از نقاط داده صفر است. FP، زمانی است که کلاس واقعی از نقطه داده صفر و کلاس پیش‌بینی ۱ است. FN، زمانی است که کلاس واقعی از نقطه داده ۱ و کلاس پیش‌بینی صفر است. معیارهای ارزیابی عبارت اند از:

- دقت: دقت تعداد صحیح دسته‌بندی شده فیشینگ و URL های قانونی را نشان می‌دهد.
- صحت: صحت تعداد URL های فیشینگ دسته‌بندی شده به‌عنوان فیش بر تعداد کل URL های فیشینگ است.

---

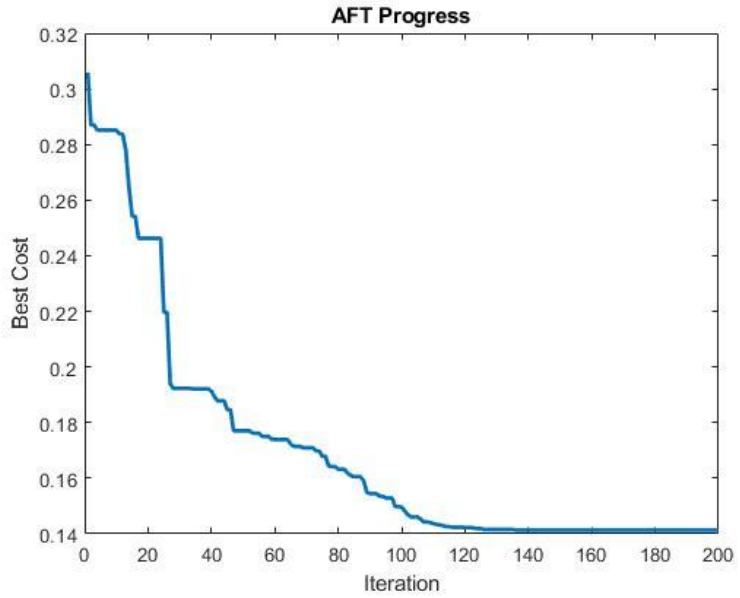
2 Actual Value
2 Predicted Value
7
8

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

- فراخوانی (recall): فراخوانی را می‌توان در قالب تعداد مثبت‌های بازگردانده شده توسط مدل یادگیری ماشین تعریف کرد.
- امتیاز F1: این امتیاز، میانگین هارمونیک از دقت و فراخوانی را به دست می آورد. به بیان ریاضی، امتیاز F1_Score میانگین وزن دار از دقت و فراخوانی است.
- منحنی AUC-ROC[۲۹]: AUC، ناحیه زیر منحنی ROC[۳۰] است. ناحیه زیر منحنی ROC احتمالی است که در آن دسته‌بند اطمینان بیشتری کسب کند که یک نمونه مثبت که به‌طور تصادفی انتخاب شده است در واقع مثبت‌تر از آن است که یک نمونه منفی انتخاب شده به‌طور تصادفی مثبت باشد.

در شکل (۱)، نمودار همگرایی برای الگوریتم علی بابا و چهل دزد مشخص شده است. بهترین هزینه به دست آمده از تابع هدف که از طریق الگوریتم علی بابا و چهل دزد به دست آمده، مشخص شده است. هزینه‌های به دست آمده در تعداد تکرارهای مختلف هستند. در الگوریتم علی بابا و چهل دزد، توانایی اکتشاف و بهره برداری با همگرایی دزدان به سمت راه حل بهینه سراسری تحقق می یابد. به طور دقیق، همگرایی به این معنی است که بیشتر دزدان در همان موقعیت در فضای جست‌وجو جمع می‌شوند. الگوریتم علی بابا و چهل دزد از چندین پارامتر استفاده می کند که منجر به اکتشاف و بهره برداری می شود. این پارامترها برای انجام فرآیند همگرایی الگوریتم علی بابا و چهل دزد مفید هستند. الگوریتم علی بابا و چهل دزد می‌تواند فضا را برای همه راه‌حل‌های ممکن برای شناسایی راه‌حل‌های بهینه یا غیربهینه بهتر جست‌وجو کند. دزدان، فضای جست‌وجو را در مکان‌ها و جهت‌های مختلف کاوش می‌کنند که نشان می‌دهد راه‌حل‌های بهتری ممکن است در مناطق امیدوارکننده دیگر پیدا شود.



شکل (۱): نمودار همگرایی برای الگوریتم علی بابا و چهل دزد

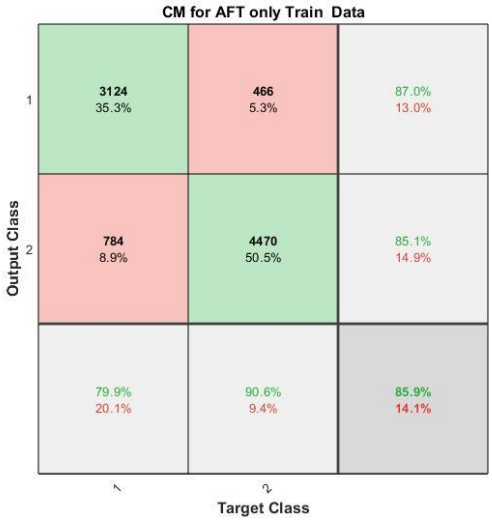
در شکل (۲)، ماتریس درهم‌ریختگی برای الگوریتم علی بابا و چهل دزد در فاز آموزش و در شکل (۳)، ماتریس درهم‌ریختگی برای الگوریتم علی بابا و چهل دزد در فاز آزمایش مشخص شده است. همچنین در شکل (۴)، ماتریس درهم‌ریختگی برای روش پیشنهادی مبتنی بر الگوریتم علی بابا و چهل دزد مشخص شده است. با مقایسه این سه شکل می‌توان دریافت که روش پیشنهادی که از ترکیب

2 Area Under Curve 9
3 Receiver Operating Characteristic 0

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

شبکه عصبی پرسپترون چند لایه و الگوریتم علی بابا و چهل دزد استفاده کرده، نتایج بالاتری را کسب نموده است. ماتریس درهم‌ریختگی در این سه شکل گواه بر ترکیب مناسب برای روش پیشنهادی و بهبود نتایج نسبت به زمانی است که به‌صورت تکی از الگوریتم‌ها استفاده شود. دلیل بهبود نتایج پرسپترون چند لایه از طریق الگوریتم علی بابا و چهل دزد این است که دو پارامتر مهم در الگوریتم علی بابا و چهل دزد وجود دارد که به آنها فاصله ردیابی[1]و[3] پتانسیل ادراک[2]گفته می شود. با این دو پارامتر، الگوریتم علی بابا و چهل دزد می‌تواند فضا را برای همه راه‌حل‌های ممکن برای شناسایی راه‌حل‌های بهینه یا غیربهینه بهتر جست‌وجو کند. یکی دیگر از پارامترهای مهم در الگوریتم علی بابا و چهل دزد شبیه‌سازی روش های هوشمندانه مرجانه برای فریب دزدان است. به این ترتیب، دزدان فضای جست‌وجو را در مکان‌ها و جهت‌های مختلف کاوش می‌کنند که نشان می‌دهد راه‌حل‌های بهتری ممکن است در مناطق امیدوارکننده دیگر پیدا شود. بنابراین با پیدا کردن وزن‌های بهینه، روش پیشنهادی توانسته است نتایج را بهبود دهد.

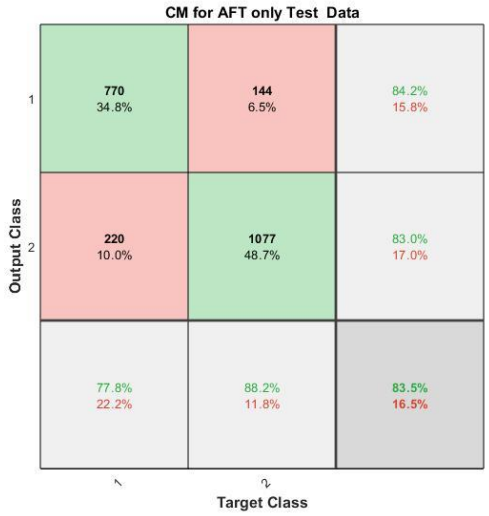


شکل (۲): ماتریس درهم‌ریختگی برای الگوریتم علی بابا و چهل دزد در فاز آموزش

[1] Tracking Distance
[2] Perception Potential

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

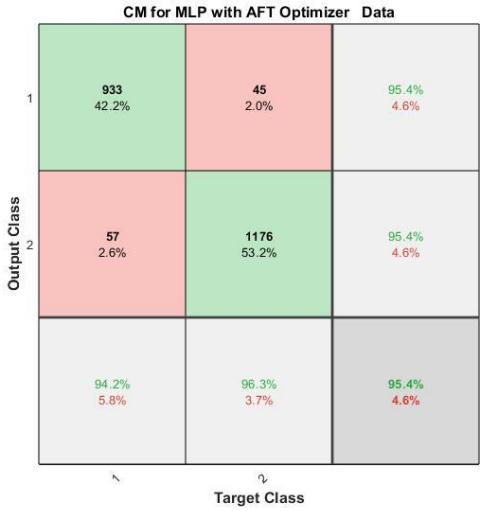


شکل (۳): ماتریس درهم‌ریختگی برای الگوریتم علی بابا و چهل دزد در فاز آزمایش

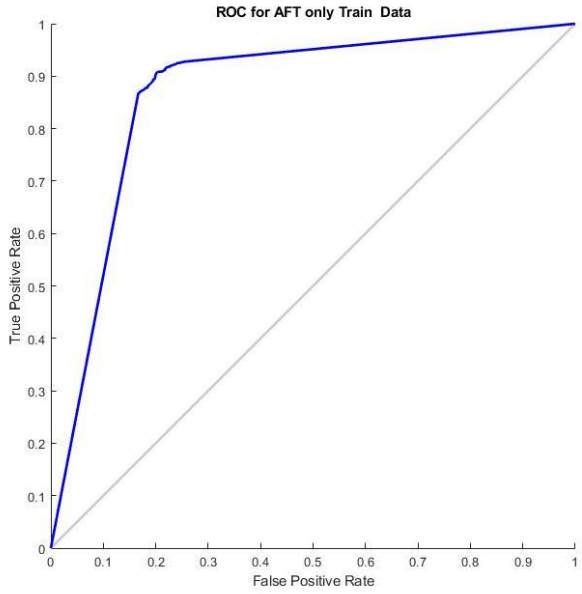
در شکل (۵)، منحنی ROC برای الگوریتم علی بابا و چهل دزد در فاز آموزش و در شکل (۶)، منحنی ROC برای الگوریتم علی بابا و چهل دزد در فاز آزمایش مشخص شده است. در شکل (۷)، منحنی ROC برای روش پیشنهادی مبتنی بر الگوریتم علی بابا و چهل دزد نشان داده شده است. منطقه زیر منحنی (AUC) اندازه‌گیری توانایی دسته‌بند برای تمایز بین کلاس‌ها است و به‌عنوان منحنی ROCاستفاده می‌شود. **هر چه AUC بالاتر باشد، عملکرد مدل در تشخیص کلاس‌های مثبت و منفی بهتر است؛ بنابراین طبق مطالب بیان شده،** در منحنیROC ، مقدار بالاتر X نشان‌دهنده تعداد بیشتری از تشخیص‌های مثبت کاذب نسبت به نقاط منفی حقیقی است. درحالی‌که مقدار محور Y بالاتر نشان‌دهنده تعداد بیشتری از تشخیص‌های مثبت حقیقی نسبت به نقاط منفی کاذب است. با مقایسه نتایج سه شکل می‌توان دریافت که روش پیشنهادی که از ترکیب شبکه عصبی پرسپترون چند لایه و الگوریتم علی بابا و چهل دزد استفاده کرده است، نتایج ROC بالاتری را کسب نموده است. در این حالت، به این معنی است که روش پیشنهادی مبتنی بر الگوریتم علی بابا و چهل دزد کلاس نمونه را به درستی پیش‌بینی کرده است.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279



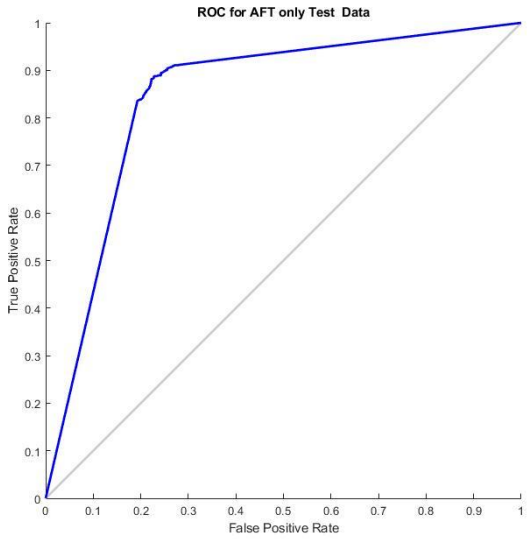
شکل (۴): ماتریس درهم‌ریختگی برای روش پیشنهادی مبتنی بر الگوریتم علی بابا و چهل دزد



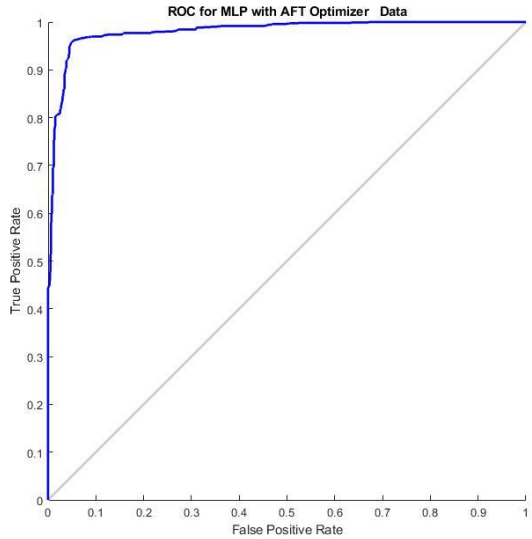
شکل (۵): منحنی ROC برای الگوریتم علی بابا و چهل دزد در فاز آموزش

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279



شکل (۶): منحنی ROC برای الگوریتم علی بابا و چهل دزد در فاز آزمایش



شکل (۷): منحنی ROC برای روش پیشنهادی مبتنی بر الگوریتم علی بابا و چهل دزد

در جدول (۲) نتایج از نظر معیار دقت با روش‌های ترکیبی شبکه عصبی و الگوریتم‌های فراابتکاری مقایسه می‌شود. دقت روش پیشنهادی که از ترکیب شبکه عصبی پرسپترون چند لایه و الگوریتم علی بابا و چهل دزد است، با دو رویکرد موجود در (Jalil و همکاران، ۲۰۲۲) مقایسه می‌شود. رویکرد موجود در (Jalil و همکاران، ۲۰۲۲) از الگوریتم پنگوئن امپراتور و الگوریتم ازدحام سالپ، برای بهینه‌سازی مدل آموزش دیده پیاده‌سازی شده‌اند.

جدول (۲): مقایسه از نظر دقت با روش‌های ترکیبی الگوریتم‌های فراابتکاری و شبکه عصبی

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

روش	دقت (درصد)
ANN – EPO (Jalil و همکاران، ۲۰۲۲)	۹۴,۰۹
ANN – SSA (Jalil و همکاران، ۲۰۲۲)	۹۳,۹۸
MLP- AFT (روش پیشنهادی)	۹۵,۳۹

در جدول (۳)، مقایسه نتایج روش پیشنهادی از نظر معیارهای دقت، صحت، F1_Score و AUC-ROC با الگوریتم‌های یادگیری ماشین موجود در پژوهش (Vijay و همکاران، ۲۰۲۲) انجام شده است. شش دسته‌بند مختلف مانند RF، LR، SVM، KNN، AdaBoost، MP (پرسپترون چند لایه)، J48 با روش پیشنهادی مقایسه شده اند. مطابق با نتایج به دست آمده، روش پیشنهادی توانسته است معیارهای ارزیابی را نسبت به الگوریتم‌های دیگر، بهبود دهد. الگوریتم علی بابا و چهل دزد دارای چندین مزیت متمایز بر اساس اصل اساسی خود است که باعث شده است نتایج به میزان مطلوبی بهبود یابد. برتری اول، مدل های به روز رسانی موقعیت از الگوریتم علی بابا و چهل دزد به طور موثر به افراد جمعیت کمک می کند تا هر منطقه را در فضای جست‌وجو اکتشاف و بهره‌برداری کنند. برتری دوم، جستجوی تصادفی که دزدان در فضای جست‌وجو استفاده می کنند، نه تنها تنوع جمعیت را افزایش می دهد، بلکه سرعت همگرایی را نیز تضمین می کندکه نشان دهنده تعادل کارآمد بین اکتشاف و بهره برداری است. برتری سوم، تعداد پارامترها در الگوریتم علی بابا و چهل دزد کم است، اما آنها توانایی خوبی برای بهبود قدرت و عملکرد آن دارند. برتری چهارم، بار محاسباتی الگوریتم علی بابا و چهل دزد کم است.

یکی از عیوب روش پیشنهادی، می تواند بار وارد شده به دلیل استفاده از دو الگوریتم باشد. به ویژه اینکه الگوریتم ها به صورت سری کار می کنند و الگوریتم علی بابا و چهل دزد مبتنی بر تکرار است.

جدول (۳): مقایسه نتایج روش پیشنهادی از نظر معیارهای مختلف

Accuracy (%)	ROC (%)	F1-score (%)	Precision (%)	دسته بند
۹۰,۴۸	۹۶,۵	۹۱,۵	۹۴,۳	LR
۸۹,۶۹	۹۰,۳	۹۰,۷	۹۵,۳	SVM
۹۱,۶۷	۹۳	۹۲,۸	۹۳,۴	KNN
۸۶,۳۶	۹۲,۴	۸۸	۹۰	AdaBoost
۹۱,۸۶	۹۷,۲	۹۲,۸	۹۵,۶	MP
۹۲,۹۳	۹۵,۴	۹۳,۸	۹۴,۸	J48
۹۴,۶۵	۹۸,۶	۹۵,۳	۹۶,۴	RF
۹۵,۳۹	۹۸,۱۲	۹۵,۸۴	۹۵,۳۹	روش پیشنهادی

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

۵- نتیجه گیری

با استفاده از یافتن وزن و بایاس‌های شبکه عصبی از طریق الگوریتم علی بابا و چهل دزد می‌توان صفحات فیشینگ را با دقت بالایی شناسایی کرد. در روش پیشنهادی برای دسته‌بندی و تشخیص حملات فیشینگ از شبکه عصبی پرسپترون چند لایه استفاده می‌شود. با داشتن مجموعه‌ای از وزن‌ها و مقدار بایاس شبکه عصبی پرسپترون، خروجی متناسب با داده‌های ورودی و وزن‌ها تولید می‌کند. وزن ها از طریق الگوریتم علی بابا و چهل دزد پیدا می‌شوند. تابع هدف در حالت کلی همان تابع هزینه‌ای است که قرار است الگوریتم علی بابا و چهل دزد آن را حل کند. دستاورد الگوریتم علی بابا و چهل دزد به‌عنوان متغیر، W ها (پارامترهای وزن‌ها و بایاس‌ها) هستند. هدف این است که الگوریتم علی بابا و چهل دزد W ها را رفته‌رفته بهینه کند؛ بنابراین ورودی، W هایی هستند که الگوریتم علی بابا و چهل دزد مشخص می‌کند. نکته مهم، انتخاب روشی است که تابع هزینه با آن محاسبه شود که شامل 'MSE'، 'RMSE' و 'Accuracy' هستند. مجموعه داده مورد استفاده از سایت https://archive.ics.uci.edu/ml/datasets/Website+Phishing است. در این مجموعه داده، ویژگی‌های مختلف مربوط به وب‌سایت‌های قانونی و فیشی را شناسایی و ۱۳۵۳ وب‌سایت مختلف را از منابع مختلف جمع‌آوری کرده‌اند. نتایج روش پیشنهادی با طرح پایه از نظر دقت، صحت، F1_Score و منحنی AUC-ROC مقایسه می‌شوند.

هدف مطالعه آینده، شناسایی وب‌سایت‌های مشکوک با افزودن لایه‌های بیشتر در شبکه عصبی و استفاده از شبکه‌های عصبی دقیق‌تر است و یک رویکرد مبتنی بر یادگیری عمیق برای شناسایی وب‌سایت‌های فیشینگ از طریق تجزیه‌وتحلیل URL ارائه خواهد شد. همچنین می‌توان از الگوریتم‌های فراابتکاری دیگر مانند بهینه‌سازی اسب برای یافتن وزن‌های شبکه عصبی استفاده کرد.

منابع

Abedin, N. F., Bawm, R., Sarwar, T., Saifuddin, M., Rahman, M. A., & Hossain, S. (2020). Phishing attack detection using machine learning classification techniques. In 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS) (pp. 1125-1130). IEEE.

Al-Andoli, M. N., Tan, S. C., Sim, K. S., Lim, C. P., & Goh, P. Y. (2022). Parallel Deep Learning with a hybrid BP-PSO framework for feature extraction and malware classification. Applied Soft Computing, 131, 109756.

Almseidin, M., Alkasassbeh, M., Alzubi, M., & Al-Sawwa, J. (2022). Cyber-Phishing Website Detection Using Fuzzy Rule Interpolation. Cryptography, 6(2), 24.

Altaher, A. (2017). Phishing websites classification using hybrid SVM and KNN approach. International Journal of Advanced Computer Science and Applications, 8(6).

Alzubi, O. A., Alzubi, J. A., Al-Zoubi, A. M., Hassonah, M. A., & Kose, U. (2022). An efficient malware detection approach with feature weighting based on Harris Hawks optimization. Cluster Computing, 25(4), 2369-2387.

Barreiro Herrera, D. A., & Camargo Mendoza, J. E. (2022). A Systematic Review on Phishing Detection: A Perspective Beyond a High Accuracy in Phishing Detection. In International Conference on Applied Informatics (pp. 173-188). Springer, Cham.

Bhagwat, S., & Gupta, G. P. (2022). Android Malware Detection Using Hybrid Meta-heuristic Feature Selection and Ensemble Learning Techniques. In International Conference on Advances in Computing and Data Sciences (pp. 145-156). Springer, Cham.

Braik, M., Ryalat, M. H., & Al-Zoubi, H. (2022). A novel meta-heuristic algorithm for solving numerical optimization problems: Ali Baba and the forty thieves. Neural Computing and Applications, 34(1), 409-455.

Darshan, S. S., Kumara, M. A., & Jaidhar, C. D. (2016). Windows malware detection based on cuckoo sandbox generated report using machine learning algorithm. In 2016 11th International Conference on Industrial and Information Systems (ICIIS) (pp. 534-539). IEEE.

Das, S., Nippert-Eng, C., & Camp, L. J. (2022). Evaluating user susceptibility to phishing attacks. Information & Computer Security.

Dhiyanesh, B., Selvanathan, N., Kiruthiga, G., & Radha, R. (2021). Effective attribute selection and classification technique for phishing attacks detection. In 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA) (pp. 1-7). IEEE.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

Jafar, M. T., Al-Fawa'reh, M., Barhoush, M., & Alshira'H, M. H. (2022). Enhanced Analysis Approach to Detect Phishing Attacks During COVID-19 Crisis. Cybernetics and Information Technologies, 22(1), 60-76.

Jain, A. K., & Gupta, B. B. (2022). A survey of phishing attack techniques, defence mechanisms and open research challenges. Enterprise Information Systems, 16(4), 527-565.

Jalil, S., Usman, M., & Fong, A. (2022). Highly accurate phishing URL detection based on machine learning. Journal of Ambient Intelligence and Humanized Computing, 1-19.

Kovač, A., Dunđer, I., & Seljan, S. (2022). An overview of machine learning algorithms for detecting phishing attacks on electronic messaging services. In 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO) (pp. 954-961). IEEE.

Minocha, S., & Singh, B. (2022). A novel phishing detection system using binary modified equilibrium optimizer for feature selection. Computers & Electrical Engineering, 98, 107689.

Palša, J., Ádám, N., Hurtuk, J., Chovancová, E., Madoš, B., Chovanec, M., & Kocan, S. (2022). MLMD—A Malware-Detecting Antivirus Tool Based on the XGBoost Machine Learning Algorithm. Applied Sciences, 12(13), 6672.

Ripa, S. P., Islam, F., & Arifuzzaman, M. (2021). The emergence threat of phishing attack and the detection techniques using machine learning models. In 2021 International Conference on Automation, Control and Mechatronics for Industry 4.0 (ACMI) (pp. 1-6). IEEE.

Sabahno, M., & Safara, F. (2022). ISHO: improved spotted hyena optimization algorithm for phishing website detection. Multimedia Tools and Applications, 81(24), 34677-34696.

Stobbs, J., Issac, B., & Jacob, S. M. (2020). Phishing web page detection using optimised machine learning. In 2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom) (pp. 483-490). IEEE.

Uplenchwar, S., Sawant, V., Surve, P., Deshpande, S., & Kelkar, S. (2022). Phishing Attack Detection on Text Messages Using Machine Learning Techniques. In 2022 IEEE Pune Section International Conference (PuneCon) (pp. 1-5). IEEE.

Vijay, J. S., Kulkarni, K., & Arya, A. (2022). Metaheuristic Optimization of Neural Networks for Phishing Detection. In 2022 3rd International Conference for Emerging Technology (INCET) (pp. 1-5). IEEE.

Zhang, X., Shi, D., Zhang, H., Liu, W., & Li, R. (2018). Efficient detection of phishing attacks with hybrid neural networks. In 2018 IEEE 18th International Conference on Communication Technology (ICCT) (pp. 844-848). IEEE.

Zhu, E., Yuan, Q., Chen, Z., Li, X., & Fang, X. (2022). CCBLA: a Lightweight Phishing Detection Model Based on CNN, BiLSTM, and Attention Mechanism. Cognitive Computation, 1-14.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

مدل تشخیص فیشینگ URL ها بر اساس یادگیری ماشین

پریسا دانشجو [۱]، سعید احمدی [۲]

دانشگاه آزاد اسلامی- واحد تهران غرب[۱]، Daneshjoo.p@wtiau.acir
دانشگاه آزاد اسلامی- واحد تهران غرب[۲]، Saeed.ahmadi.edu@gmail.com[۲]

چکیده

حمله های فیشینگ همیشه تهدیدات قابل توجهی برای امنیت اینترنت بوده اند. یکی از معمولی‌ترین شکل‌های فیشینگ، از طریق URLها است، جایی که مهاجمان، URLهای تقلبی را به شکل URLهای معتبر در می‌آورند تا کاربران گول بخورند و برروی آنها کلیک کنند. فنون یادگیری ماشینی، امیدهایی برای شناسایی URLهای فیشینگ به‌وجود آورده اند، اما اثربخشی آنها براساس رویکرد استفاده شده می‌تواند تغییر کند.

اهداف: هدف این پژوهش، پیشنهاد دو روش یادگیری ماشینی، «شبکه های عصبی کانوالی» (CNN) و «خود توجهی چندسره» (MHSA)، برای شناسایی URLهای فیشینگ است. علاوه بر آن، ارزیابی و مقایسه اثربخشی این رویکرد در مقایسه با روش ها و مدل های دیگر است.

روش تحقیق: یک مجموعه داده از URLها گردآوری و به آنها برچسب فیشینگ یا معتبر داده شد. عملکرد چندین مدل استفاده کننده از روش های یادگیری ماشینی مختلف، شامل CNN و MHSA، برای دسته بندی این URLها با استفاده از معیارهای مختلف، مانند دقت، صحت، فراخوانی و نمره F1، ارزیابی شد.

نتایج: نتایج نشان می‌دهند که ترکیب مدل‌های CNN و MHSA عملکرد بهتری نسبت به دیگر مدل‌های انفرادی دارد و به دقت ۹۸/۳% می‌رسد که در مقایسه با روش‌های نوین موجود، بهبود قابل توجهی در شناسایی URLهای فیشینگ فراهم می‌کند.

نتیجه‌گیری: ترکیب CNN و MHSA رویکردی موثر برای آشکارسازی URLهای فیشینگ است. این روش نسبت به روش‌های نوین موجود عملکرد بهتری دارد و روشی دقیق و مطمئن‌تر برای آشکارسازی URLهای فیشینگ فراهم می‌کند. نتایج این مطالعه، پتانسیل استفاده از روش‌های ترکیبی در بهبود دقت و اطمینان روش‌های آشکارسازی URL فیشینگ مبتنی بر یادگیری ماشینی را نشان می‌دهند.

کلمات کلیدی – فیشینگ؛ آدرس URL؛ یادگیری عمیق؛ لایه کانوالی؛ خود توجهی چندسر

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

مقدمه

فیشینگ یکی از رایج ترین روش های دسترسی به داده های شخصی است [۱]. برای به حداقل رساندن آسیب یک حمله فیشینگ، لازم است که آن را در کوتاه ترین زمان ممکن شناسایی کنیم. تقریباً تمامی انواع حمله فیشینگ از URLهای فیشینگ استفاده می‌کنند [۱].

URLهای فیشینگ به وبسایت‌ها یا صفحات وبی لینک شده اند. این وبسایت ها یاصفحات وب به‌نحوی طراحی شده اند که شبیه وبسایت‌های معتبر و قانونی به نظر برسند، اما در واقعیت کار سایت‌های مخربی هستند که توسط مجرمان سایبری ایجاد شده اند. آن‌ها با استفاده از این وبسایت‌ها اطلاعات شخصی همچون اطلاعات ورود به حساب (لاگین)، شماره های کارت اعتباری و سایر داده های حساس را سرقت می کنند [۱].

نمی‌توان بیشتر از این حد روی اهمیت تشخیص URLهای فیشینگ تاکید کرد، زیرا آن‌ها تهدید مهمی نسبت به افراد و سازمان‌ها هستند. در اینجا دلایلی راجع به ضرورت تشخیص URLهای فیشینگ را بیان می‌کنیم [۱]:

1- حفاظت در برابر سرقت هویت: URLهای فیشینگ معمولاً برای فریب افراد و افشای اطلاعات لاگین، جزئیات حساب بانکی و سایر اطلاعات شخصی طراحی شده اند. افراد و سازمان‌ها می‌توانند با تشخیص این URLها از خود در برابر سرقت هویت، حفاظت کنند [۱].
2- جلوگیری از زیان‌های مالی: حملات فیشینگ زیان‌های مالی شدیدی را به افراد و سازمان‌ها تحمیل می‌کنند. سازمان‌ها می‌توانند با تشخیص و مسدودکردن URLهای فیشینگ، از سرقت پول و داده های حساس توسط مجرمان سایبری جلوگیری کنند [۱].
3- حفاظت در برابر بدافزار: URLهای فیشینگ معمولاً حاوی لینک‌هایی به نرم افزارهای مخرب هستند و این نرم افزارها در شرایط خاصی به رایانه یا شبکه آسیب می‌زنند. سازمان‌ها با تشخیص و مسدودسازی این URLهای فیشینگ، از آلوده شدن به بدافزار و رخنه داده ها جلوگیری کنند [۱].
4- حفظ اعتماد: ممکن است سازمان‌هایی که قربانی حملات فیشینگ هستند اعتماد مشتریان، کارفرمایان و شرکای خود را از دست بدهند. سازمان‌ها با تشخیص و پیشگیری از حملات فیشینگ، اعتبارشان را حفظ و از گسترش اخبار منفی جلوگیری می کنند [۱].

به‌طور خلاصه، تشخیص URLهای فیشینگ، نقش اساسی را در حفاظت در برابر سرقت هویت، زیان مالی، خنثی‌سازی آلودگی بدافزار و حفظ اعتماد دارد. افراد و سازمان‌ها باید هوشیاری لازم را برای شناسایی و گزارش دهی URLهای فیشینگ داشته باشند تا در جهان دیجیتال امن باقی بمانند.

ادبیات پژوهش

فیشینگ چیست؟

فیشینگ یکی از رایج‌ترین روش‌های دسترسی به داده های شخصی است[۱]. برای به حداقل رساندن آسیب یک حمله فیشینگ، لازم است که آن را در کوتاه‌ترین زمان ممکن، شناسایی کنیم. تقریباً تمامی انواع حمله فیشینگ از URLهای فیشینگ استفاده می‌کنند[۱].

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

URLهای فیشینگ به وبسایت‌ها یا صفحات وبی لینک شده اند. این وبسایت‌ها یا صفحات وب به‌نحوی طراحی شده اند که شبیه وبسایت‌های معتبر و قانونی به نظر برسند، اما در واقعیت کار سایت‌های مخربی هستند که توسط مجرمان سایبری ایجاد شده اند. آن‌ها با استفاده از این وبسایت‌ها اطلاعات شخصی همچون اطلاعات ورود به حساب (لاگین)، شماره‌های کارت اعتباری و سایر داده‌های حساس را سرقت می‌کنند [۱].

نمی توان بیشتر از این حد روی اهمیت تشخیص URLهای فیشینگ تاکید کرد، زیرا آن‌ها تهدید مهمی نسبت به افراد و سازمان ها هستند. در اینجا دلایلی راجع به ضرورت تشخیص URLهای فیشینگ را بیان می کنیم [۱]:

5- حفاظت در برابر سرقت هویت: URLهای فیشینگ معمولاً برای فریب افراد و افشای اطلاعات لاگین، جزئیات حساب بانکی و سایر اطلاعات شخصی طراحی شده اند. افراد و سازمان ها می توانند با تشخیص این URLها از خود در برابر سرقت هویت حفاظت کنند [۱].
6- جلوگیری از زیان های مالی: حملات فیشینگ زیان های مالی شدیدی را به افراد و سازمان ها تحمیل می کنند. سازمان ها می توانند با تشخیص و مسدودکردن URLهای فیشینگ، از سرقت پول و داده های حساس توسط مجرمان سایبری جلوگیری کنند [۱].
7- حفاظت در برابر بدافزار: URLهای فیشینگ معمولاً حاوی لینک هایی به نرم افزارهای مخرب هستند، و این نرم افزارها در شرایط خاصی به رایانه یا شبکه آسیب می زنند.سازمان ها با تشخیص و مسدودسازی این URLهای فیشینگ، از آلوده شدن به بدافزار و رخنه داده ها جلوگیری کنند [۱].
8- حفظ اعتماد: ممکن است سازمان هایی که قربانی حملات فیشینگ هستند اعتماد مشتریان، کارفرمایان و شرکای خود را از دست بدهند. سازمان ها با تشخیص و پیشگیری از حملات فیشینگ، اعتبارشان را حفظ می کنند و از گسترش اخبار منفی جلوگیری می کنند [۱].

به‌طور خلاصه، تشخیص URLهای فیشینگ، نقش اساسی را در حفاظت در برابر سرقت هویت، زیان مالی، خنثی سازی آلودگی بدافزار و حفظ اعتماد دارد. افراد و سازمان ها باید هوشیاری لازم برای شناسایی و گزارش دهی URLهای فیشینگ را داشته باشند تا در جهان دیجیتال امن باقی بمانند.

URLهای فیشینگ و یادگیری ماشین چه هستند؟

یک URL فیشینگ، یک لینک مخرب است که مهاجم روی اینترنت توزیع می کند. هدف وی این است که کاربران را فریب داده و به داده های حساس آن ها همچون رمزهای عبور، شماره‌های کارت اعتباری و سایر اطلاعات شخصی دسترسی پیدا کند.

یادگیری ماشین یک تکنیک هوش مصنوعی است که طی آن الگوریتم های رایانه ای بر اساس حجم های وسیعی از داده ها آموزش دهی می شوند. در مقوله تشخیص URLهای فیشینگ، می‌توان از یادگیری ماشین برای تشخیص الگوهای مشکوک در آدرس های لینک استفاده کرد. این الگوها می‌توانند نشانه ای از حملات فیشینگ بالقوه باشند.

تشخیص URLهای فیشینگ با یادگیری ماشین از تحلیل حجم های وسیعی از داده ها استفاده می کند. این داده ها حاوی ویژگی های مختلفی همچون URL، شرایط ظاهری صفحه وب، زمینه و غیره هستند. می توان مدل های یادگیری ماشینی که برای تشخیص URLهای فیشینگ به‌کار رفته را روی نمونه های واقعی سایت های فیشینگ و سایت های معتبر (غیرفیشینگ) آموزش داد. بدین ترتیب آن ها امکان شناسایی لینک های مشکوک بر اساس مدل آموزش دهی شده را به‌دست می آورند.

بر این اساس، استفاده از یادگیری ماشین برای تشخیص URLهای فیشینگ، روش موثری برای حفاظت از کاربران در برابر حملات سایبری مرتبط با فیشینگ است.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

ارزیابی روش های مختلف تشخیص فیشینگ

مشکل تشخیص URLهای فیشینگ این است که آن ها به‌نحوی طراحی شده که دقیقاً شبیه URLهای قانونی به نظر برسند، بنابراین وضعیت کاربران برای تفکیک آن ها از URLهای معتبر دشوار می شود. این URLها اکثراً شبیه وبسایت های شناخته شده همچون وبسایت های بانکی یا تجارت الکترونیک طراحی می شوند، و با فریب کاربران باعث افشای اطلاعات حساس می شوند [۴].

یکی از چالش های تشخیص URLهای فیشینگ این است که آن ها به شدت هدفمند و شخصی سازی شده هستند، در نتیجه تشخیص شان با روش های مرسوم مبتنی بر قوانین دشوار است. علاوه بر این، روش های فیشینگ پیچیده تر شده اند و برای تشخیص به تکنیک های پیشرفته تری نیاز دارند [۳].

برای مقابله با حملات فیشینگ، چندین روش برای تشخیص URLهای فیشینگ توسعه یافته‌اند. این روش ها عبارتند از:

1- لیست های سیاه: لیست های سیاه حاوی URLهای فیشینگ معلومی هستند که توسط متخصصان امنیت شناسایی شده اند. مرورگرها، ارائه دهندگان ایمیل و نرم افزارهای امنیتی می توانند از این لیست ها برای جلوگیری از دسترسی کاربران به وبسایت های فیشینگ شناخته شده استفاده کنند [۱]. بسیاری از برندهای محبوب همچون گوگل، مایکروسافت، اپل و خیلی موارد دیگر، از لیست سیاه به عنوان ابزاری برای محافظت در برابر URLهای فیشینگ استفاده می کنند. این شرکت ها از روش های مختلفی برای نگهداری و به‌روزرسانی لیست های سیاه خود استفاده می کنند که خزنده های خودکار و گزارش‌های کاربر دو نمونه از آن ها هستند. به عنوان مثال، سرویس مرورگری امن گوگل، به‌طور مرتب لیست وبسایت های غیرامن خود، از جمله وبسایت های فعال در حملات فیشینگ را به‌روزرسانی می‌کند و قبل از بازدید کاربران از آن ها هشدار صادر می‌کند. فیلتر کنترل هوشمند مایکروسافت در مرورگرهای اج و اینترنت اکسپلورر ادغام شده و با لیست سیاه از کاربران در برابر وبسایت های به‌طور بالقوه مخرب حفاظت می کند [۱۲].
2- فیلترهای سیستم نام دامنه (DNS): می توان از این فیلترها برای مسدودسازی دسترسی به URLهای فیشینگ شناخته شده استفاده کرد. هنگامی که کاربری سعی می‌کند به یک وبسایت فیشینگ شناخته شده دسترسی پیدا کند، فیلتر DNS کاربر را به یک صفحه امن انتقال داده یا دسترسی به سایت را به‌طور کامل قطع می‌کند[٦]. چندین برند محبوب وجود داشته که از فیلترینگ DNS به عنوان روشی برای حفاظت در برابر URLهای فیشینگ استفاده می کنند. برخی از این برندها شامل سیسکو، باراکودا نتورکز، سوفوس، مک آفی و سیمانتک هستند. این شرکت ها خدمات فیلترینگ DNS را ارائه داده و به سازمان ها کمک کرده دسترسی به سایت های فیشینگ شناخته شده و نیز آدرس های IP و دامنه های مخرب را ممنوع کنند. سازمان ها با اتکا به این روش، به صورت بیش فعالانه ای از شبکه ها و کاربران خود در برابر حملات فیشینگ حفاظت می کنند.
3- آموزش های آگاه سازی کاربر: آموزش دهی به کاربران راجع به خطرات حملات فیشینگ و نحوه آشکارسازی (تشخیص) URLهای فیشینگ، می تواند روش موثری برای جلوگیری از این حملات باشد. می توان به کاربران آموزش داد که نشانه های URLهای فیشینگ، از جمله اشتباه در املای نام دامنه یا وجود کاراکترهای غیرعادی در URL را جستجو کنند [٦]. شرکت های محبوب زیادی همچون مایکروسافت، گوگل و آمازون آموزش های آگاه سازی کاربر را به عنوان بخشی از پروتکل های امنیتی شان به کاربران ارائه می کنند. برای نمونه، مایکروسافت آموزش های زیادی از جمله وبینارها و دوره های آنلاین را فراهم کرده است و به کارکنانش کمک می کند طرح های فیشینگ را شناسایی و از آن ها اجتناب کنند. گوگل نیز منابع مشابهی فراهم می کند که حملات فیشینگ شبیه سازی شده یکی از آن ها است. این حملات میزان آگاهی کارکنان را تست می کنند. همچنین بسته های آموزشی فراهم شده که به ارتقای مهارت کارکنان می انجامند [۱۵].

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

4- الگوریتم های یادگیری ماشین: می توان از الگوریتم های یادگیری ماشین برای تشخیص URLهای فیشینگ استفاده کرد. این کار با تحلیل خصوصیات URL از جمله نام دامنه، طول URL و وجود کلیدواژه های خاص انجام می شود. این الگوریتم ها توانایی تشخیص شباهت های بین URLهای فیشینگ و وبسایت های شناخته شده فیشینگ را هم دارند [۲]. یادگیری ماشین در مقایسه با لیست سیاه یا فیلتر DNS رویکرد موثرتری برای آشکارسازی URLهای فیشینگ محسوب می شود. دلیلش این است که می تواند با تهدیدهای جدید و در حال تکامل تطبیق پیدا کند. لیست سیاه و فیلترینگ DNS به نگهداری از لیست های URLهای فیشینگ شناخته شده یا دامنه ها وابسته هستند و هنگامی که مهاجمان URLها یا دامنه های جدیدی ایجاد می کنند بلافاصله منسوخ می شوند [۱۳].

از سوی دیگر، مدل های یادگیری ماشین می توانند الگوها و ویژگی های URLها و صفحات وب را تحلیل کنند و حملات فیشینگ جدید و ناشناخته را حتی در صورت رویت نشدن در گذشته تشخیص دهند. همچنین این مدل ها می توانند از داده های گذشته اطلاعات مفیدی را استخراج کنند و دقت را در طول زمان افزایش دهند، لذا در تشخیص URLهای فیشینگ کارآمدتر می شوند.

علاوه بر این، یادگیری ماشین می تواند ویژگی های مختلف فرای URL یا دامنه را تحلیل کند که از جملهآنها ظاهر صفحه وب و زمینه ای است که لینک در آن قرار داشته است. بدین ترتیب کار مهاجمان برای ممانعت از تشخیص، صرفاً با استفاده از دامنه ها یا URLهای مختلف دشوارتر می شود [۱۴]. به طور کُلی، یادگیری ماشین رویکرد بیش فعالانه تر و تطبیقی برای آشکارسازی URLهای فیشینگ است، بنابراین ابزار مناسب تری برای حفاظت در برابر تهدیدهای سایبری رو به گسترش تلقی می شود.

به عنوان جمع بندی بحث، باید گفت: آشکارسازی URLهای فیشینگ بخش اساسی جلوگیری از حملات فیشینگ است. روش های مختلفی برای تشخیص URLهای فیشینگ توسعه داده شده اند، از جمله لیست های سیاه، فیلترهای DNS، الگوریتم های یادگیری ماشین و آموزش آگاهی کاربر. این روش ها به افراد و سازمان ها کمک کرده در برابر حملات فیشینگ امن باقی بمانند و از اطلاعات حساس شان حفاظت کنند.

می توان از این روش ها و مدل های یادگیری ماشین به عنوان راه حل مسئله استفاده کرد:

1- یادگیری نظارت شده: این روش با آموزش دهی یک مدل یادگیری ماشین؛ روی مجموعه داده برچسب گذاری شده از URLهای فیشینگ و URLهای قانونی اجرا می‌شود. می توان از این مدل برای طبقه بندی URLهای جدید به دو دسته فیشینگ و قانونی، بر اساس الگوهای یادگیری شده در طول آموزش استفاده کرد [۱۷].
2- یادگیری نظارت نشده: در این روش، مدل یادگیری ماشین بر روی مجموعه داده غیربرچسب گذاری شده از URLها آموزش داده می شود و تشخیص الگوها و ناهنجاری ها در داده ها که نشانه بالقوه URLهای فیشینگ بوده یاد گرفته می شود [۱۷].
3- یادگیری نیمه نظارتی: این روش مولفه های یادگیری نظارت شده و غیرنظارت شده را ترکیب می کند. مدل یادگیری ماشین روی مجموعه کوچکی از داده های برچسب گذاری شده URLهای فیشینگ و قانونی آموزش داده می شود. همچنین یادگیری روی مجموعه داده غیربرچسب گذاری شده برای تشخیص الگوهای جدید و ناهنجاری ها در داده ها به‌کار می رود [۱۷].
4- یادگیری عمیق: روش های یادگیری عمیق از جمله شبکه های عصبی کانوالی (CNN) یا شبکه های عصبی بازگشتی (RNN) برای تشخیص URLهای فیشینگ کاربرد دارند. آن ها ویژگی ها را مستقیماً از داده های خام همچون اسکرین شات های وبسایت یا لاگ های ترافیک شبکه یاد می گیرند [۱۷].
5- یادگیری ترکیبی: در این روش چندین مدل یادگیری ماشین با هم ترکیب شده تا عملکرد کُلی ارتقا یابد. روش های ترکیبی می توانند به طور مشخص برای آشکارسازی URLهای فیشینگ مفید باشند. در حقیقت آن ها انواع مختلف مدل ها با نقاط قوت و ضعف مختلف را با یکدیگر ترکیب می کنند [۳][۱۷].

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

هر روش نقاط ضعف و قوت خاص خودش را دارد، لذا باید چندین روش را آزمایش کرد تا موثرترین روش تشخیص URLهای فیشینگ با یادگیری ماشین شناسایی شود.

روش های یادگیری عمیق

یادگیری عمیق یکی از زیرمجموعه های یادگیری ماشین است و از شبکه های عصبی مصنوعی برای تحلیل و طبقه بندی داده ها استفاده می کند. اما روش های مرسوم یادگیری ماشین به انتخاب و مهندسی غیرماشینی(دستی) ویژگی ها برای طبقه بندی وابسته هستند. این مدل ها ویژگی ها را از روی داده های خام یاد می گیرند و می توانند الگوهای پیچیده تری را کشف کنند. این روش ها معمولاً URL عادی را به یک ماتریس تبدیل می کنند [۱۸] سپس شاخص هایی را برای مدل منتخب فراهم می کنند تا قانونی بودن یا فیشینگ بودن URL را مشخص کنند.

از لحاظ آشکارسازی URLهای فیشینگ، رویکردهای یادگیری عمیق نرخ های دقت بالاتری را نسبت به روش های مرسوم یادگیری ماشین فراهم می کنند. علتش این است که مدل های یادگیری عمیق می توانند الگوها و ویژگی های حساس را شناسایی کنند که ممکن است در حالت عادی برای انسان ها یا روش های مرسوم یادگیری ماشین مخفی باقی می مانند.

البته بهتر است اشاره کنیم که رویکردهای یادگیری عمیق به داده ها و منابع محاسباتی بیشتری در مقایسه با روش های مرسوم یادگیری ماشین نیاز دارند و این حالت برای برخی سازمان ها دردسرساز است. البته انتخاب بین رویکردهای یادگیری عمیق و روش های مرسوم یادگیری ماشین، به نیازها و منابع خاص هر سازمان بستگی دارد.

شبکه عصبی کانوالی

شبکه عصبی کانوالی (CNN) یک نوع الگوریتم یادگیری عمیق است که برای وظایف تشخیص تصویر و الگو بسیار کارآمد است. این روش یک تصویر ورودی را دریافت می کند، یک سری فیلترها را اعمال کرده تا ویژگی ها را در سطوح انتزاعی مختلف استخراج کند سپس از ویژگی ها برای طبقه بندی تصویر به دسته های مختلف استفاده می کند [۱۶].

CNNها به‌طور متداول در کاربردهای بینایی رایانه به‌کار می روند و آشکارسازی URLهای فیشینگ یک مورد از آن ها است. دلیلش این است که URLهای فیشینگ غالباً حاوی تصاویر یا لوگوهایی هستند که برای تقلید وبسایت های قانونی و فریب کاربران برای کلیک روی آن ها طراحی شده اند. با آموزش دهی یک CNN روی یک مجموعه داده بزرگ از URLهای فیشینگ و قانونی، شبکه الگوها و ویژگی هایی را که نشانه های خرابکاری هستند، شناسایی می کنند.

یکی از دلایلی که CNNها بهترین روش تشخیص URLهای فیشینگ هستند، توانایی شان برای یادگیری خودکار از ویژگی های مطلوب داده های خام ورودی است. در حقیقت به‌جای اعمال دستی ویژگی ها بر اساس دانش دامنه، شبکه می تواند استخراج مناسب ترین ویژگی ها را با توجه به وظیفه موجود انجام دهد. علاوه بر این، CNNها به‌شدت مقیاس پذیر هستند و این یعنی آن ها می توانند مجموعه داده های بزرگ را مدیریت کنند و روی خوشه‌های رایانشی قوی آموزش داده شده تا دقت شان افزایش یابد.

خودتوجهی چندسره

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

خودتوجهی چندسره، تکنیکی است که در یادگیری عمیق، بخصوص در حوزه پردازش زبان طبیعی (NLP) به‌کار می رود. این حالت نمونه تعمیم یافته ای از خودتوجهی است که به شبکه عصبی این امکان را داده تا اهمیت بخش های مختلف توالی ورودی را در حین پردازش بسنجد. خودتوجهی چندسره این مفهوم را گسترش داده و عملیات خودتوجهی را به‌صورت موازی انجام می دهد. لذا مدل می تواند نمایش های متعدد از توالی ورودی را یاد بگیرد و الگوهای پیچیده تر را تشخیص دهد.

در زمینه آشکارسازی های URLهای فیشینگ، خودتوجهی چندسره برای استخراج ویژگی ها از متن URL و تشخیص الگوهای مهم مرتبط با فیشینگ کاربرد دارد. مدل با توجه به سرهای توجه چندگانه جنبه های مختلف URL را یاد می گیرد که وجود کلیدواژه های مشکوک یا نام های دامنه غیرعادی دو مورد هستند. آن ها به صورت یک نمایش نهایی ترکیب شده و برای طبقه بندی کاربرد دارند.

هرچند CNNها یک روش محبوب برای آشکارسازی URLهای فیشینگ هستند، اما خودتوجهی چندسره می تواند در شرایط خاص مزیت هایی را فراهم کند. برای مثال در پردازش توالی های طولانی متن موثرتر است، اما فیلترهای کانوالی مرسوم در تشخیص این الگوها دچار مشکل می شوند. علاوه بر این، خودتوجهی چندسره، تفسیرپذیرتر از CNNها است و به محققان اجازه می دهد بخش های مختلف ورودی را که در تصمیم نهایی طبقه بندی تاثیر بیشتری دارند بهتر درک کنند. این در حالی است که اثربخشی خودتوجهی چندسره در نهایت به مجموعه داده ها و مسائل مشخص بستگی دارد.

یافته‌ها

ما می توانیم با مدل های یادگیری ماشین وبسایت های فیشینگ را تشخیص دهیم. دقت این مدل ها به مجموعه داده های به‌کاررفته برای آموزش و تست، ویژگی های استخراجی از وبسایت ها و الگوریتم ها و طبقه بندی‌های به‌کاررفته بستگی دارد. مجموعه داده های مختلفی برای آموزش کارایی دارند که Alexa و خزنده عمومی (common crawl) برای سایت های قانونی برای فیشینگ قابل استفاده هستند. phish tank و open-fish هم برای URLهای مشکوک گزارشی توسط کاربران کاربرد دارند. از چندین ویژگی برای مقایسه روش‌های یادگیری ماشین استفاده شده که صحت، نرخ مثبت کاذب، بازیابی، دقت و امتیاز F-1 از آن جمله هستند. رایج ترین و موثرترین روش های تشخیص URL فیشینگ شامل بیز ساده، جنگل تصادفی، CNN، MLP، MHSA، LSTM، CNN+RNN هستند. صحت در تمام روش‌های بالا و بین ۹۶ تا ۹۹٫۸۴ % است، اما عوامل دیگری همچون زمان آموزش، منابع محاسباتی و مقاومت در برابر نویز هم برای تعیین روش مناسب تر و بهتر به‌کار می روند.

شناسایی تکنیک های انتخاب و مهندسی ویژگی

همانطور که گفتیم، چندین ویژگی هستند که برای مقایسه روش های یادگیری ماشین در امر تشخیص URLهای فیشینگ کاربرد دارند. برخی ویژگی های پرکاربرد به این شرحند:

1- صحت: درستی کُلی مدل طبقه بندی را می سنجد. محاسبه آن با تقسیم تعداد نمونه های به‌درستی طبقه بندی شده بر تعداد کل نمونه ها انجام می شود.
2- نرخ مثبت کاذب (FPR): نسبت پیش بینی های مثبت کاذب به تعداد کل نمونه های منفی را می سنجد. FPR بالا بدین معنی است که مدل URLهای غیرفیشینگ را به عنوان فیشینگ شناسایی کرده است.
3- بازیابی: نسبت پیش بینی های مثبت صحیح به تعداد کل نمونه های مثبت را می‌سنجد. بازیابی بالا یعنی مدل به‌درستی درصد بالایی از URLهای فیشینگ را تشخیص می دهد.
4- دقت: نسبت پیش بینی های مثبت صحیح به تعداد کل نمونه های پیش بینی شده مثبت را می سنجد. دقت بالا یعنی مدل با دقت URLهای فیشینگ را شناسایی می‌کند.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

5- امتیاز F-1: به عنوان میانگین موزون دقت و بازیابی تعریف شده است. یک امتیاز واحد و متعادل بین دقت و بازیابی را ارائه می دهد.

صحت، FPR، بازیابی، دقت، و امتیاز F1، مناسب ترین ویژگی ها برای مقایسه روش های یادگیری ماشین در تشخیص URLهای فیشینگ هستند. دلیلش این است که آن ها ارزیابی کاملی از عملکرد مدل فراهم می کنند. هرچند صحت درستی کُلی را می سنجد، اما دقت، FPR، بازیابی و امتیاز F-1 اطلاعاتی از توانایی مدل برای تشخیص URLهای فیشینگ فراهم کرده و از مثبت های کاذب ممانعت می کنند. این شاخص ها نقش اساسی را در تشخیص فیشینگ دارند و ناتوانی در تشخیص URLهای فیشینگ منجر به تهدیدهای امنیتی می شود. نرخ مثبت کاذب بالا نیز باعث ناامیدی کاربر و بی‌اعتمادی به سیستم می شود.

برای درک بهتر نحوه تحلیل شاخص ها، نتایج را اینگونه بیان می کنیم:

- مثبت صحیح (TP) تعداد URLهای فیشینگی است که به درستی طبقه بندی شده است.
- منفی صحیح (TN) تعداد URLهای قانونی است که به عنوان قانونی طبقه‌بندی شده است.
- مثبت کاذب (FP) تعداد URLهای قانونی است که به عنوان فیشینگ طبقه‌بندی شده است.
- منفی کاذب (FN) تعداد URLهای قانونی است که به عنوان قانونی طبقه‌بندی شده است.

پس از انتخاب ویژگی ها برای ارزیابی و مقایسه مدل ها، می توان آن ها را اینگونه تعریف کرد:

- TP + TN/ TP + TN + FN =Acc
- FP/ FP + TN =FRP
- TP/ TP + FN = Rec
- TP / TP + FP =Pre
- 2* Pre*Rec / Pre + Rec =F1

مقایسه عملکرد

برخی از پرکاربردترین و موثرترین روش‌های به‌کاررفته برای تشخیص URLهای فیشینگ به این شرح هستند:

1- بیز ساده: یک الگوریتم احتمالاتی مبتنی بر قضیه بیز است. فرض می شود تمام ویژگی ها مستقل از هم هستند. این الگوریتم با محاسبه احتمال فیشینگ بودن URL بر اساس وقوع ویژگی های خاص در آن عمل می کند. الگوریتم روی مجموعه داده های برچسب گذاری شده آموزش داده شده و در حین تست، از احتمالات فراگیری شده برای طبقه بندی URLهای جدید به دو دسته قانونی یا فیشینگ استفاده می کند. بیز ساده یک روش آسان و سریع است و کارایی خوبی در تشخیص URLهای فیشینگ نشان داده است [۱].

2- جنگل تصادفی: یک الگوریتم یادگیری ترکیبی است که درخت های تصمیم متعدد را ایجاد و پیش بینی ها را ترکیب کرده تا به تصمیم نهایی برسد. هر درخت جنگل روی یک زیرمجموعه تصادفی از مجموعه اطلاعات آموزش داده شده است. در حین تست، الگوریتم پیش بینی های تمام درخت ها را ترکیب کرده تا تصمیم نهایی را اتخاذ کند. جنگل تصادفی به‌خاطر دقت و مقاومت در برابر داده های نویزی مشهور است [۱].

3- CNN (شبکه عصبی کانوالی): یک الگوریتم یادگیری عمیق بوده که با الهام گیری از ساختار مغز انسان ساخته شده است. آن ها با اعمال فیلترهای کانوالی برای استخراج ویژگی ها از داده های ورودی عمل می کنند. این ویژگی ها از لایه های نورون عبور کرده و نمایش های پیچیده از داده ها را یاد می گیرند. آن ها توانایی عملکرد مطلوب در تشخیص URLهای فیشینگ با یادگیری ویژگی هایی همچون نام دامنه، طول URL، و n-گرم های کاراکتر را نشان داده اند [۲۱].

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

4- MLP (پرسپترون چندلایه): یک نوع شبکه عصبی مصنوعی است که از چند لایه نورون تشکیل شده است. هر نورون شبکه ورودی ها را از لایه قبلی دریافت و تابع فعالسازی غیرخطی را اعمال می کند تا خروجی ایجاد کند. MLPها با پس انتشار، یک الگوریتم یادگیری نظارت شده آموزش داده می شوند. این الگوریتم وزن های نورون ها را به‌نحوی تنظیم کرده که خطای بین خروجی های واقعی و پیش بینی شده حداقل شود. MLPها عملکرد مناسبی در تشخیص URLهای فیشینگ دارند و این کار را با یادگیری عمر دامنه، مجوزهای SSL و طول دامنه انجام می دهند [۲۷].

5- MHSA (خودتوجهی چندسره): یک مدل تحولی و یکی از الگوریتم های یادگیری عمیق است که در پردازش زبان طبیعی کاربرد دارد. این الگوریتم با اعمال چند سر خودتوجهی به داده های ورودی عمل می کند تا اطلاعات را استخراج کند سپس خروجی های این سرها ادغام شده و از چند لایه نورون عبور کرده و برای یادگیری نمایش های داده به‌کار می روند. MHSA عملکرد مطلوبی در تشخیص URLهای فیشینگ داشته و ویژگی هایی همچون نام دامنه، طول URL، n-گرم های کاراکتر را یاد می گیرد [۲۷].

6- LSTM (حافظه طولانی کوتاه مدت): یک شبکه عصبی بازگشتی (RNN) است که برای غلبه بر مشکل گرادیان محوشونده RNNهای معمولی طراحی شده است. این شبکه ها قادر به یادگیری وابستگی های بلندمدت در داده های متوالی، با یادآوری و فراموشی گزینشی اطلاعات در طول زمان هستند. آن ها این توانایی را با سلول های حافظه به‌دست می‌آورند که به عنوان «واحدهای گیت دار»، ورود و خروج جریان اطلاعات از سلول را کنترل می کنند [۲۰].

7- CNN+ RNN: یک مدل یادگیری عمیق ترکیبی است که نقاط قوت CNN و RNN را در هم ادغام کرده است. برای تشخیص URLهای فیشینگ، بخش CNN ویژگی های محلی URLها را یاد می گیرد اما بخش RNN وابستگی های متوالی بین آن ها را نشان می دهد [۲۲].

با توجه به این جدول، مشاهده می کنیم که میزان صحت در تمام روش ها نسبتاً زیاد است و بین ۹۶ % الی ۹۹،۸۴ % است. اما صحت به تنهایی برای اثبات برتری کافی نیست زیرا عوامل دیگری همچون زمان آموزش دهی، منابع محاسباتی، مقاومت در برابر نویز هم باید در نظر گرفته شوند.

جدول ۴-۱: مقایسه عملکرد

ردیف	الگوریتم تشخیصی ML	صحت	FPR	بازیابی	دقت	F1
۱	بیز ساده [۱۹]	۹۷،۱۸				
۲	جنگل تصادفی [۲۱]	۹۷				
۳	CNN [۲۰]	۹۶،۶۱	۳،۵	۹۷،۰۹	۹۶،۶۱	۹۶،۸۵
۴	LSTM [۲۰]	۹۷،۲۰	۱،۸	۹۸،۶۳	۹۶،۴۵	۹۷،۵۳
۵	MLP [۲۰]	۹۶،۶۵		۹۶،۶۵	۹۶،۶۵	۹۶،۶۵
۶	CNN+ RNN [۲۲]	۹۷،۹	۳،۱۰	۹۸،۳۹	۹۶،۷۶	۹۷،۵۷
۷	CNN + LSTM [۲۳]	۹۳،۲۸	۱،۸۰	۹۷،۱۳	۹۹،۱۲	۹۸،۱۱

ارزیابی کارایی راه حل های موجود

با توجه به جدول ارائه شده، کارایی روش های مختلف یادگیری ماشین در تشخیص URLهای فیشینگ بر حسب شاخص های صحت، FPR، بازیابی، دقت و امتیاز F-1 ارزیابی می شود.

1- بیز ساده [۱۹]: صحت آن ۹۷،۱۸ % است که نسبتاً بالا است. البته هیچ اطلاعاتی راجع به FPR، بازیابی، دقت، F1 وجود ندارند.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

2- جنگل تصادفی [۲۱]: صحت ۹۷ % بوده اما اطلاعاتی راجع به FPR، بازیابی، دقت و F1 وجود ندارند.
3- CNN [۲۱]: صحت ۹۶,۶۱ % بوده، و FPR نسبتاً زیاد و ۳,۵ % است. بازیابی آن زیاد و ۹۷,۰۹ % است و این یعنی اکثر URLهای فیشینگ را به‌درستی تشخیص داده است، اما دقتش مقداری از ۹۶,۶۱ % کمتر است یعنی برخی URLهای قانونی را به عنوان URLهای فیشینگ طبقه بندی کرده است. امتیاز F1 ۹۶,۸۵ % است.
4- LSTM [۲۰]: بالاترین صحت به میزان ۹۷,۲% را دارد و FPR کم در حد ۱,۸ % دارد. بازیابی آن زیاد و ۹۸,۶۳ % است یعنی اکثر URLهای فیشینگ را به درستی شناسایی کرده است. البته دقتش مقداری کمتر از ۹۶,۴۵ % است و این یعنی برخی URLهای قانونی را به عنوان URL فیشینگ طبقه بندی می کند. امتیاز F1 این روش ۹۷,۵۳ % است.
5- MLP [۲۰]: صحت آن ۹۶,۶۵ % است، اما اطلاعاتی راجع به FPR، بازیابی، دقت و F1 وجود ندارند.
6- CNN + RNN [۲۲]: صحت آن ۹۷,۷ % است و FPR نسبتاً بالا در حد ۳,۱۰ % دارد. بازیابی آن ۹۸,۳۹ % است که مقدار زیادی است و نشان داده که اکثریت URLهای فیشینگ را به درستی تشخیص می دهد. اما دقتش کمی کمتر از ۹۶,۷۶ % است ، یعنی برخی URLهای قانونی را به عنوان فیشینگ طبقه بندی کرده است. امتیاز F1 آن ۹۷,۵۷ % است.
7- CNN + LSTM [۲۳]: این روش کمترین صحت به میزان ۹۳,۲۸ % را دارد و FPR آن کم و ۱,۸ % است. بازیابی آن بالا و ۹۷,۱۳ % است. این نشان داده که اکثر URLهای فیشینگ را به درستی تشخیص داده است. دقت آن بیشترین مقدار و ۹۹,۱۲ % است و این یعنی تعداد خیلی کمی از URLهای قانونی را به عنوان URL فیشینگ طبقه بندی کرده است. امتیاز F1 این روش ۹۸,۱۱ % است.

به عنوان جمع بندی، می توان گفت LSTM موثرترین روش از لحاظ صحت و FPR است. CNN + LSTM از بالاترین دقت برخوردار است و می توان گفت URLهای قانونی خیلی کمی را به عنوان URL فیشینگ طبقه بندی کرده است. CNN + RNN و بیز ساده صحت نسبتاً بالایی داشته اما FPR شان هم نسبتاً زیاد است. جنگل تصادفی و MLP صحت بالایی دارند اما اطلاعاتی راجع به FPR، بازیابی، دقت و F1 ارائه نمی دهند.

ترکیب دو روش بهتر از به‌کارگیری یک روش است زیرا ضعف ها و نقاط قوت هر یک را پوشش می دهد، و حالت ترکیبی منجر به بهبود عملکرد می شود. برای مثال CNN + RNN در داخل جدول صحت بیشتری از بیز ساده داشته، اما بیز ساده سریع تر است و به منابع محاسباتی کمتری نیاز دارد. با ترکیب این دو روش، سیستم تشخیص URL های مخرب دقیق تر و کارآمدتری را خواهیم داشت.

CNN (شبکه عصبی کانوالی) و MHSA (خودتوجهی چندسره) هر دو در زمره معماری های قدرتمند یادگیری عمیق هستند. آن ها در وظایف مختلف پردازش زبان طبیعی از جمله تشخیص URLهای فیشینگ عملکرد موفقی داشته اند. ترکیب CNN و MHSA می تواند با به‌کارگیری نقاط قوت هر دو روش، عملکرد تشخیصی را ارتقا دهد.

CNN یک نوع شبکه عصبی است که فیلترهای کانوالی را به داده های ورودی اعمال می‌کند و بیشتر برای تشخیص تصویر به‌کار می رود. در حوزه NLP، CNN می تواند ویژگی های مهم متن را با اعمال «پنجره لغزان» به توالی ورودی یاد بگیرد و در گام بعدی الگوهای محلی کلمات را استخراج می کند. این ویژگی های محلی ترکیب شده و به نمایش سطح بالاتری از متن ورودی تبدیل می شوند. کارایی CNN در تشخیص URLهای فیشینگ با استخراج ویژگی های n- گرم از URLها و استفاده از آن ها برای آموزش طبقه بند نشان داده شده است.

از سوی دیگر، MHSA یک مدل تبدیل محور است که با خودتوجهی مجموع وزنی توکن های ورودی را محاسبه می کند. لذا مدل می تواند وابستگی های سراسری و روابط گسترده (محدوده طولانی) بین کلمات و توالی را تشخیص دهد. کارایی MHSA در NLP برای

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

مدلسازی معنی ظاهری متن ثابت شده است. این کار با توجه به کلمات مناسب در توالی ورودی انجام می شود. می توان از MHSA در بحث تشخیص URLهای فیشینگ برای یادگیری «نمایشی از URL» استفاده کرد. این نمایش معنی ظاهری و زمینه را شناسایی می کند.

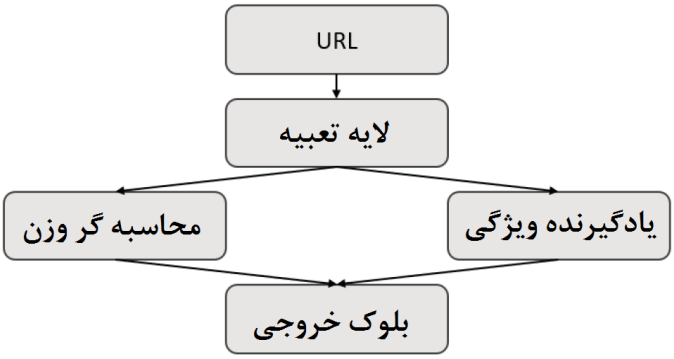
ترکیب CNN و MHSA باعث دستیابی به مزایای هر دو مدل می شود. همچنین با اتکا به توانایی های مکمل آن ها عملکرد تشخیصی را ارتقا می دهد. CNN از قدرت تشخیص الگوهای محلی و ویژگی های n- گرم URL برخوردار است، اما MHSA معنی ظاهری و زمینه را مدلسازی می کند. روش ترکیبی دو پیش بینی را ادغام کرده و صحت و مقاومت بیشتری در برابر URLهای فیشینگ مختلف ایجاد می کند.

نمای کُلی از مدل

به‌دلیل اینکه خودتوجهی چندسره (MHSA) از عملکرد عالی در پردازش زبان طبیعی (NLP) برخوردار است و از قدرت محاسبه وزن های ویژگی ها و تشخیص وابستگی ها بین کاراکترهای مختلف متن برخوردار است، می تواند در تحلیل URLها نیز مفید باشد و حتی بهتر از LSTM عمل کند. همچنین به خاطر کارایی CNNها در یادگیری خودکار ویژگی‌ها بدون مداخله انسان، ترکیب نمودن این تکنیک برای بهره‌مندی از نقاط قوت آن و بهبود قدرت تشخیص وبسایت های فیشینگ امکان پذیر است.

هنگامی که مدلی را برای ترکیب دو تکنیک یادگیری ماشین ایجاد می کنیم، می توان بخش هایی از آن را تفکیک کرد. به عنوان گام نخست، «لایه تعبیه» ای را تعریف می‌کنیم که URL ورودی را با کدگذاری وان هات به نمایش ماتریسی تبدیل می کند. علاوه بر این، به‌خاطر استفاده از دو تکنیک، می توان دو نسخه از ماتریس را برای یادگیری ویژگی و محاسبه وزن ویژگی به‌کار گرفت. یکی از نسخه ها در حین محاسبه وزن به لایه های MHSA اعمال شده تا وزن ها را محاسبه کند. اما به‌طور همزمان در فرایند استخراج ویژگی، نسخه بعدی از ماتریس URL برای یادگیری ویژگی ها به لایه های کانوالی اعمال می شود. بنابراین خروجی لایه قبلی به منزله ورودی لایه بعدی تلقی می‌شود. بعد از اتمام دو فرایند همزمان، دو بخش خروجی به بلوک خروجی خورانده می‌شوند تا نتیجه نهایی طبقه بندی محاسبه شود.

بلوک خروجی ابتدا ویژگی های اصلی و دو نسخه را به عنوان وزن های ویژگی ورودی دریافت می کند. خروجی به یک لایه تماماً مرتبط با تابع فعال‌سازی سیگموئید اعمال می‌شود و خروجی بین ۰ تا ۱ ایجاد می کند. اگر خروجی بیشتر از ۰/۵ باشد، URL ورودی قانونی فرض می شود و در غیر این صورت «فیشینگ» تلقی می شود.



شکل ۴-۱: نمای کلی از مدل

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

عملکرد مدل

با در نظرگرفتن مزایای ترکیب CNN و خودتوجهی چندسره، معماری مدل شامل ۳ مولفه اصلی خواهد بود: یک لایه تعبیه، یک یادگیرنده ویژگی و یک محاسبه‌گر وزن.

لایه تعبیه، رشته URL را به ماتریسی تبدیل می کند. تعداد سطرهای آن ماتریس برابر طول URL و دارای ۸۴ ستون است که ۸۴ کاراکتر بالقوه در URL را نشان می دهند. برای نمایش هر کاراکتر از کدگذاری وان هات استفاده می شود و ماتریس از طریق شبکه عصبی به ۶۴ ستون خلاصه می شود. URLهایی با طول های مختلف به‌وسیله رشته ای با طول ثابت، با تریمینگ یا پدینگ پردازش می شوند.

یادگیرنده ویژگی، ویژگی ها را از ماتریس خروجی لایه تعبیه خارج می کند. این کار با استفاده از یک لایه کانوالی، دو لایه باقیمانده و یک لایه تماماً متصل انجام می‌شود. لایه کانوالی حاوی ۵ کرنل مرسوم و لایه پولینگ حداکثری است. لایه های باقیمانده مسئله تجزیه اشباع صحت را حل می کنند. این کار با جمع کردن ورودی و خروجی لایه کانوالی انجام می شود. لایه تماماً متصل، قدرت بیان شبکه عصبی و کارایی استخراج ویژگی را افزایش می دهد.

محاسبه گر وزن حاوی یک لایه MHSA، دو لایه باقیمانده و یک لایه تماماً متصل است. این واحد مسئول محاسبه وزن های ویژگی ها است. خروجی لایه تعبیه با کدگذاری موقعیتی اعمال شده و حاوی اطلاعات موقعیت نسبی کاراکترها در توالی رشته URL است. ماتریس کدگذاری موقعیتی به‌وسیله توابع سینوس و کسینوس به‌دست می آید. سپس ماتریس خروجی به لایه MHSA که حاوی ۸ سر است اعمال می شود. در نهایت ماتریس ویژگی حاصل شده و برای طبقه بندی یا پیش بینی به‌کار می رود.

نتیجه بلوک خروجی یک مقدار بین ۰ الی ۱ است. هر چقدر مقدار خروجی بیشتر باشد (بزرگتر از ۰/۵)، احتمال فیشینگ بودن URL کمتر است.

شاخص های عملکردی مدل به‌طور کلی این‌گونه هستند:

- صحت: ۰,۹۸۳٤
- نرخ مثبت کاذب : ۰,۰۱۷٦
- دقت: ۰,۹۸٤٤
- بازیابی: ۰,۹۸۱٤
- امتیاز F1: ۰,۹۸۳۰

تحلیل معنی این شاخص های عملکردی:

1- صحت: صحت مدل ۰,۹۸۳٤ است و بدین معنی است که ۹۸,۳٤ % از URLهای فیشینگ مجموعه داده را به درستی تشخیص داده است.
2- FPR: ۰,۰۱۷٦ است ، بدین معنی که ۱,۷٦ % URLهای غیرفیشینگ به اشتباه، URL فیشینگ تلقی شده اند.
3- دقت: ۰,۹۸٤٤ است به این معنی که ۹۸,٤٤ % URLهای شناسایی شده به عنوان فیشینگ توسط مدل، واقعاً URL فیشینگ هستند.
4- بازیابی: ۰,۹۸۱٦ است، یعنی مدل به‌درستی ۹۸,۱٦ % تمام URLهای فیشینگ واقعی مجموعه داده را شناسایی کرده است.
5- امتیاز F1: ۰,۹۸۳۰ است، بدین معنی که مدل توازن خوبی بین دقت و بازیابی حاصل کرده است.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

این نتایج به‌طور کلی بدین معنی هستند که ترکیب CNN و MHSA عملکرد بسیار خوبی در تشخیص URLهای فیشینگ داشته است. علاوه بر این، صحت، دقت، بازیابی و F1 بالا و FPR پایین هستند.

تحلیل نتایج

جدول زیر با توجه به نتایج تحقیق ارائه شده است و عملکرد تکنیک های مختلف یادگیری ماشین برای تشخیص وبسایت های فیشینگ را مقایسه می کند.

جدول ۴-۲: مقایسه عملکرد با روش ترکیبی پیشنهادی

ردیف	الگوریتم تشخیص ML	صحت	FPR	بازیابی	دقت	F1
۱	بیز ساده [۱۹]	۹۷,۱۸				
۲	جنگل تصادفی [۲۱]	۹۷				
۳	CNN [۲۰]	۹۶,۶۱	۳,۵	۹۷,۰۹	۹۶,۶۱	۹۶,۸۵
۴	LSTM [۲۰]	۹۷,۲۰	۱,۸	۹۸,۶۳	۹۶,۴۵	۹۷,۵۳
۵	MLP [۲۰]	۹۶,۶۵		۹۶,۶۵	۹۶,۶۵	۹۶,۶۵
۶	CNN + RNN [۲۲]	۹۷,۹	۳,۱۰	۹۸,۳۹	۹۶,۷۶	۹۷,۵۷
۷	CNN + LSTM [۲۳]	۹۳,۲۸	۱,۸	۹۷,۱۳	۹۹,۱۲	۹۸,۱۱
	CNN + MHSA	۹۸,۳۴	۱,۷۶	۹۸,۴۴	۹۸,۱۶	۹۸,۳۰

تحلیل و مقایسه کامل تکنیک های مختلف یادگیری ماشین که برای تشخیص URL فیشینگ مبتنی به‌کاررفته بر اساس جدول ارائه شده اند:

1- بیز ساده [۱۹]: صحت آن ۹۷,۱۸ % است. یک الگوریتم ساده و محبوب است که با داده هایی با ابعاد زیاد با فرض مستقل بودن تمام ویژگی ها از هم به‌خوبی عمل می کند، اما شرایط اکثراً این‌گونه نیست.
2- جنگل تصادفی [۲۱]: صحت آن ۹۷ % است. این الگوریتم یک روش یادگیری ترکیبی است که درخت های تصمیم مختلف را ایجاد و پیش بینی هایشان را ترکیب می کند تا صحت را افزایش و بیش برازش را کاهش دهد.
3- CNN [۲۰]: صحت این الگوریتم ۹۶,۶۱ % است. FPR ۳,۵ %، بازیابی ۹۷ %، دقت ۹۶,۶۱% و F1 ۹۶,۸۵ % هستند. CNNها به وفور برای طبقه بندی تصویر کاربرد دارند، اما برای طبقه بندی متن نیز قابل استفاده هستند. FPR ۳,۵ % است و نشان می دهد ۳,۵ % URLهای قانونی به اشتباه URL فیشینگ تلقی شده اند.
4- LSTM [۲۰]: صحت آن ۹۷,۲۰ % ، FPR آن ۱,۸ %، بازیابی ۹۸,۶۳ % ، دقت ۹۶,۴۵ % و F1 ۹۷,۵۳ هستند. LSTMها یک نوع شبکه عصبی بازگشتی هستند که وابستگی های بلندمدت در داده های متوالی را شناسایی می کنند. FPR کم و بازیابی زیاد بیانگر کارایی مدل در تشخیص URLهای فیشینگ با مثبت های کاذب حداقلی هستند.
5- MLP [۲۰]: صحت آن ۹۶,۶۵ % است و هیچ شاخص دیگری در جدول مطرح نشده است. MLPها یک نوع شبکه عصبی فیدفوروارد بوده که می توانند روابط غیرخطی بین داده های ورودی و خروجی را یاد بگیرند.
6- CNN + RNN [۲۲]: صحت آن ۹۷,۹ % است، FPR ۳,۱ ، بازیابی ۹۸,۳۹ ، دقت ۹۶,۷۶ %، F1 ۹۵,۵۷ هستند. ترکیب شدن CNN با RNN باعث تشخیص ویژگی های مکانی و توالی داده های ورودی می شود، لذا عملکرد در مقایسه با استفاده یک جنبه ای از هر مدل بهبود می یابد.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

7- CNN + LSTM [۲۳]: صحت آن ۹۳,۲۸، FPR ۱,۸۰ %، بازیابی ۹۷,۱۳ % ، دقت ۹۹,۱۲ % ، F1 ۹۸,۱۱ هستند. ترکیب CNN با LSTM باعث تشخیص ویژگی های محلی و سراسری در داده های ورودی می شود و عملکرد را ارتقا می دهد. البته صحت و F1 کم نشان داده مدل همچون بقیه در این کار موثر نیست.

8- CNN+ MHSA: این الگوریتم دارای بالاترین صحت به میزان ۹۸,۳۴ % است، FPR آن کمترین به مقدار ۱,۷۶ % است. بازیابی ۹۸,۴۴%، دقت ۹۸,۱۶ % و F1 ۹۸,۳۰ هستند. CNN + MHSA ترکیب این دو روش است و وابستگی های بلندمدت در داده های ورودی را تشخیص می دهد. همچنین به‌طور همزمان به بخش های مختلف توالی توجه می کند. هرچند FPR با تکنیک پس پردازش، یادگیری فعال، ارزیابی مستمر و حلقه های بازخورد قابل بهبود است. این رویکرد عملکرد مدل را بهبود داده و توازن بهتری بین مثبت های کاذب و منفی های کاذب ایجاد می کند. این مدل در این کار بهترین گزینه است و صحت بالا، مثبت های کاذب کم و دقت و بازیابی زیاد از قابلیت هایش هستند.

به عنوان جمع بندی، باید گفت CNN + MHSA بهترین عملکرد را دارد، و LSTM، CNN+ RNN و جنگل تصادفی در رده های بعدی قرار می گیرند.

1. بحث و بررسی

۱-۵: سوال ۱- کدام نوع از الگوریتم های یادگیری ماشین برای تشخیص URLهای فیشینگ به‌کار رفته است و چطور می توان آن ها را آموزش داد و بهینه سازی کرد؟

برای مقابله با حملات فیشینگ، روش های مختلفی برای تشخیص URLهای فیشینگ توسعه داده شده اند. این روش ها شامل لیست سیاه، فیلترهای DNS، آموزش های آگاه کردن کاربر و الگوریتم های یادگیری ماشین هستند. هر روش نقاط ضعف و قوت خودش را دارد. ترکیب روش ها می تواند حفاظت موثری نسبت به این حملات فراهم کند.

لیست های سیاه و فیلترهای DNS به لیست های تهیه شده از URLها یا دامنه های مخرب بستگی دارند و با ایجاد دامنه ها و URLهای جدید توسط مهاجمان، به‌سرعت منسوخ می‌شوند. البته آن ها در مسدودسازی سایت های فیشینگ و جلوگیری از دسترسی کاربران موثر عمل می کنند. آموزش های آگاه سازی کاربران می توانند به تشخیص کلاهبرداری های فیشینگ کمک کنند، اما در برابر حملات پیچیده و شخصی روی قربانیان ناکارآمد هستند.

ما می توانیم از الگوریتم های یادگیری ماشین برای تشخیص URLهای فیشینگ استفاده کنیم. خصوصیات URLها از جمله نام دامنه، طول URL، وجود کلیدواژه های خاص تحلیل می شوند. این الگوریتم ها شباهت های بین URLهای فیشینگ و وبسایت های فیشینگ شناخته شده را تشخیص می دهند. یادگیری ماشین در زمینه تشخیص URLهای فیشینگ موثرتر از لیست سیاه یا فیلتر DNS است. دلیلش این است که با تهدیدهای جدید و تکاملی هماهنگ می شوند. مدل های یادگیری ماشین می توانند الگوها و ویژگی های URL ها و صفحات وب را تحلیل و حملات فیشینگ جدید و نامشخص را پیش بینی کنند، هرچند قبلاً رخ نداده باشند. همچنین این مدل ها از داده های گذشته نکاتی را یاد می گیرند و دقت شان را در طول زمان ارتقا می دهند، بنابراین در زمینه تشخیص این URLها موثرتر عمل می کنند.

یادگیری ترکیبی باعث ترکیب چندین مدل یادگیری ماشین می شود و عملکرد کلی را بهبود می بخشد.

سوال ۱-۱: کدام نوع از مجموعه‌های داده برای آموزش الگوریتم های یادگیری ماشین به‌کار می روند؟

چند نوع الگوریتم یادگیری ماشین موجود هستند که برای تشخیص URLهای فیشینگ به‌کار می روند. آن ها بر اساس مجموعه داده های منتخب عمل کرده و شامل یادگیری نظارت

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

شده، یادگیری غیرنظارت شده، یادگیری نیمه نظارت شده، یادگیری عمیق و یادگیری ترکیبی هستند. هر روش نقاط ضعف و قوت خودش را دارد. انتخاب الگوریتم ممکن است به نیازهای خاص یک سازمان و ماهیت حملات فیشینگ در حال تشخیص وابسته باشد.

یادگیری نظارت شده با آموزش دهی یک مدل یادگیری عمیق روی مجموعه داده برچسب گذاری شده همراه است. URLهای فیشینگ و قانونی در این مجموعه داده مستقر هستند. می توان از مدل برای طبقه بندی URLهای جدید به عنوان فیشینگ یا قانونی استفاده کرد. این کار با توجه به الگوهای یادگیری شده در طول آموزش انجام می شود. یادگیری نظارت شده می تواند در تشخیص حملات فیشینگ معلوم موثر باشد، اما در زمینه حملات فیشینگ ناشناخته یا جدید چنین تضمینی وجود ندارد.

در یادگیری غیرنظارت شده، یک مدل یادگیری ماشین روی مجموعه داده غیربرچسب گذاری شده از URLها آموزش داده می شود. بدین ترتیب الگوها و ناهنجاری های داده ای که نشانه وجود URLهای فیشینگ هستند مشخص می شوند. یادگیری غیرنظارت شده برای تشخیص حملات فیشینگ جدید و ناشناخته مفید است، اما ممکن است مثبت های کاذب را ایجاد کند.

یادگیری نیمه نظارت شده، مولفه های یادگیری نظارت شده و غیرنظارت شده را در هم ترکیب می کند. این مدل روی مجموعه داده کوچک برچسب گذاری شده از URLهای معتبر و فیشینگ آموزش داده می شود، اما یادگیری بر حسب مجموعه داده برچسب گذاری نشده انجام می شود تا الگوها و ناهنجاری های جدید در داده ها شناسایی شوند. یادگیری نیمه نظارتی می تواند در شناسایی حملات جدید و ناشناخته موثر باشد و مثبت های کاذب را به حداقل برساند.

روش های یادگیری عمیق همچون CNN یا RNN، برای تشخیص URLهای فیشینگ کاربرد دارند. آن ها ویژگی ها را مستقیماً از روی داده های خام ، همچون اسکرین شات های وبسایت یا لاگ های ترافیک شبکه یاد می گیرند. یادگیری عمیق می تواند در شناسایی حملات جدید و ناشناخته موثر باشد اما به حجم زیادی از داده های برچسب گذاری شده و منابع محاسباتی نیاز دارد.

هرچند روش های مختلفی برای آموزش داده ها بر حسب مجموعه داده وجود دارند، مورد بهکاررفته در اینجا به دو بخش یکسان تقسیم شده است: ۲۰۰۰ URL قانونی و ۲۰۰۰ URL فیشینگ. این تفکیک متعادل تضمین می کند که مدل به سمت یکی از دو دسته منحرف نشده است، و هر دو نوع URL را با دقت شناسایی می کند.

سوال ۱-۲: چه شاخص های دقتی (صحت) برای مقایسه الگوریتم ها بهکار می روند؟

این شاخص های دقت الگوریتم ها را در حوزه تشخیص فیشینگ با هم مقایسه می کنند:

1- صحت: نشان می دهد مدل تا چه اندازه URL را بهدرستی به عنوان فیشینگ یا قانونی طبقه بندی می کند؟ به صورت نسبت URLهای بهدرستی طبقه بندی شده به تعداد URLهای موجود در مجموعه تست تعریف می شود.
2- دقت: شاخصی است که نشان داده مدل به چه شکلی URLهای فیشینگ را بهدرستی شناسایی می کند. به عنوان نسبت تعداد مثبت های صحیح (URLهای فیشینگ که به درستی فیشینگ تلقی می شوند) به مجموع تعداد URLهای شناسایی شده به همین عنوان توسط مدل تعریف می شود.
3- بازیابی: شاخصی که نشان داده مدل تا چه اندازه تمام URLهای فیشینگ را درست شناسایی می کند. به صورت نسبت تعداد مثبت های صحیح به تعداد کل URLهای فیشینگ واقعی در مجموعه تست تعریف می شود.
4- امتیاز F1: شاخصی از عملکرد کلی مدل است و دقت و بازیابی را شامل می شود. به عنوان میانگین موزون دقت و بازیابی تعریف می شود.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

سوال ۱-۳: کدام الگوریتم های یادگیری ماشین بهترین نتایج را در تشخیص وبسایت های فیشینگ دارند؟

به‌دلیل اینکه ارزیابی کارایی الگوریتم های مختلف به چند شاخص از جمله صحت، FPR ، بازیابی، دقت و امتیاز F1 بستگی دارد، نتایج ارائه شده به این صورت هستند:

صحت روش بیز ساده زیاد و به اندازه ۹۷،۱۸ % است. صحت جنگل تصادفی ۹۷% است، صحت CNN ۹۶،۶۱ % است اما FPR بالا در حد ۳،۵ % دارد و این نشانه طبقه بندی برخی URLهای قانونی به عنوان فیشینگ است. LSTM بالاترین صحت را دارد (۹۷،۲۰%) و FPR اندکی دارد (۱،۸ %). یعنی اکثر URLهای فیشینگ را به‌درستی طبقه بندی کرده و تعداد کمی URL قانونی را به عنوان فیشینگ طبقه بندی می کند.

روش MLP دارای صحت ۹۶،۶۵ % است. روش CNN + RNN صحت ۹۷،۹ % دارد اما FRP نسبتاً زیادی دارد (۳،۱۰ %). روش CNN+ LSTM دارای کمترین صحت (۹۳،۲۸ %) است، اما بالاترین دقت به میزان ۹۹،۱۲ % را هم دارد. این یعنی تعداد بسیار کمی از URLهای قانونی را به عنوان فیشینگ طبقه بندی می کند.

به‌طور خلاصه، روش LSTM موثرترین روش از لحاظ صحت و FPR است، ولی CNN+ LSTM بیشترین دقت را دارد. به‌خاطر نقاط ضعف و قوت، ترکیب این دو روش بهتر از استفاده تکی است و باعث بهبود عملکرد کلی می شود.

برای مثال، روش CNN+ RNN صحت بیشتری نسبت به بیز ساده دارد، اما بیز ساده سریع تر است و به منابع محاسباتی کمتری نیاز دارد. ترکیب دو روش سیستم تشخیصی دقیق تر و موثرتری را به وجود می آورد.

۵-۲: سوال ۲- روش ترکیبی پیشنهادی تا چه اندازه در شناسایی URLهای فیشینگ موثر است؟

روش پیشنهادی ترکیبی (MHSA و CNN) عملکرد بهتری در تشخیص URLهای فیشینگ داشته است. ترکیب شدن باعث ارتقای عملکرد نسبت به CNN و LSTM می شود. نتایج جدول ۲ نشان داده که صحت و F1 روش پیشنهادی به ترتیب ۹۸،۳۴ و ۹۸،۳۰ % هستند. البته FPR روش پیشنهادی ۱،۷۶ % است و این یعنی تعداد بیشتری از صفحات قانونی را فیشینگ تلقی می کند.

علاوه بر این، زمان آموزش دهی روش پیشنهادی نسبتاً کم است و میانگین ۳۲ دقیقه در هر اجرا (دوره زمانی) را دارد.

به طور خلاصه، باید گفت روش ترکیبی دو تایی (خودتوجهی چندسره و CNN) عملکرد بهتری در تشخیص URLهای فیشینگ دارد. این روش صحت و امتیاز F1 بالایی دارد. زمان آموزش آن نسبتاً کم و FPR آن اندکی بیشتر است، لذا کارایی روش ترکیبی پیشنهادی در شناسایی URLهای پیشنهادی نسبتاً بالا است.

۵-۳: سوال ۳- روش ترکیبی پیشنهادی تا چه اندازه در شناسایی URLهای فیشینگ در مقایسه با سایر روش ها کارآمد است؟

با توجه به نتایج، روش پیشنهادی در مقایسه با بقیه روش ها عملکرد بسیار خوبی دارد. ساختارهای مختلف از جمله CNN، LSTM، CNN-CNN، CNN- LSTM با روش ترکیبی پیشنهادی (CNN و خودتوجهی چندسره) مقایسه شده اند.

این روش در تشخیص URLهای فیشینگ نسبت به بقیه روش ها بسیار موثر است. مدل پیشنهادی با ۵ روش پرکاربرد مقایسه شده و کمترین FPR به میزان ۰،۲۶ %، بیشترین صحت به میزان ۹۹،۸۴% و F1 به میزان ۹۹،۸۴ را حاصل می کند. همچنین تمام روش های قبلی را از لحاظ بازیابی با نرخ ۹۹،۹۵ % شکست می دهد. هرچند FPR روش پیشنهادی

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

بیشتر از CNN-LSTM (۰٫۸۲%) است، اما کماکان از بقیه روش ها کمتر است.نتیجه این است که ترکیب دو روش برای تشخیص URLهای فیشینگ در مقایسه با سایر روش ها به‌شدت موثر است.

اطلاعات زیر نشان می دهند که ترکیب دو شبکه به بهبود عملکرد مدل کمک می کند. زمان آموزش دهی روش پیشنهادی نیز از CNN-LSTM و سایر روش ها کمتر است، لذا می توان گفت که روش پیشنهادی (ترکیبی) کارایی بالاتری دارد.

نتیجه گیری و پیشنهادها

حملات فیشینگ یکی از تهدیدهای اساسی نسبت به امنیت آنلاین هستند. روش های مختلفی برای تشخیص و پیشگیری از آن ها پیشنهاد شده اند. الگوریتم های یادگیری ماشین به عنوان یک رویکرد امیدوارکننده برای شناسایی URLهای فیشینگ معرفی شده اند، زیرا توانایی یادگیری از داده ها و تطبیق پذیری با تهدیدهای جدید و توسعه یافته را دارند. ما در این مطالعه انواع الگوریتم های به‌کاررفته برای تشخیص URLهای فیشینگ، مجموعه داده های به‌کاررفته برای آموزش، شاخص های صحت ارزیابی عملکرد را بررسی کرده ایم.

نتایج، بیانگر این هستند که الگوریتم های مختلف یادگیری ماشین، دارای نقاط ضعف و قوت مختلفی در تشخیص URLهای فیشینگ هستند. یادگیری نظارت شده روشی موثر برای تشخیص حملات شناخته شده است، اما یادگیری غیرنظارت شده می تواند برای شناسایی حملات جدید و مجهول به‌کار رود. یادگیری نیمه نظارتی، توازنی بین این دو روش تلقی می شود. یادگیری عمیق قدرت یادگیری مستقیم ویژگی ها از روی داده های خام را دارد و برای تشخیص حملات جدید و ناشناخته به‌کار می‌رود اما به انبوهی از داده های برچسب گذاری شده و منابع محاسباتی نیاز دارد.

از لحاظ شاخص های صحت، می توان گفت بیز ساده و جنگل تصادفی، دارای صحت های بالاتر به میزان ۹۷٫۱۸ % و ۹۷ % هستند. صحت روش CNN اندکی کمتر و ۹۶٫۶۱ % است، اما FPR بیشتری به اندازه ۳٫۵ % دارد. این یعنی تعداد بیشتری از URLهای قانونی به عنوان URL فیشینگ دسته بندی شده اند. LSTM دارای بالاترین صحت به میزان ۹۷٫۲۰ % است و نرخ مثبت کاذب پایین به اندازه ۱٫۸ % دارد، لذا اکثر URLهای فیشینگ را به‌درستی تشخیص داده و فقط چند مورد به اشتباه قانونی تلقی می شوند.

با توجه به این تحقیق، کارآمدترین روش تشخیص، یک روش ترکیبی شامل CNN و خودتوجهی چندسره است. دلیلش این است که این روش بهترین عملکرد را داشته و صحت آن ۹۸٫۳۴ % است. FPR آن کمترین مقدار (۱٫۷۶%). بازیابی ۹۸٫۴۴، دقت ۹۸٫۱۶ و F1 ۹۸٫۳ هستند. این مدل CNN را با MHSA ترکیب کرده و به تشخیص وابستگی های طولانی مدت و تحلیل همزمان چند بخش کمک می کند. به‌طور کلی بهتر از سایر مدل ها از جمله LSTM، CNN + RNN و جنگل تصادفی عمل می کند.

به‌طور کلی، این مطالعه پتانسیل(ظرفیت) الگوریتم های یادگیری ماشین برای تشخیص URLهای فیشینگ را بررسی کرده است. همچنین ترکیب کردن روش ها از جمله لیست های سیاه، فیلتر DNS، آموزش آگاهی به کاربران و الگوریتم های یادگیری ماشین را پیشنهاد می‌کند که می توانند حفاظت لازم در برابر حمله را فراهم کنند. از سوی دیگر، انتخاب الگوریتم به نیازهای خاص سازمان و ماهیت حمله فیشینگ بستگی دارد. برای بررسی کارایی سایر روش ها و ظرفیت شان در شرایط واقعی به تحقیقات بیشتری نیاز داریم.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

منابع

[1] James, L. (2006). Banking on phishing. In James, L. (Ed.), Phishing Exposed (pp. 1-35). Syngress. ISBN 9781597490306

[2] Sundara Pandiyan, S., Selvaraj, P., Burugari, V. K., Benadit P, J., & Kanmani, P. (2022). Phishing attack detection using Machine Learning. Measurement: Sensors, 24, 100476. ISSN 2665-9174

[3] Ahammad, S. K. H., Kale, S. D., Upadhye, G. D., Pande, S. D., Babu, E. V., Dhumane, A. V., & Bahadur, M. D. K. J. (2022). Phishing URL detection using machine learning methods. Advances in Engineering Software, 173, 103288. ISSN 0965-9978

[4] Berners-Lee, T., Masinter, L., & McCahill, M. (Eds.). (1994). Uniform Resource Locators (URL). Request for Comments: 1738. Network Working Group. CERN. Standards Track. Updated by: 1808, 2368, 2396, 3986, 6196, 6270, 8089. Obsoleted by: 4248, 4266. Errata Exist

[5] L. Wenyin, G. Liu, B. Qiu and X. Quan, "Antiphishing through Phishing Target Discovery," in IEEE Internet Computing, vol. 16, no. 2, pp. 52-61, March- April 2012, doi: 10.1109/MIC.2011.103

[6] Safi, A., & Singh, S. (2023). A systematic literature review on phishing website detection techniques. Journal of King Saud University - Computer and Information Sciences, 35(2), 590-611. ISSN 1319-1578

[7] Vrbančič, G., Fister, I., & Podgorelec, V. (2020). Datasets for phishing websites detection. Data in Brief, 33, 106438. ISSN 2352-3409

[8] Zheng, F., Yan, Q., Leung, V. C. M., Yu, F. R., & Ming, Z. (2022). HDP-CNN: Highway deep pyramid convolution neural network combining wordlevel and character-level representations for phishing website detection. Computers & Security, 114, 102584. ISSN 0167-4048

[9] Wei, W., Ke, Q., Nowak, J., Korytkowski, M., Scherer, R., & Woźniak, M. (2020). Accurate and fast URL phishing detector: A convolutional neural network approach. Computer Networks, 178, 107275. ISSN 1389-1286

[10] Sahingoz, O. K., Baykal, S. I., & Bulut, D. (2018). Phishing detection from urls by using neural networks. Computer Science & Information Technology (CS & IT), 41-54.

[11] Remmide, M. A., Boumahdi, F., Boustia, N., Feknous, C. L., & Della, R. (2022). Detection of Phishing URLs Using Temporal Convolutional Network. Procedia Computer Science, 212, 74-82. ISSN 1877-0509.

[12] Marwa M. Emam, Nagwan Abdel Samee, Mona M. Jamjoom, Essam H. Houssein, Optimized deep learning architecture for brain tumor classification using improved Hunger Games Search Algorithm, Computers in Biology and Medicine, Volume 160, 2023, 106966, ISSN 0010-4825

[13] Sundara Pandiyan S, Prabha Selvaraj, Vijay Kumar Burugari, Julian Benadit P, Kanmani P, Phishing attack detection using Machine Learning, Measurement: Sensors, Volume 24, 2022, 100476, ISSN 2665-9174,

[14] Kai Florian Tschakert, Sudsanguan Ngamsuriyaroj, Effectiveness of and user preferences for security awareness training methodologies, Heliyon, Volume 5, Issue 6, 2019, e02010, ISSN 2405-8440

[15] Mohsen Soori, Behrooz Arezoo, Roza Dastres, Machine learning and artificial intelligence in CNC machine tools, A review, Sustainable Manufacturing and Service Economics, 2023, 100009, ISSN 2667-3444,

[16] Tianyuan Liu, Hangbin Zheng, Pai Zheng, Jinsong Bao, Junliang Wang, Xiaojia Liu, Changqi Yang, An expert knowledge-empowered CNN approach for welding radiographic image recognition, Advanced Engineering Informatics, Volume 56, 2023, 101963, ISSN 1474-0346,

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

[17] Jun Ma, Guolin Yu, Weizhi Xiong, Xiaolong Zhu, Safe semisupervised learning for pattern classification, Engineering Applications of Artificial Intelligence, Volume 121, 2023, 106021, ISSN 0952-1976

[18] Benavides-Astudillo, E., Fuertes, W., Sanchez-Gordon, S., Rodriguez- Galan, G., Martínez-Cepeda, V., Nuñez-Agurto, D. (2023). Comparative Study of Deep Learning Algorithms in the Detection of Phishing Attacks Based on HTML and Text Obtained from Web Pages. In: Botto-Tobar, M., Zambrano Vizuete, M., Montes León, S., Torres-Carrión, P., Durakovic, B. (eds) Applied Technologies. ICAT 2022. Communications in Computer and Information Science, vol 1755. Springer, Cham. https://doi.org/10.1007/978-3-031-24985-3_28

[19] J. Kumar, A. Santhanavijayan, B. Janet, B. Rajendran and B. S. Bindhumadhava, "Phishing Website Classification and Detection Using Machine Learning," 2020 International Conference on Computer Communication and Informatics (ICCCI), Coimbatore, India, 2020, pp. 1-6, doi: 10.1109/ICCCI48352.2020.9104161.

[20] Do, Q.N.; Selamat, A.; Krejcar, O.; Yokoi, T.; Fujita, H. Phishing Webpage Classification via Deep Learning-Based Algorithms: An Empirical Study. Appl. Sci. 2021, 11, 9210. https://doi.org/10.3390/ app11199210

[21] M. N. Alam, D. Sarma, F. F. Lima, I. Saha, R. -E. -. Ulfath and S. Hossain, "Phishing Attacks Detection using Machine Learning Approach," 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2020, pp. 1173-1179, doi: 10.1109/ICSSIT48917.2020.9214225.

[22] Y. Huang, Q. Yang, J. Qin and W. Wen, "Phishing URL Detection via CNN and Attention-Based Hierarchical RNN," 2019 18th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/13th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE), Rotorua, New Zealand, 2019, pp. 112-119, doi: 10.1109/TrustCom/BigDataSE.2019.00024.

[23] M. A. Adebowale, K. T. Lwin and M. A. Hossain, "Deep Learning with Convolutional Neural Network and Long Short-Term Memory for Phishing Detection," 2019 13th International Conference on Software, Knowledge, Information Management and Applications (SKIMA), Island of Ulkulhas, Maldives, 2019, pp. 1-8, doi: 10.1109/SKIMA47702.2019.8982427.

[24] Bahnsen, A. C., Bohorquez, C. E., Villegas, S., Vargas, J., & González, F. A. (2017). Classifying phishing URLs using recurrent neural networks. In 2017 APWG symposium on electronic crime research (eCrime) (pp. 1–8). Scottsdale, AZ, USA.

[25] Bahnsen, A. C., Bohorquez, C. E., Villegas, S., Vargas, J., & González, F. A. (2017). Classifying phishing URLs using recurrent neural networks. In 2017 APWG symposium on electronic crime research (eCrime) (pp. 1–8). Scottsdale, AZ, USA.

[26] Zhang J., Li X. Phishing detection method based on borderline-smote deep belief network security, privacy, and anonymity in computation, communication, and storage. SpaCCS 2017, Lecture notes in computer science, vol. 10658, Springer, Cham (2017), pp. 45-53

[27] Yang P., Zhao G., Zeng P. Phishing website detection based on multidimensional features driven by deep learning IEEE Access, 7 (2019), pp. 15196-15209

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

ارائه روش مقابله با حمله DDOS درشبکه بی‌سیم پهن باند چند رسانه ای

*** مهدی قهرمانی[1]، محمد مهدی شیر محمدی [2]**

گروه کامپیوتر،واحد همدان،دانشگاه آزاد اسلامی،همدان،ایران [1]mehdighahremani934@gmail.com

گروه کامپیوتر،واحد همدان،دانشگاه آزاد اسلامی،همدان،ایران [2]Mmshirmohammadi@gmail.com

چکیده

امروزه بسیاری از موسسات و مراکز آموزشی و سازمانها و ادارات موفق جهت معرفی خود و یا با هدف افزایش بهره وری کاری خود، از پیشرفتهای تکنولوژیکی برای ارائه محصولات جدید استفاده می‌کنند. یکی از مهمترین نمودها و محصولات این پیشرفتها، نرم افزارهای چندرسانه‌ای یا مالتی مدیا هستند که روزبه‌روز استفاده از آنها در زمینه های مختلف مانند آموزش و تجارت افزایش می یابد. با رشد سریع برنامه های چندرسانه‌ای از طریق اینترنت، حفظ کیفیت خدمات (QoS) ضروری شده است. ارائه تضمین خدمات معتبر از طریق اینترنت، بزرگترین چالش فعلی برای خدمات مبتنی بر پروتکل اینترنتی است. استفاده از ترافیک چندرسانه ای در ارتباط با رسانه های جریانی مانند: ویدئوکنفرانس با استفاده از شبیه ساز آپنت افزایش یافته است. در این مقاله، عملکرد سیستم های ویدئوکنفرانس را می‌توان بر اساس سناریوهای متفاوت مدل‌سازی کرد. این سناریوها: شامل شرایط سنگین و سبک و بررسی میزان بار و تاخیر شبکه است، همچنین امکان بررسی حملات DDOS که یک نوع حمله سایبری است که به دنبال از کار انداختن و مختل کردن سرویس‌های آنلاین و منابع شبکه‌ای مختلف است. در این حملات، مهاجم از تعداد زیادی دستگاه کامپیوتری (که به طور معمول توسط بدافزار تحت کنترل او درآمده‌اند) برای ارسال درخواست‌های تقلبی و بی‌فایده به سیستم مورد هدف استفاده می‌کند.هدف اصلی چنین حملاتی، اشباع کردن منابع شبکه و سیستمی مانند پهنای باند، پردازنده و حافظه است تا به این ترتیب سرویس‌های آن سیستم را از دسترس خارج و غیرفعال سازد. این موضوع باعث افزایش چشمگیر زمان پاسخ‌دهی و در نهایت فروپاشی کامل سرویس مورد هدف می‌شود.

کلمات کلیدی: بار زیاد، بارکم، تاخیر، حمله DDOS، شبکه محلی بی سیم(WLAN)

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

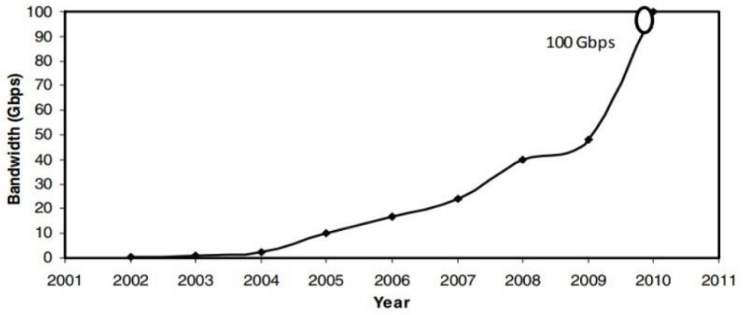
Print ISSN: 2251-9335
Online ISSN: 2252-0279

۱-مقدمه

افزایش تقاضا برای انتقال داده موجب پیشرفت فناوری‌های شبکه‌های کامپیوتری مانند LAN، VLAN و WLAN شده است تا برنامه‌های کاربردی مفیدی را برای مشتریان فراهم کنند. VLAN به مهندسان شبکه اجازه می‌دهد تا به جای شبکه‌های فیزیکی، شبکه‌های منطقی را طراحی کنند و برای جداسازی شبکه به چندین دامنه پخش بدون مشکل تأخیر استفاده شود.[۱].پیشرفت فناوری‌های الکترونیکی موجب افزایش اهمیت ارتباطات بی‌سیم، به‌ویژه کاربردهای چندرسانه‌ای شده است. در شبکه‌های چندخدماتی، تخصیص مناسب و قابل اعتماد پهنای باند برای ارائه کیفیت خدمات تضمین‌شده، یکی از مهم‌ترین مسائل است. [۲]

در دهه اخیر، ارتباطات با استفاده از ابزارهای فناورانه محبوب شده است و بهبود برآورده کردن انتظارات مشتریان را به همراه دارد. ابتکارات جدید در شبکه های سلولی به منظور پاسخگویی به نیازهای کاربران در حال توسعه هستند و تنوع خدمات در دسترس است. بهبود کیفیت خدمات (QoS) می‌تواند برای ارتقاء کیفیت خدمات و رضایت مشتریان مهم باشد. فناوری اینترنت اخیراً روش‌های ارتباطی کاربران را تغییر و استفاده از کاربردهای IP را افزایش داده است. توسعه دسترسی پرسرعت IP و شبکه‌های مبتنی بر آن، توجه زیادی به خدمات تلفن اینترنتی (VoIP) جلب کرده است. برای جذابیت بیشتر VoIP، کیفیت آن باید مشابه با تلفن سنتی باشد.[۳].دسترسی عمومی به جریان‌های چند رسانه‌ای، اکنون اصلی‌ترین انگیزه در طراحی شبکه‌های کامپیوتری و ارتباطی نسل بعدی است. محصولات در حال توسعه هستند تا قابلیت‌های ترافیک چند رسانه‌ای را در تمام اتصالات شبکه افزایش دهند. این تغییر از شبکه تلفن آنالوگ به شبکه‌ای را نشان می‌دهد که از پروتکل داده‌ای اینترنت استفاده می‌کند. تکامل سریع این شبکه‌ها باعث افزایش انتظارات عمومی و فرصت‌های کارآفرینی شده است. پژوهشگران و تولیدکنندگان علاقه‌مند به انتقال جریان‌های چند رسانه‌ای در شبکه‌ها هستند و ما نیاز داریم تا شبکه بتواند همزمان چندین رسانه را پشتیبانی کند. دو نکته مهم در اینجا وجود دارد؛ اول اینکه رسانه‌ها به شکل دیجیتال نمایش داده می‌شوند و شبکه‌های ارتباطات دیجیتال از آنها استفاده می‌کنند. دوم اینکه پایانه‌های کاربر نیز تأثیر زیادی بر ارتباطات چند رسانه‌ای و دسترسی به آنچه در دسترس است، دارند.[۴].

حملات DDoS به تهدیدی رایج برای کسب‌وکارهای آنلاین تبدیل شده‌اند. این حملات به شکل مرئی و پُرهزینه‌ای از جرم‌های سایبری درآمده‌اند و کسب‌وکارهای آنلاین به منظور اجتناب از هزینه‌های ویرانگر تعطیلی مرتبط با DDoS، به صورت پیشگیرانه در این زمینه فعالیت می‌کنند. مطالعات نشان می‌دهد که میزان حملات DDoS افزایش یافته و ترافیک آنها نیز درحال افزایش است. باگذشت زمان، حملات DDoS تکامل یافته‌اند و مهاجمان چندین ماشین آسیب‌پذیر را بر روی قربانی هماهنگ می‌کنند تا اثر انکار سرویس را بیشتر کنند. شکل۱افزایش حملات DDOS را نشان می دهد[۵].



شکل ۱: افزایش حملات DDOS

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

۲-کارهای مرتبط

در این مقاله[۶] ما درباره مسائل امنیتی مهم و تهدیدات امنیتی مختلف، به‌ویژه ریسک‌های فعال و غیرفعال، در شبکه‌های بی‌سیم پهن‌باند بحث کردیم. حملات سرویس‌انکاری (DoS) شکل شدید حمله فعال به همه انواع شبکه‌های بی‌سیم، به‌ویژه شبکه‌های پهن‌باند هستند و قادرند یک گره‌تکی، قسمتی از شبکه بی‌سیم، کل شبکه بی‌سیم یا منابع شبکه بی‌سیم را هدف قرار دهند. حملات DoS می‌توانند دو ویژگی مهم شبکه‌های بی‌سیم امن، یعنی صحت داده و در دسترسی سرویس را تهدید کنند. شبکه‌های متناسب با ۸۰۲,۱۱ و ۸۰۲,۱۶ در مقابل انواع مختلف حملات DoS، نسبت به WMN به دلیل معماری چند قفسه، پوشش مساحت گسترده و اتصال کاربران ادهاک، آسیب‌پذیرتر هستند. در لایه فیزیکی، تمامی سه شبکه پهن‌باند به‌اندازه یکدیگر در برابر تهدیدات DoS آسیب‌پذیر هستند، در حالی که در لایه پیوند، ۸۰۲/۱۶ نسبت به بقیه نسبتاً امن‌تر است. WMN به دلیل هزینه‌های چند قفسه از نظر لایه شبکه به حملات DoS آسیب‌پذیرتر است نسبت به ۸۰۲,۱۱ و ۸۰۲,۱۶ که هزینه‌های مسیریابی کمتری دارند. با توجه به اهمیت WMN، نیاز به تلاش‌هایی در جهت مبارزه با حملات DoS وجود دارد. طراحی و پیاده‌سازی رادیوی شناور، مکانیزم‌های رمزنگاری و احراز هویت بهبود یافته و مکانیزم‌های تشخیص تجاوز انحصاری، اقدامات مقابله ممکن هستند که باید مورد بررسی قرار گیرند. نیاز به تحقیقات بیشتر در زمینه امنیت شبکه‌های بی‌سیم پهن‌باند وجود دارد. فقط یک شبکه بی‌سیم پهن‌باند امن مورد پذیرش بالا برای استقرار تجاری در سطح گسترده خواهد بود.

- آنتونی تانوری و همکاران از نقشه تفاوت (disparity map) برای جمع‌آوری دید گسترده‌تری از داده‌های چندرسانه‌ای استفاده کردند. این نقشه با استفاده از تجزیه و تحلیل عمق ۳ بعدی، اطلاعات مربوط به عمق اشیا را فراهم می‌کند.

- این روش باعث کاهش زمان محاسبات و ارائه راه‌حل برای پردازش داده‌ها در زمان واقعی می‌شود. حسگرها با مطابقت دادن تصاویر استریو، نقشه تفاوت را ایجاد می‌کنند که باعث کاهش تأثیر ترافیک بر پهنای باند و افزایش طول عمر شبکه حسگر چندرسانه‌ای بی‌سیم (WMSN) می‌شود[۷].

- احمد متین و همکاران در مقاله خود[۸] به مقایسه شبکه‌های حسگر بی‌سیم (WSN) و شبکه‌های حسگر چندرسانه‌ای بی‌سیم (WMSN) پرداخته‌اند. WSN به طور کلی ارزان، مصرف انرژی کم و تعداد زیادی حسگر همراه با ایستگاه‌های پایه دارد. در مقابل، WMSN برای پردازش جریان‌های ویدئویی، صوتی و داده‌های حسگر اسکالار طراحی شده است.

- وائل علی حسین و همکاران در مقاله [۹] طراحی و تحلیل عملکرد پروتکل مسیریابی قابل اطمینان-مطمئن را برای شبکه‌های حسگر چندرسانه‌ای بی‌سیم متحرک انجام داده‌اند. آنها یک پروتکل مسیریابی جدید را طراحی کرده‌اند که بر اساس مسیریابی چند-مسیری با عبور از بهترین گره از نظر نرخ انتقال داده و نزدیکی به مقصد است. این پروتکل را GFTEM Greedy Forwarding with Throughput and Energy Metric نامیده‌اند. عملکرد را با سایر پروتکل‌های مسیریابی موجود مانند: AODV ، DYMO و پروتکل مسیریابی stateless درمحیط

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

شبکه‌های Wi-Fi مقایسه کرده‌اند. نتایج نشان می‌دهد که GFTEM در مقایسه با سایر پروتکل‌ها، تأخیر پایان-به-پایان کمتر، نرخ از دست رفتن بسته‌های کمتر و کارایی انرژی بهتری دارد.

-ناصم عباس و فنگکی یو در مقاله [۱۰] یک الگوریتم کنترل ترافیک ازدحام (TCCA) را برای شبکه‌های حسگر چندرسانه‌ای پیشنهاد داده‌اند: TCCAاز ترکیب دو شاخص ازدحام برای تشخیص ازدحام استفاده می‌کند: اشغال بافر و نرخ تغییر اشغال بافر همچنین یک کنترل کننده نرخ همراه با بازخورد را توسعه داده‌اند تا کیفیت جریان ویدیویی را بهبود بخشند. مکانیزم‌های مختلفی برای بازانتقال بسته‌های از دست رفته ارائه شده است، به طوری که بسته‌های گم شده موقتاً ذخیره و زمانی که ازدحام برطرف شد، بازانتقال می‌شوند. این طرح با استفاده از ۱۴ مبدل Pi مورد آزمایش قرار گرفته است. نتایج نشان می‌دهد که TCCA با کنترل نرخ و کاهش از دست رفتن بسته‌ها، عملکرد بهتری نسبت به روش‌های متداول دارد.

در مقاله [۱۱] مهم‌ترین جنبه چالش‌برانگیز در انتخاب سرور چند رسانه‌ای پویا، نرخ بیت تطبیقی هر جریان چند رسانه‌ای است که با وضعیت شبکه تغییر می‌کند. بنابراین، در این مقاله، ما بر روش تخمین پهنای باند موجود یک پیوند بی‌سیم تمرکز کرده‌ایم.

این روش باید ویژگی‌های زیر را داشته باشد:

(الف) قابل اجرا در برنامه‌های کاربردی بر روی زمان واقعی مانند خدمات پخش چند رسانه‌ای باشد.

(ب) ساده و مؤثر در تخمین پهنای باند موجود باشد.

(ج) بار مصرفی پایینی داشته باشد.

این مقاله [۱۲]چارچوب QoS IEEE 802.16e، IEEE 802.16m و LTE را توضیح می‌دهد و ویژگی‌های QoS آنها را با یکدیگر مقایسه می‌کند.ارائه QoS مورد نیاز برای تحویل تجربه کاربری خوب در اینترنت موبایل بسیار مهم است. مفهوم QoS حتی اهمیت بیشتری پیدا می‌کند، زیرا قابلیت‌های دستگاه‌ها تمایل مصرف‌کنندگان برای استفاده از محتوای رسانه‌ای غنی‌تر مانند ویدئو را نشان داده است. فناوری‌های بی‌سیم نسل چهارم مانند IEEE 802.16e، IEEE 802.16m و LTE برای پشتیبانی از نیازهای QoS کنونی و آینده طراحی شده‌اند. پشتیبانی QoS مبتنی بر جریان-محور و یک‌طرفه در IEEE 802.16e، به انواع مختلف جریان سرویس مانند: UGS و BE امکان ارائه ترافیک زمان‌واقعی و غیرزمان‌واقعی را می‌دهد. مکانیزم درخواست و اعطای پهنای باند صعودی به MSها امکان درخواست و دریافت منابع مورد نیاز برای انتقال داده در جهت صعودی را می‌دهد. ویژگی‌های پیشرفته مانند یک سرویس برنامه‌ریزی جدید، دسترسی سریع و درخواست پهنای باند به تأخیر انداخته شده در IEEE 802.16m، قابلیت‌های ارائه QoS مورد نیاز برای برنامه‌های کاربردی اینترنت موبایل نسل بعدی را بیشتر بهبود می‌دهد. مکانیزم‌های QoS LTE از کنترل QoS آغازشده توسط شبکه بر اساس GBR و غیر-GBR بیننده‌ها پیروی می‌کنندکه یک رفتار انتقال بسته بر اساس کلاس برای ارائه ترافیک زمان‌واقعی و غیرزمان‌واقعی است.

۳-مکانیزم RTS/CTS

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

در شبکه های بی‌سیم، ایستگاه پایه برای انتقال آماده است و قبل از انتقال هر فریم داده، یک فریم کوتاه درخواست برای ارسال (RTS) می‌فرستد. تأثیر یک فریم RTS کمتر از تأثیر فریم داده واقعی است زیرا اختلاف اندازه آنها کمتر است. وقتی ایستگاه پایه در گیرنده برای دریافت آماده است، فریم RTS با یک فریم ارسال "مجاز برای ارسال" (CTS) برای فرستنده تأیید می‌شود و بنابراین همه ترافیک از ایستگاه دیگر مسدود می‌شود. علاوه بر این، اگر فرستنده، فریم CTS را دریافت کند، یک فریم داده در صورتی که کانال برای طول انتقال کامل رزرو شده باشد، ارسال می‌شود. در نهایت، فریم تأیید (ACK) توسط گیرنده به فرستنده بر اساس دریافت فریم ارسال می‌شود. بنابراین، ارزیابی کارایی

قابل توجه مکانیزم اختیاری مصافحه RTS/CTS بر عملکرد شبکه های محلی بی‌سیم مبتنی بر IEEE 802.11 اهمیت دارد.کیفیت سرویس (QoS) در شبکه‌های ارتباطی به مجموعه‌ای از ویژگی‌ها اشاره دارد که بیانگر کیفیت و قابلیت اطمینان سرویس ارائه شده توسط آن شبکه است.

برخی از این ویژگی‌ها عبارتند از:

۱. تأخیر (Delay): زمان لازم برای انتقال بسته‌ها از منبع به مقصد برای برنامه‌های زنده مانند صوت و تصویر، تأخیر کم اهمیت دارد.

۲. نوسان تأخیر (Jitter): تغییرات در تأخیر بسته‌ها که می‌تواند باعث قطع و وصل در ارتباطات شود.

۳. نرخ از دست رفتن بسته‌ها (Packet Loss Rate): درصد بسته‌هایی که در طول مسیر از بین می‌روند.

۴. باندای عبوری (Throughput): حداکثر میزان داده‌ای که شبکه می‌تواند با کیفیت مطلوب انتقال دهد.

سرویس‌دهی با QoS مناسب به معنای تضمین این پارامترها در حد مطلوب برای کاربردهای مختلف است. این امر نیازمند مکانیزم‌هایی برای مدیریت ترافیک و منابع شبکه است که متناسب با نیازهای کاربردی تنظیم می‌شود.

4-پیاده‌سازی و شبیه‌سازی OPNET

4-1.شبیه ساز OPNET

OPNET (Optimized Network Engineering Tool) شرکتی است که در سال ۱۹۸۶ در انستیتوی فناوری ماساچوست (MIT) تأسیس شد. یک سال بعد، در سال ۱۹۸۷، اولین نرم‌افزار شبیه‌سازی عملکرد شبکه‌های تجاری توسط این شرکت منتشر شد که می‌توانست ابزار بهینه‌سازی عملکردهای مهم شبکه را فراهم کند و یک انقلاب در شبیه‌سازی شبکه ایجاد کرد. ایجاد مدیریت عملکردهای تحلیلی شبکه با شبیه‌سازی امکان‌پذیر شد.

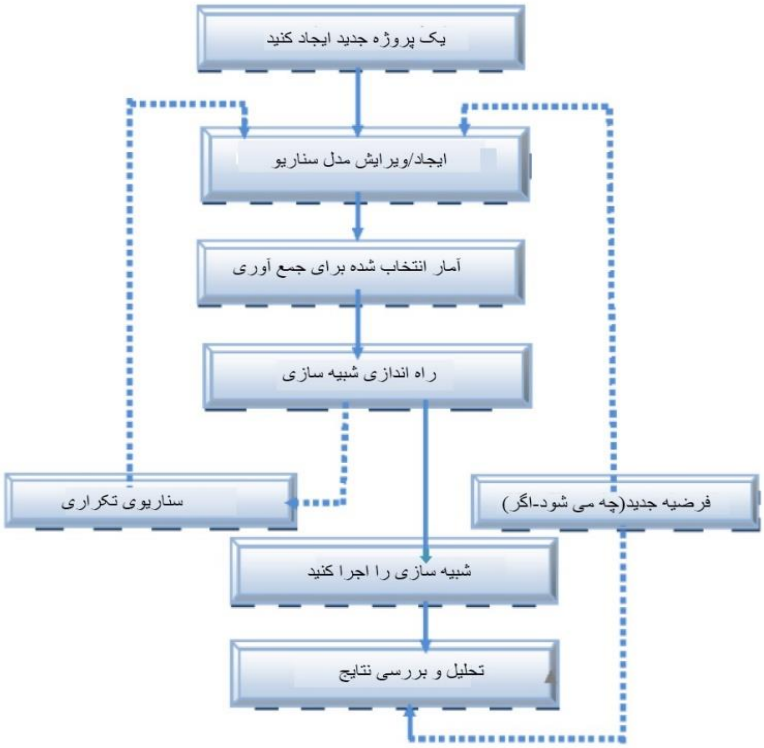
علاوه بر مدلر OPNET، توسعه محصولات دیگری مانند: OPNET Development Kit و WDM Guru نیز انجام شده است. شبیه‌سازی به عنوان یک روش فزاینده محبوب برای مطالعه عملکردها و کارکردهای مدل‌های پیشنهادی در سناریوهای مختلف در نظر گرفته می‌شود. شبیه‌سازی یک رویه آزمایشی از یک نمونه طراحی شده در یک پلتفرم است که محیط واقعی را تقلید می‌کند و فرصتی برای مطالعه، ایجاد و اصلاح عملکرد طرح پیشنهاد شده با هدف تقویت و ضعف انتظارات قبل از پیاده‌سازی مدل در محیط واقعی فراهم می‌کند.

۵-مدل پیشنهادی

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

شبیه‌سازی هر سیستمی با مراحلی که در فرآیند شبیه‌سازی صورت می‌گیرد، آغاز می‌شود. برای هر سناریو، ترافیک شبکه و پیکربندی‌ها اصلاح می‌شوند و شبیه‌سازی اجرا می‌شود. نمودار جریان برای نشان دادن مراحل کلیدی انجام شده برای ارزیابی عملکرد یک شبکه محلی بی‌سیم در شکل ۲ ارائه شده است.



شکل ۲: نمودار جریان برای ارزیابی عملکرد WLAN

این نمودار جریان مراحل اصلی را به شرح زیر نشان می‌دهد:

۱. تعریف سناریوهای شبیه‌سازی

۲. پیکربندی پارامترهای شبکه

۳. اجرای شبیه‌سازی

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

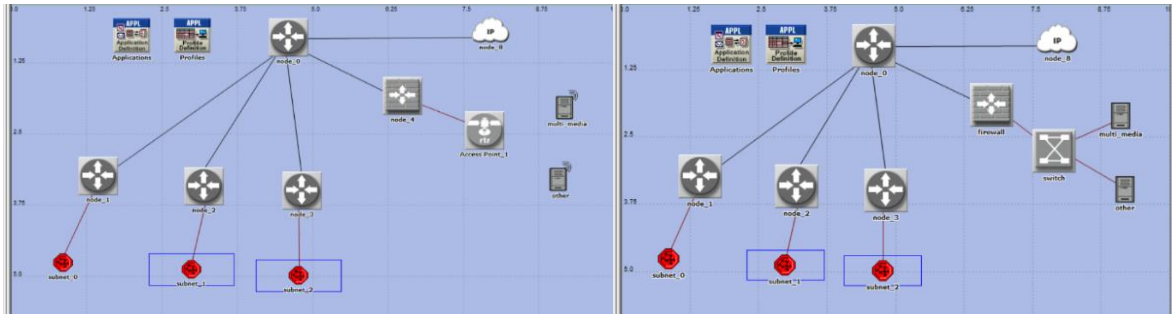
Print ISSN: 2251-9335
Online ISSN: 2252-0279

۴. جمع‌آوری و تحلیل نتایج

۱-۵. سناریوی پایه

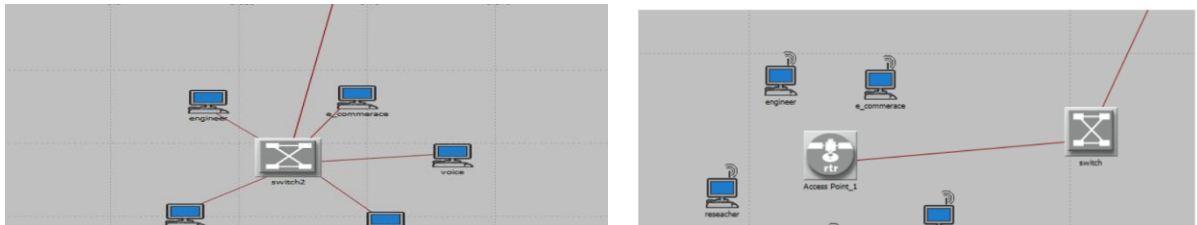
پیکربندی پایه ۸۰۲,۱۱g به عنوان یک سناریوی پایه با استفاده از یک مدل استاندارد از تنظیمات شبکه محلی بی‌سیم OPNET 14.5 ایجاد شده است. در این سناریو، رفتار یک شبکه محلی بی‌سیم ۸۰۲/۱۱g با یک زیرساخت در چارچوب شبکه محلی بی‌سیم سازمان‌یافته برای بهتر نمایش پیکربندی واقعی شبکه، مورد بررسی قرار می‌گیرد.(همانطور که در شکل (الف) نشان داده شده است).

یک ابر پروتکل اینترنتی (IP) می‌تواند برای نشان دادن اتصال اینترنت اصلی به یک لینک سریال نقطه به نقطه (1.544Mbps) T1 استفاده شود. سه زیرشبکه در هر طرف ابر IP از طریق یک دروازه IP که به یک لینک T1 نقطه به نقطه متصل است، قرار گرفته‌اند. همچنین دو سرور از طریق یک سوییچ مرکزی با استفاده از یک لینک BasetT 100متصل هستند که در شکل (الف) نشان داده شده و به صورت بی‌سیم در شکل (ب) قابل مشاهده است.



الف- چارچوب‌های سرور سیم‌دار شبیه‌سازی شده ب- چارچوب‌های سرور بی‌سیم شبیه‌سازی شده

یک زیرشبکه اول در سمت راست ابر پروتکل اینترنتی (IP) قرار دارد و سرورهای شبکه ترافیک از طریق اترنت BasetT 100 به هم متصل هستند. این سرورها از طریق لینک اترنت Baset 100 به دیوار آتش متصل هستند و می‌توانند به عنوان منابع و مقاصد در همه برنامه‌ها مانند ویدیو کنفرانس، پروتکل انتقال فایل (FTP)، پروتکل انتقال فایل (HTTP)، برنامه‌های صوتی، پست الکترونیکی (ایمیل) و شبیه‌سازی پایگاه داده در شبکه کامل که ترافیک مبادله شده با گره‌های موبایلWLAN۸۰۲,۱۱ g را مشخص می‌کند، استفاده شوند (همانطور که در شکل های زیر نشان داده شده است).



دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

شکل الف : زیرشبکه ۲ 802.11 WLAN شکل ب : زیرشبکه ۱ 802.11g WLAN

زیرشبکه دوم، دفتر شعبه دورکار را نشان می‌دهد که شامل پنج ایستگاه کاری در LAN دفتر است که با لینک 100 BaseT به هم متصل هستند. این LAN دفتر از طریق یک سوییچ مرکزی با لینک اترنت 100 Base به هم متصل است تا یک موقعیت دفتر واقعی را که از استاندارد LAN اترنت سریع استفاده می‌کند، شبیه‌سازی کند. دروازه IP LAN را به ابر IP متصل می‌کند، این دروازه به LAN دفتر از طریق لینک

اترنت 100 Base T متصل است، در حالی که لینک سریال نقطه به نقطه T1 ابر IP را به دروازه IP متصل می‌کند (همانطور که در شکل(ب) نشان داده شده است).

سرانجام، زیرشبکه سوم در سمت دیگر ابر پروتکل اینترنتی قرار دارد. اتصال شبکه بی‌سیم محلی (WLAN) از طریق نقطه دسترسی AP به LAN دفتر از طریق یک سوئیچ مرکزی با استفاده از سیم‌کشی اترنت انجام می‌شود تا محیطی از دفتر واقعی در استانداردهای LAN اترنت سریع را شبیه‌سازی کند که شامل گسترش یک WLAN به یک منطقه از سیم‌کشی دشوار یا زیبایی‌شناسی مورد نیاز مانند یک اتاق رسانه (صوت یا ویدیو) یا کنفرانس است (همانطور که در شکل ۳ نشان داده شده است).

شکل ۳: زیرشبکه ۳ WLAN دفاتر

LAN دفتر به طور مساوی بین پروفایل‌ها تقسیم شده است، به طوری که هر پروفایل یک ایستگاه کاری دارد. در اینجا هدف اصلی تحلیل عملکرد مفید ادراک شده توسط اپراتور مربوط به کاربران WLAN است. یک زیرشبکه WLAN با پنج گره موبایل به پروفایل‌های مختلف اختصاص داده شده است، همانطور که در جدول ۱ نشان داده شده است.

جدول ۱: نمایه‌های تخصیص داده شده به گره‌های موبایل در زیرشبکه WLAN

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

مشخصات کاربر	گره موبایل
مهندس	ایستگاه۱
محقق	ایستگاه۲
مشتری تجارت الکترونیک	ایستگاه۳
شخص فروش	ایستگاه۴

۲-۵. سناریوهای بار چندرسانه‌ای

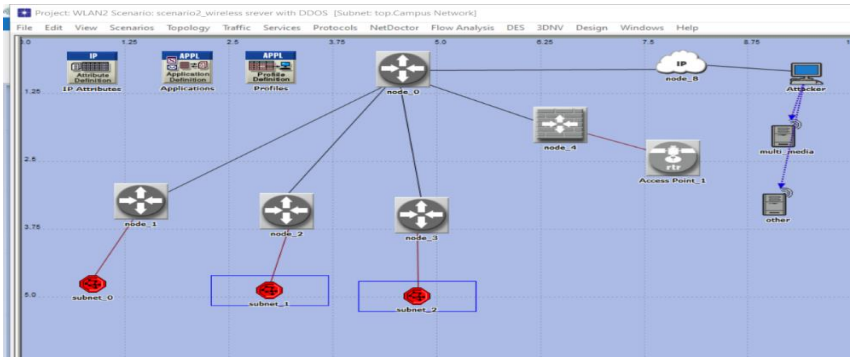
ترافیک چندرسانه‌ای می‌تواند به شکل یک ویدئو کنفرانس در شبکه نمایش داده شود. این شکل از کنفرانس شامل تصاویر، داده و صدا، و نمایش ترافیک چندرسانه‌ای تعریف شده است. در حالی که سرور ویدیو برای کمک به برنامه ویدئو کنفرانس در جزئیات شبکه ارائه شده است.

5-3. سناریوی حمله ی DDOS

همان طور که مشاهده می‌کنید یک نوع حمله از نوع حمله (DDOS) به آن اضافه می شود و 3 سناریو :

سناریوی اول :حمله DDOS به وایرلس سرور(با بار کم) شکل ۴

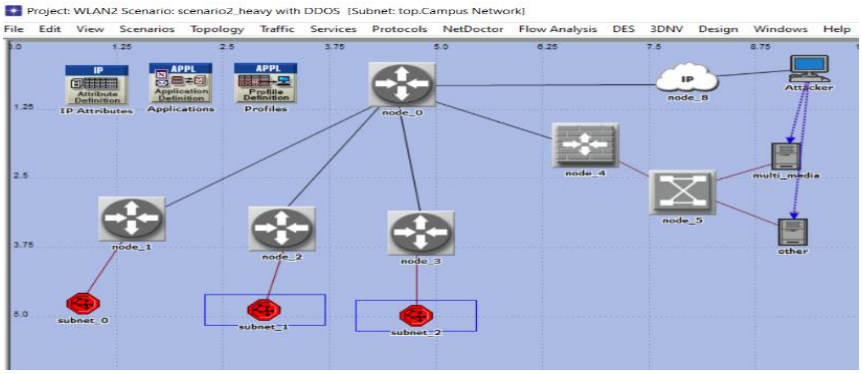
سناریوی دوم و سوم حمله DDOS به سرور با سیم (با بارکم و بار سنگین) شکل ۵



شکل ۴سناریوی اول : حمله DDOS به وایرلس سرور(با بار متوسط)

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

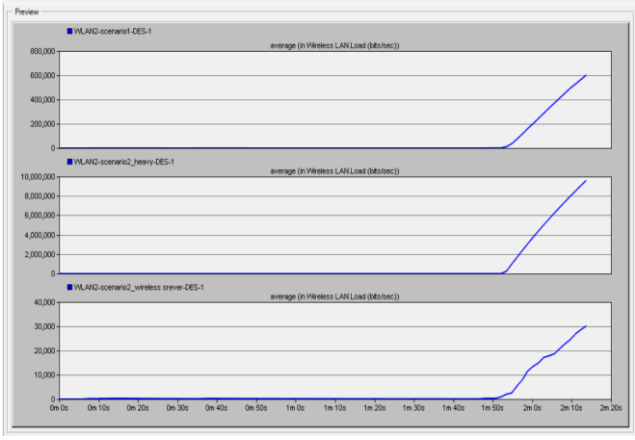


شکل ۵ سناریوی دوم :حمله DDOS به سرور با سیم (بار سنگین و بار کم)

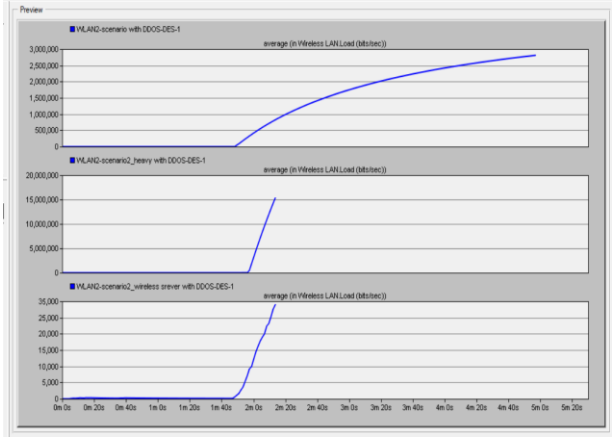
6. نتایج و بحث شبیه‌سازی

۶,۱. بار

بار اولین پارامتری است که بر کل عملکرد کارایی بی‌سیم تأثیر می‌گذارد. ارزیابی بار مربوط به دریافت داده‌های ارسال شده است، داده‌ها دارای میانگین کلی بار WLAN با مقدار تقریبی (۴۳۰/۷۴۰۷) Kbps در مدت ۵ دقیقه اجرا شده است. شکل ۶ نتایج مربوط به سه سناریو قبل از حمله را نشان می‌دهد وشکل ۷ نتایج حمله DDOS را نشان می دهد.



شکل ۶: قبل از حمله DDOS



شکل۷: در زمان حمله DDOS

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

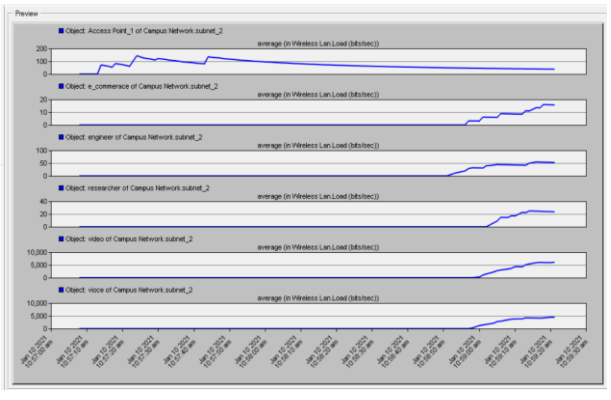
Print ISSN: 2251-9335
Online ISSN: 2252-0279

در مقایسه این دو تصویر:

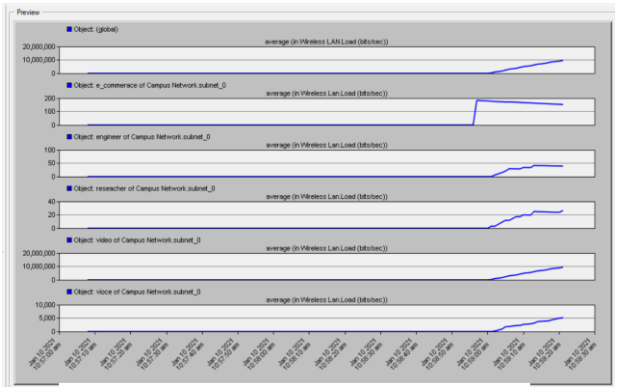
هر ۳ سناریو مقایسه شده است (قبل از حمله و بعد از حمله):

۱. هر دو تصویر نمودارهایی را نمایش می‌دهند که انرژی در ((Wireless LAN/Local (M bits/s) را در طول زمان برای پیکربندی‌های شبکه مختلف نشان می‌دهند.در شکل ۶ سناریوی اول مقدار بار به ۶۰۰/۰۰۰ رسیده است و در سناریوی دوم (با بار زیاد) به مقدار نزدیک ۱۰/۰۰۰/۰۰۰ و در سناریوی سوم وایرلس سرور(با بار کم) به مقدار :۳۰/۰۰۰ رسیده است. ۲-مشاهده می شود که در شکل ۷ میزان بار بیش تر شده است چون حمله DDOS به شبکه اتفاق افتاده است که در سناریوی اول مقدار بار به ۳/۰۰۰/۰۰۰رسیده است و در سناریوی دوم (با بار زیاد) به مقدار نزدیک ۱۵ /۰۰۰/۰۰۰ و در سناریوی سوم وایرلس سرور(با بار کم) به مقدار :۳۵/۰۰۰ رسیده است. از نظر زمانی هم در شکل ۶ در وایرلس سرور در زمان ۱ دقیقه و ۵۰ ثانیه بوده و بعد از حمله نیز در زمان: ۱ دقیقه و ۵۰ ثانیه اتفاق افتاده است. وقتی یک حمله DDoS به یک شبکه رخ می‌دهد، میزان ترافیک ورودی به آن شبکه به طور شدیدی افزایش می‌یابد. این افزایش بار باعث می‌شود که منابع شبکه مانند باند پهنای اینترنت، ظرفیت سرور و پهنای باند ارتباطی به سرعت تحت فشار قرار گیرند و در نتیجه، دسترسی به سرویس‌های شبکه برای کاربران واقعی

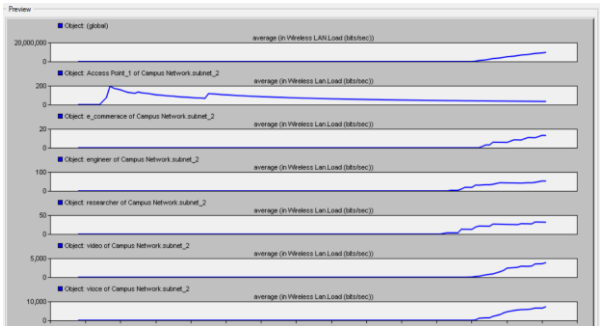
دچار اختلال شود. بنابراین، نمودارهای مربوط به چنین حملاتی معمولاً افزایش چشمگیر در میزان انرژی یا ترافیک شبکه را نشان می‌دهند که ناشی از بار سنگین وارد شده به سیستم است. این افت عملکرد و دسترسی‌پذیری شبکه می‌تواند منجر به اختلال در خدمات و افت رضایت کاربران شود.



شکل ۹ مقادیر بار فردی با بار زیاد (سابنت ۲)



شکل ۸ مقادیر بار فردی با بار زیاد (سابنت 0)



دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

شکل ۱۱: مقادیر فردی با بار زیاد بعد از حمله DDOS (سابنت۲)

شکل ۱۰ :مقادیر فردی با بار زیاد بعد از حمله DDOS (سابنت۰)

در این سناریوی شبیه‌سازی، دو حالت مختلف بار شبکه نمایش داده شده است: برای بار زیاد

۱. قبل از حمله DDoS (شکل‌های ۸ و۹):

- در این حالت، مقادیر بار برای subnet=0 و subnet=2 در تاخیر کمتری قرار دارند.

- میزان ترافیک و انرژی استفاده شده در شبکه در این مرحله در محدوده‌ای طبیعی و قابل قبول است.

2. بعد از حمله DDoS (شکل‌های ۱۱ و۱۰):

- در این حالت، مقادیر بار برای subnet=0 و subnet=2 افزایش یافته‌اند.

- این افزایش بار ناشی از حمله DDoS به سیستم است.

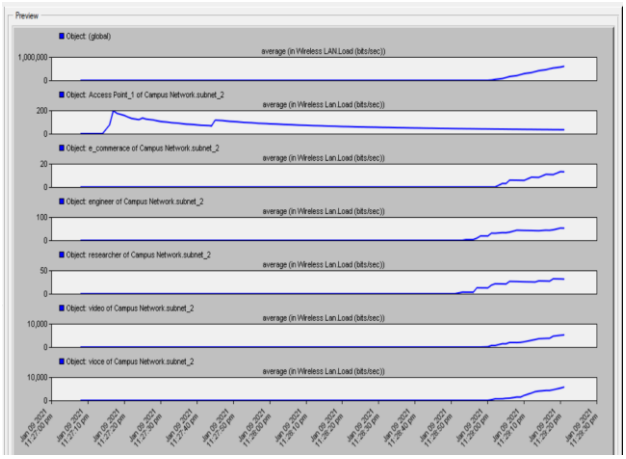
بنابراین، مقادیر بار در حالت بعد از حمله DDoS به طور قابل ملاحظه‌ای بالاتر از حالت قبل از حمله است و این افزایش بار، نشان‌دهنده اثرات مخرب این نوع حملات بر روی زیرساخت‌های شبکه است.

جدول ۲ سناریوی مقادیر بار انفرادی سناریوی ۱ (سبک) و سناریوی ۲ (سنگین)

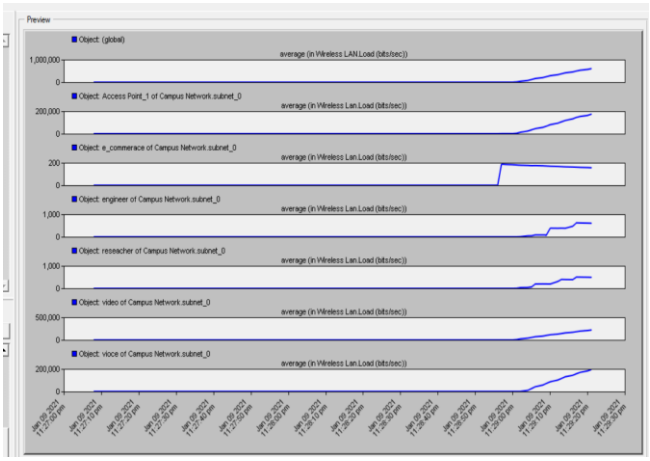
نوع گره	بار WLAN (سنگین)(kbps) زیر شبکه1	بار WLAN(سنگین) (kbps) زیر شبکه سه	بار WLAN (سبک)(kbps) زیر شبکه یک	بار WLAN (سبک)(kbps) زیر شبکه 3
نقطه دسترسی	(AP2) 42.07	(AP1) 200.2	(AP1) 916124	(AP3) 899128
تجارت الکترونیک	0	188.4	73	146
مشتری				
مهندس	0	0	223	232
محقق	0	0	202	260
ویدئو	0	0	893783	87652
کنفرانس				
صدا	0	0	17878	18201
مقدار فرعی	42.07	388.6	1828285	1794502
جمع کل	430.7407407		3622787.467	

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279



شکل۱۳. مقادیر بار فردی برای سناریوی (با بار کم)



شکل۱۲. مقادیر بار فردی برای سناریوی (با بار کم)

در این حالت نیز سناریوی شبیه‌سازی، دو حالت مختلف بار شبکه نمایش داده شده است: برای بار کم

۱. قبل از حمله DDoS (شکل‌های ۱۳ و ۱۲):

- در این حالت، مقادیر بار برای subnet=0 و subnet=2 در تاخیرخیلی کمتری قرار دارند.

- میزان ترافیک و انرژی استفاده شده در شبکه در این مرحله در محدوده‌ای طبیعی و قابل قبول است.

۲. بعد از حمله DDoS (شکل‌های ۱۵ و ۱۴):

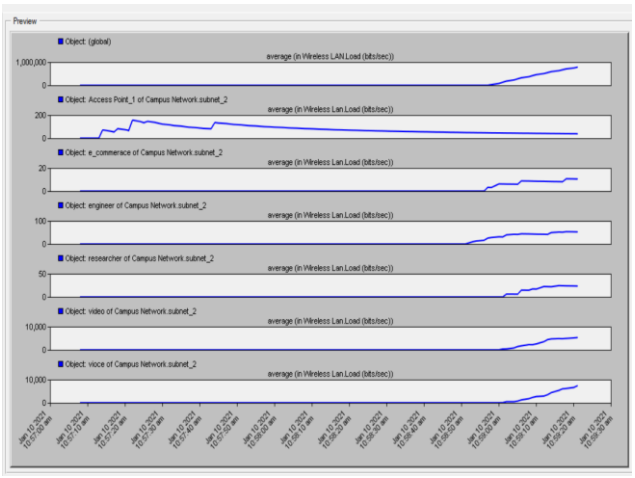
- در این حالت نیز، مقادیر بار برای subnet=0 و subnet=2 افزایش کمی را داشته اند.

- این افزایش بار ناشی از حمله DDoS به سیستم است.

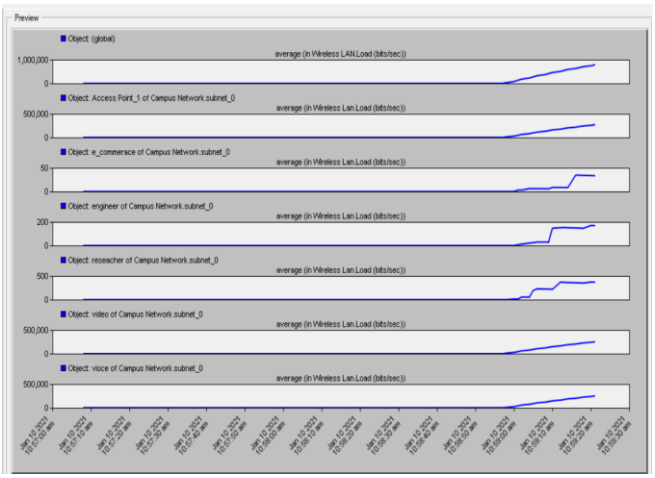
دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

بنابراین، مقادیر بار در حالت بعد از حمله DDoS به طور قابل ملاحظه‌ای بالاتر از حالت قبل از حمله است.



شکل۱۵. مقادیر بار فردی برای سناریوی (با بار کم) بعد از حمله



شکل۱۴. مقادیر بار فردی برای سناریوی (با بار کم) بعد از حمله

۶,۲. تاخیر

تاخیر می‌تواند یک معیار بنیادی برای توصیف کیفیت سرویس (QoS) هر شبکه ای، به ویژه در زمان واقعی باشد. در کاربرد چندرسانه‌ای، تاخیر می‌تواند یک پارامتر مرکزی برای انتخاب عملکرد مؤثر لایه MAC، زمان عملیات آن و مکانیسم Required To Send/ Clear To Send (RTS/CTS) باشد. تاخیر دو نوع اصلی دارد: تاخیر دسترسی به رسانه و آمار تاخیر کلی انتقال بسته. مشاهده می‌شود که تاخیر WLAN که تاخیر پایان تا پایان همه بسته‌های دریافتی توسط MAC های WLAN توسط همه گره‌های شبکه WLAN و سپس هدایت شده به لایه‌های بالاتر را مشخص می‌کند، بسیار بالا است که نشان دهنده تلاش‌های مجدد زیاد است.

در شکل ۱۶، نمودارها مربوط به سه سناریوی مختلف هستند:

WLAN2-scenario2_heavy_with_DDOS-DES-**1**

WLAN2-scenario2_heavy_DES-**1**

WLAN2-scenario1_wireless_server_DES-**1**

هر سناریو با یک رنگ مشخص شده است و روند میانگین تاخیر ترافیک شبکه را در طول زمان نشان می‌دهد. درشکل ۱۶، تاخیر ترافیک شبکه برای سه سناریو نشان داده شده است.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

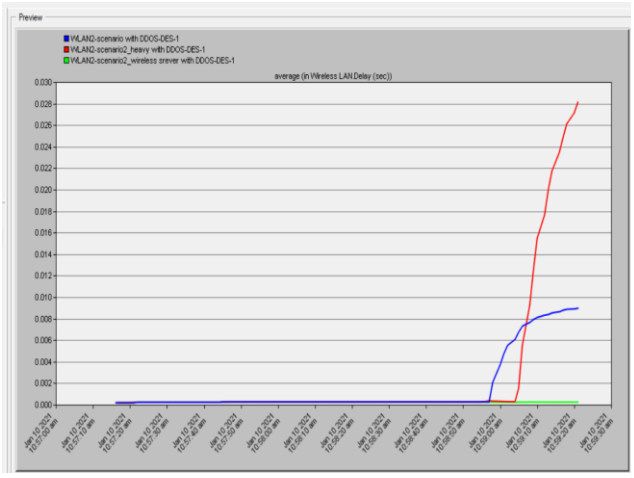
شکل۱۷ نمودارهایی هستند که تاثیر حمله DDOS را بر روی ترافیک شبکه نشان می‌دهند.

WLAN2-scenario_with_DDOS-DES-**1**

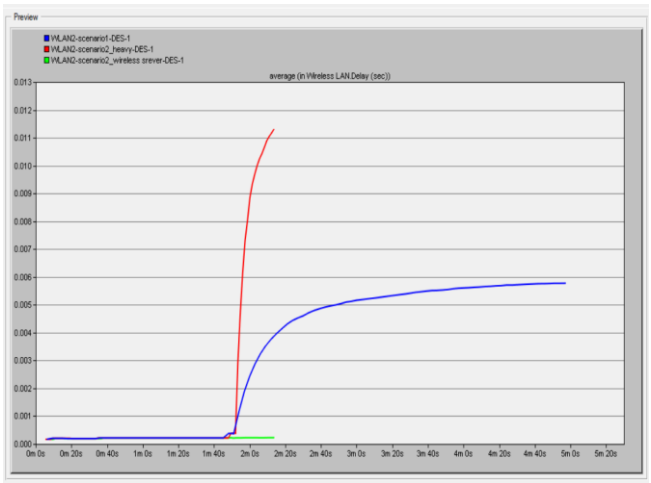
WLAN2-scenario2_heavy_with_DDOS-DES-**1**

WLAN2-scenario2_wireless_server_with_DDOS-DES-**1**

درشکل ۱۷، تاخیر ترافیک شبکه بی‌سیم محلی به صورت واضح تر و با جزئیات بیشتری نشان داده شده است. مقایسه این دو تصویر نشان می‌دهد که حمله DDOS چگونه تاخیر ترافیک شبکه را به شکل قابل توجهی افزایش می‌دهد.



شکل ۱۷ : میزان تاخیر بعد از حمله(به صورت یک نمودار)



شکل ۱۶ میزان تاخیر قبل از جمله(به صورت یک نمودار)

در یک حمله DDoS به شبکه، عامل اصلی افزایش تأخیر در شبکه است:

۱. کاهش پاسخ‌گویی سیستم:

- با اشباع منابع شبکه، سیستم قادر به پردازش و پاسخ‌گویی به همه درخواست‌ها به موقع نخواهد بود.

- این موضوع منجر به افزایش چشمگیر در زمان پاسخ‌گویی و تأخیر در ارائه سرویس‌ها به کاربران واقعی می‌شود.

۲. اختلال در مسیریابی و قطع ارتباط:

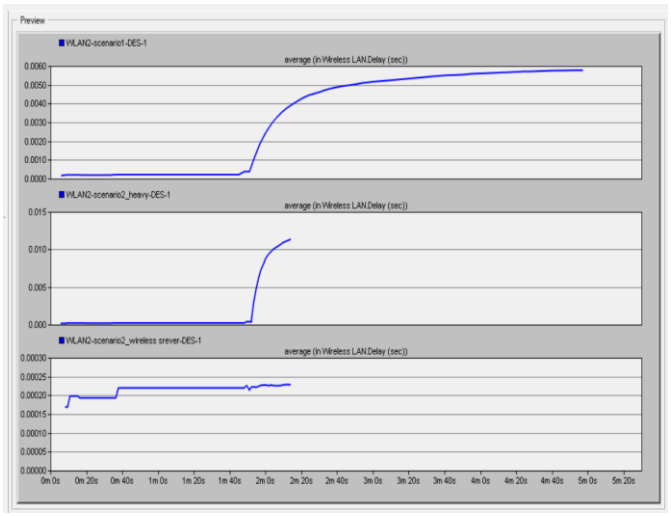
- در شرایط اشباع شبکه، مسیریابی ترافیک و انتقال داده‌ها دچار اختلال می‌شود.

- در نتیجه، بسته‌های داده به درستی مسیریابی نشده و ممکن است ارتباطات قطع شوند.

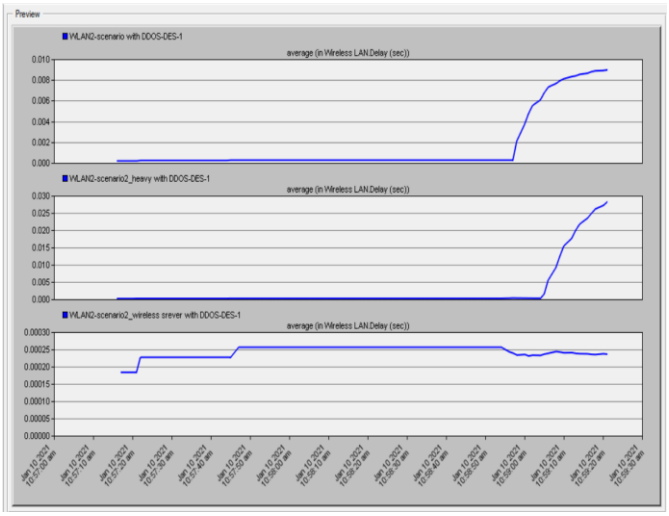
دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

بنابراین، نمودارهای مربوط به چنین حملاتی معمولاً افزایش چشمگیر در میزان تأخیر در شبکه را نشان می‌دهند. این افزایش تأخیر ناشی از بار سنگین وارد شده به سیستم و ناتوانی آن در پاسخ‌گویی به موقع به همه درخواست‌ها است. این موضوع می‌تواند منجر به قطع ارتباطات و اختلال جدی در خدمات ارائه شده به کاربران نهایی شود. درشکل۱۷ نمودارهای بعد از حمله: سناریوی اول مقدار تاخیر به ۰,۰۱۰ رسیده است و در سناریوی دوم (با بار زیاد) مقدار تاخیر به مقدار نزدیک ۰,۰۲۸ و در سناریوی سوم وایرلس سرور(با بار کم) آن چنان تاخیری نداشته است. در مقابل، شکل های۱۸ و ۱۹ که قبل از حمله و بعد از حمله را به صورت ۳ سناریو و نمودارهای تکی ارائه داده اند، این افزایش تأخیر به خوبی نمایش داده شده است.



شکل ۱۸ : میزان تاخیر قبل از حمله(به صورت نمودارهای جدا از هم



شکل۱۹: میزان تاخیر بعد از حمله (به صورت نمودار های جدا از هم)

7- نتیجه‌گیری

در این مقاله، عملکرد شبکه‌های WLAN قبل از حمله و بعد از حمله از نظر تأخیر انتهایی به انتهایی، با استفاده از ترافیک چندرسانه‌ای در سناریوی کنفرانس ویدئویی مورد ارزیابی قرار گرفته است. شبیه‌سازی‌ها با استفاده از مدل‌ساز OPNET 14.5 تنظیم شده است. مقاله به استفاده از مکانیزم RTS/CTS (درخواست برای ارسال/واضح برای ارسال) در شبکه‌های بی‌سیم اشاره می‌کند. این یک پروتکل لایه MAC است که در شبکه‌های WLAN IEEE 802.11 برای کمک به هماهنگی دسترسی به رسانه بی‌سیم مشترک و کاهش مشکل گره پنهان استفاده می‌شود. در مجموع، پروتکل کلیدی مورد بحث در این مقاله، پروتکل WLAN IEEE 802.11 است که از مکانیزم RTS/CTS به عنوان بخشی از هماهنگی دسترسی لایه MAC استفاده می‌کند. عملکرد کیفیت سرویس (QoS) در شبکه‌های IP می‌تواند به کارگیری کارآمدترین روش از منابع شبکه موجود را برای به حداقل رساندن تأخیرهای ترافیک شبکه که دارای خدمات چندرسانه‌ای متنوعی از جمله صدا، ویدئو و پایگاه داده است، به کار گیرد. چهار سناریو برای شبکه ایجاد شده است: سناریوی ۱ (بار سنگین)، سناریوی ۲ (بار متوسط تا پایین) و سناریوی ۳ (بار سنگین)سناریوی4 (حمله DDOS) مقایسه بین آنها انجام شده است و ارزیابی‌های عملکردی هر دو به بار و تأخیر عملکردها مربوط بوده است. در این آزمون‌های عملکرد، نتایج در دو سناریوی قبل از حمله و بعد از حمله مقایسه شده است.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

مراجع

[۱] Al-Khaffaf, D.A.J. and M.G. Al-Hamiri, Performance evaluation of campus network involving VLAN and broadband multimedia wireless networks using OPNET modeler. TELKOMNIKA (Telecommunication .Computing Electronics and Control), 2021. **19**(5): p. 1490-1497

[۲] Atmaca, S., et al., A new MAC protocol for broadband wireless communications and its performance .evaluation. Telecommunication Systems, 2014. **57**: p. 13-23

[۳] Lafta, S.A., et al., Quality of service performances of video and voice transmission in universal mobile telecommunications system network based on OPNET. Bulletin of Electrical Engineering and Informatics, .2021. **10**(6): p. 3202-3210

[٤] Ghazala, M.M.A., M.F. Zaghloul, and M. Zahra, Performance evaluation of multimedia streams over Wireless Computer Networks (WLANs). International Journal of Advanced Science and Technology, .2009. **13**: p. 63-76

[٥] Srivastava, A., et al. A recent survey on DDoS attacks and defense mechanisms. in International .and Applications. 2011. Springer Conference on Parallel Distributed Computing Technologies

[٦] .Khan, S., et al., Denial of service attacks and challenges in broadband wireless networks. 8; 7, 2008

[٧] Tannoury, A., et al. Efficient and accurate monitoring of the depth information in a Wireless Multimedia Network based surveillance. in 2017 Sensors Networks Smart and Emerging Technologies Sensor .(SENSET). 2017. IEEE

[٨] Mateen, A., et al. Comparative analysis of wireless sensor networks with wireless multimedia sensor networks. in 2017 IEEE International Conference on Power, Control, Signals and Instrumentation .Engineering (ICPCSI). 2017. IEEE

[٩] Hussein, W.A., et al. Design and performance analysis of high reliability-optimal routing protocol for mobile wireless multimedia sensor networks. in 2017 IEEE 13th Malaysia International Conference on .Communications (MICC). 2017. IEEE

[١٠] Abbas, N. and F. Yu. A traffic congestion control algorithm for wireless multimedia sensor networks. in .2018 IEEE SENSORS. 2018. IEEE

[١١] wireless lans for multimedia streaming services. Advances in Lee, H.K., et al., Bandwidth estimation in .Multimedia, 2007. **2007**(1): p. 070429

[١٢] Alasti, M., et al., Quality of service in WiMAX and LTE networks [Topics in Wireless Communications]. .١١١-IEEE Communications Magazine, 2010. **48**(5): p. 104

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

توسعه استراتژیک و ایجاد نوآوری در کسب‌وکار با تکیه بر هوش مصنوعی و فناوری بلاکچین

مصطفی قبادی*

* کارشناس ارشد مدیریت فناوری اطلاعات-کسب و کار الکترونیک-دانشگاه آزاد اسلامی-واحد تهران شمال info@ghobadi.ir

نوع مقاله: علمی پژوهشی

چکیده:

در عصری که تحولات کسب‌وکار با سرعت بی‌سابقه‌ای در حال وقوع است، فناوری‌های پیشرفته مانند هوش مصنوعی (AI) ، امکانات تازه‌ای را برای ارتقاء عملکرد تجاری فراهم می‌کنند. این پیشرفت‌ها، تعاملات شرکتی با مشتریان و کارمندان را از طریق خدمات مبتنی بر فناوری اطلاعات تحول می‌بخشند. با گسترش استفاده از AI ، کسب‌وکارها باید بر استراتژی‌های فعلی خود بازنگری کرده و به طور فعال به دنبال کشف فرصت‌های جدید بازاری باشند. با افزایش توجه به تحقیقات در زمینه نوآوری‌های تجاری، بلاکچین به عنوان یک راهکار برای تضمین امنیت داده‌ها مطرح شده است. در این مقاله، مدل نوآوری کسب‌وکار مبتنی بر AI و بلاکچین (BI-AIBT) به منظور تقویت فرآیندهای تجاری و تضمین تعاملات امن بین مشتریان متنوع معرفی شده است. این مدل با استفاده از داده‌های تجربی کیفی از شرکت‌کنندگان در دو حوزه کسب‌وکار مورد بررسی قرار گرفته است BI-AIBT .با تحلیل تأثیر استفاده از فناوری اطلاعات بر ایجاد ارزش، پیشنهادها و جذب کسب‌وکار مورد سنجش قرار گرفته و نشان داده است که بلاکچین می‌تواند در تقویت تعاملات بین ظرفیت‌های سازمانی و مهارت‌های کارکنان مؤثر باشد. نتایج آزمایشی این مدل نشان می‌دهد که تحول ناشی از فناوری اطلاعات به عنوان یک عنصر مهم در تقویت استراتژی‌های نوآوری کسب‌وکار شناخته شده و مدل BI-AIBT نسبت‌های پیش‌بینی تقاضا (۹۷،۱%)، کیفیت محصول (۹۸،۳%)، توسعه کسب‌وکار (۹۸،۹%)، تحلیل رفتار مشتری (۹۶،۳%)، و رضایت مشتری (۹۷،۲%) را تقویت می‌کند

کلید واژگان: بلاکچین، فناوری های پیشرفته، نوآوری، هوش مصنوعی

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

۱. بررسی اجمالی نوآوری کسب وکار مبتنی بر هوش مصنوعی و فناوری بلاکچین:

پیشرفت تجاری اغلب با ادغام رویکردها، خدمات یا محصولات نوینی همراه است که به بهبود عملکرد سازمانی منجر می‌شود. این فرایند که به عنوان پیشرفت تجاری شناخته می‌شود، می‌تواند از تغییرات جزئی در فرآیندهای موجود گرفته تا ابداعات کاملاً جدید را در بر گیرد. (Morkunas, Paschen & Boon, 2019) هوش مصنوعی به سیستم‌هایی اطلاق می‌شود که قادر به انجام وظایف پیچیده‌ای هستند که پیشتر فقط توسط انسان‌ها انجام می‌شد. بلاکچین یک شبکه غیرمتمرکز از رایانه‌ها است که داده‌ها را ثبت و ذخیره می‌کند تا یک سلسله رویدادهای متوالی را در یک سیستم ثبت شفاف و غیرقابل تغییر، نشان دهد.

این پیشرفت‌های فناورانه می‌توانند موجب تحول یا ایجاد مجدد روش‌ها و عملیات‌های مرسوم شوند. (Hu et al., 2018) در دنیای کنونی، فناوری‌هایی مانند بلاکچین و هوش مصنوعی قادرند تا ساختارهای موجود را متحول کنند، (Nguyen, Liu, Chu & Weng, 2018) مدل‌های کسب‌وکار پیشرفته‌ای را پرورش دهند و تحولات عمده‌ای را در صنایع ایجاد کنند. بلاکچین، با فراهم کردن یک دفتر کل توزیع‌شده و غیرمتمرکز، می‌تواند به افزایش سطح اعتماد، پاسخگویی، امنیت و حریم خصوصی در بخش‌های تجاری منجر شود. (J. Gao, Wang & Shen, 2020; Sun, Yan & Zhang, 2016) این فناوری که قابلیت نگهداری انواع دارایی‌ها را دارد، به دفتر کل ثبت یا دفتر کل توزیع‌شده تشبیه می‌شود. دفتر کل توزیع‌شده که به صورت اجماعی میان اشخاص مختلف در مکان‌ها و سازمان‌های گوناگون به اشتراک گذاشته می‌شود، (Pham, Nguyen, Nguyen, Pham & Nguyen, 2020) همان پایه و اساسی است که بلاکچین بر آن استوار است و همان فناوری‌ای است که در بیت‌کوین به کار رفته است. در نوعی از فناوری دفتر کل توزیع‌شده، معاملات با استفاده از یک نشانه رمزنگاری دائمی به نام هش ثبت می‌شوند و به این صورت بلاکچین نامیده می‌شوند. (Asghar et al., 2021; Ruan et al., 2019) معاملات در بلوک‌هایی جمع‌آوری شده و هر بلوک جدید حاوی هش بلوک پیشین است که این امر باعث اتصال آن‌ها به یکدیگر می‌شود. این جزئیات بیشتر با مسائل مالی و هویتی در ارتباط هستند. در مناطق آلمانی‌زبان و اروپا، اینترنت اشیاء (IoT) به بهینه‌سازی بخش‌های مختلف و تسهیل فرآیندهای کسب‌وکار کمک می‌کند و در نهایت، هوش مصنوعی با شناسایی و بهینه‌سازی نتایج فرآیندهای کسب‌وکار، به ارتقاء عملکرد آن‌ها کمک می‌کند. (Manogaran, Alazab, Shakeel & Hsu, 2021)

هدف از نوآوری در شرکت‌ها، افزایش منافع مالی است که از طریق ایجاد بازارهای جدید و افزایش درآمدها در سیستم‌های فعلی، (Jan et al., 2020) یا کاهش زمان و منابع مصرفی و ارتقاء عملکرد به دست می‌آید. (Kumar et al., 2020) نوآوری در زمینه بلاکچین، به عنوان یک پیشرفت انقلابی، مهارت‌های موجود را دگرگون می‌سازد، چرا که باعث منسوخ شدن فناوری‌های پیشین می‌شود. (Arjun & Suprabha, 2020; J. Gao, Wang & Shen, 2020) این نوع نوآوری، با تغییر ساختارهای کسب‌وکار موجود در بازار، به عنوان یک نوآوری معماری مخرب شناخته می‌شود. (Fu et al., 2020) بلاکچین، با ایجاد بازارها و پلتفرم‌های همکاری غیرمتمرکز که شامل توانایی‌های محاسباتی، داده‌ها و الگوریتم‌های مورد استفاده در بخش‌های مختلف هوش مصنوعی می‌شود، (G. Manogaran et al., 2020; P et al., 2020) زمینه‌ساز توسعه‌هایی گسترده‌تر در این حوزه است. هوش مصنوعی به دستگاه‌هایی اطلاق می‌شود که برای انجام کارهای ذهنی طراحی شده‌اند. (Filimonau & Naumova, 2020; Wang, Huang, Hsu & Yang, 2016) بلاکچین، به عنوان یک شبکه کامپیوتری غیرمتمرکز، داده‌ها را در یک دفتر کل شفاف و ثابت ثبت و نگهداری می‌کند و رویدادها را به ترتیب نمایش می‌دهد. (Khelifi et al., 2020) ترکیب بلاکچین و هوش مصنوعی، می‌تواند به افزایش یادگیری ماشین و دسترسی به محصولات مالی برای هوش مصنوعی کمک کند. (Amin, Faragallah & El-Latif, 2010; Kaur, Garg, Kaddoum, Ahmed & Atiquzzaman, 2019) بلاکچین، امنیت و اشتراک‌گذاری داده‌ها را تسهیل می‌کند. پیشرفت‌های اخیر در بلاکچین ممکن است تأثیرات عمیقی بر زنجیره‌های تأمین داشته باشند، (Feng, He, Zeadally,

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

(Khan & Kumar, 2019; Ur-Rehman, Gondal, Kamruzzaman & Jolfaei, 2020) با حذف واسطه‌ها و هموارسازی فرآیندها، امنیت و مدیریت داده‌ها را ارتقاء دهند. هوش مصنوعی با استانداردسازی و خودکارسازی فرآیندها، پتانسیل(ظرفیت) کاهش زمان و هزینه‌ها را دارد و به افزایش بهره‌وری و کارایی کمک می‌کند (Mistry, Tanwar, Tyagi & Kumar, 2020) همچنین تصمیم‌گیری سریع‌تر مدیریت را بر اساس داده‌های تحلیلی فراهم می‌آورد. هوش مصنوعی به شرکت‌ها امکان می‌دهد تجربه‌ای شخصی‌تر به مشتریان ارائه دهند و در تحلیل داده‌های بزرگ بسیار کارآمد است، به سرعت الگوهایی را در داده‌ها شناسایی می‌کند. در نهایت، هوش مصنوعی نقشی مهم در بازاریابی دیجیتال و ترویج برند دارد، به ویژه در کمپین‌هایی که نیاز به سرعت عمل بالا دارند.

سیستم‌های پیشرفته هوش مصنوعی قادر به شناسایی و تحلیل داده‌های مربوط به مشتریان هستند تا از این طریق، ارتباطات مؤثرتری را در لحظات کلیدی برقرار سازند. این فناوری‌ها امکان ارائه پیام‌های متناسب با نیازهای فردی را بدون نیاز به مداخله نیروی انسانی فراهم می‌آورند که این امر به افزایش کارایی کمک می‌کند. استفاده از هوش مصنوعی به مؤسسات، این امکان را می‌دهد که نظارت دقیق‌تری بر برند دیجیتالی خود داشته باشند و حضور آنلاین(برخط) خود را مستحکم‌تر کنند. این امر از طریق تجزیه و تحلیل دقیق صفحات وب، شبکه‌های اجتماعی و دیگر پلتفرم‌ها توسط مدیران برند و خدمات میسر می‌شود. فناوری بلاکچین ظرفیت دگرگون‌سازی عملکرد کسب‌وکارها و ساختارهای اجتماعی و اقتصادی را دارد. این فناوری می‌تواند در شرایط خاص برای تأیید معاملات با هزینه‌ای اندک به کار رود. بلاکچین، امنیت احراز هویت را از طریق شناسه‌های الکترونیکی که به مشتریان و کارکنان اختصاص داده شده، تقویت می‌کند. با ذخیره‌سازی این داده‌ها در یک بلاکچین عمومی، احتمال وقوع جرایمی مانند کلاهبرداری هویتی و مالی کاهش می‌یابد. نوآوری در کسب‌وکار مبتنی بر هوش مصنوعی و بلاکچین (BI-AIBT) به منظور ارتقاء فرآیندهای کاری و ایجاد یک ارتباط امن با مشتریان طراحی شده است. مطالعات موردی از دو بخش کسب‌وکار متفاوت نشان‌دهنده تأثیرات متفاوت دیجیتالی‌سازی بر ایجاد ارزش و جذب مشتری است. همچنین، فناوری بلاکچین می‌تواند به تقویت همکاری بین توانایی‌ها و مهارت‌های درون سازمانی کمک کند.

در این تحقیق، دستاوردهای اصلی به شرح زیر هستند:

- ابتکار در طراحی کسب‌وکار با استفاده از قابلیت‌های هوش مصنوعی و فناوری زنجیره بلوک(BI-AIBT) ، که به منظور تقویت پیشرفت‌های تجاری به کار گرفته شده است.
- بررسی و تعیین کاربردهای فناوری زنجیره بلوک در ایجاد لایه‌های امنیتی برای حفاظت از داده‌های ذخیره‌شده در شبکه.
- تحلیل‌های عددی نشان‌دهنده افزایش نسبت‌های پیش‌بینی تقاضا، کیفیت محصول، پیشرفت‌های تجاری، تحلیل رفتار مشتریان و رضایت مشتریان در مقایسه با مدل‌های قبلی است.

ساختار باقی مقاله به صورت زیر تنظیم شده است: بخش دوم به بررسی مطالعات مرتبط با نوآوری‌های کسب‌وکار اختصاص دارد. بخش سوم، خلاصه‌ای از مطالعه پیشنهادی را که در این تحقیق به کار رفته است، ارائه می‌دهد. بخش چهارم به توضیح نتایج شبیه‌سازی و بحث‌های مربوطه می‌پردازد. در نهایت، بخش پنجم با تحلیل دقیق مشاهدات و نتایج به پایان می‌رسد.

۲. آثار مرتبط با نوآوری در کسب و کار:

در حوزه نوآوری کسب‌وکار، تحقیقات گسترده‌ای بر اجرای دانش علمی و فناوری‌های پیشرفته در شرکت‌های کوچک و متوسط (SMEs) تاکید دارند. Stratan et al. (Stratan, Novac & Vinogradova, 2020) به این نتیجه رسیدند که ارتقاء کارایی در SME ها از طریق به‌کارگیری رویکردهای نوین در مدیریت و همکاری با موسسات دیگر، از جمله دانشگاه‌ها، امکان‌پذیر است. این

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

امر می‌تواند به تقویت توانایی‌های نوآورانه شرکت‌ها منجر شود. با این وجود، بسیاری از SME های مستقر در مولداوی با چالش‌هایی در زمینه اجرای اختراعات و تحقیقات علمی مواجه هستند.

Thuethongchai et al. (Belazi, Khan, El-Latif & Belghith, 2016) بر اهمیت تحلیل‌های بزرگ داده‌ای (BDA) برای ارتقاء فرآیندهای تصمیم‌گیری و طراحی خدمات در شرکت‌ها تاکید کردند. آن‌ها معتقد بودند که BDA می‌تواند به شناسایی نیازهای دقیق مشتریان از طریق داده‌های دیجیتال و در نتیجه به افزایش خلاقیت در خدمات کمک کند.

Trad et al. (Thuethongchai, Taiphapoon, Chandrachai & Triukose, 2020; Trad, 2021)، یک مدل ریاضی جامع کاربردی (AHMM) را توسعه دادند که برای بررسی تحولات در کسب وکارها، ریاضیات کاربردی، مهندسی کسب‌وکار و نیز علم مالی به کار می‌رفت. این مدل به تقلید از فرآیندهای شناختی انسانی و استفاده از ابزارهای اوریستیکی در تصمیم‌گیری‌ها می‌پرداخت. AHMM همچنین به تسهیل هماهنگی میان مفاهیم مختلف معماری سازمانی (EA) و روش‌های انتقال دانش کمک می‌کرد، که این امر در نهایت به پشتیبانی از ابتکارات تحولی در شرکت‌ها کمک می‌کرد.

در مطالعه‌ای که توسط Hakala et al. (Nguyen, Leu, Zeadally, Liu & Chu, 2018) انجام شد، روشی نوین تحت عنوان مدل‌سازی داستانی (MNR) معرفی شد که امکان پیوند دادن فعالیت‌های برجسته در بازار و اکوسیستم‌های کارآفرینی و نوآوری را فراهم می‌آورد. این روش، با روشن‌سازی تعدادی از تفسیرهای پنهان و انتظارات اساسی، به افزایش شفافیت و وضوح در گفتمان‌های مربوط به اکوسیستم کمک می‌کند. MNR با تلفیق و بازتاب استراتژیک تشابهات و تفاوت‌های میان مفاهیم مرتبط، ساختاری را برای تولید داستان‌های مدلی فراهم می‌آورد که می‌تواند جایگزینی مناسب برای پرسشنامه‌های تحقیقاتی باشد.

Borah et al. (Hakala, O'Shea, Farny & Luoto, 2020) ، مفهوم مداخلات بازاریابی خلاقانه (IMI) را مطرح کردند که شامل اقدامات مرتبط با رسانه‌های اجتماعی است که به صورت متنی و همزمان با وقوع رویدادهای واقعی انجام می‌پذیرد. مطالعات متعددی شامل آزمایش‌های شبه و تحلیل‌های داده‌ای به ثبت رسیده و نظریه‌ای که تأثیر IMI را با استفاده از عناصر طنز و غافلگیری توضیح می‌دهد، تأیید شده است. این تحقیقات، اهمیت IMI در رسانه‌های اجتماعی و ویژگی‌هایی را که شرکت‌ها باید برای بهره‌برداری فعال از شبکه‌های آنلاین و کسب ارزش‌های مالی دنبال کنند، مورد تأکید قرار داده‌اند.

Mustafa et al., (Mustafa & Khan, 2020) به بررسی فرصت‌ها و چالش‌های پیش روی بانکداری اسلامی در برابر شرکت‌های فین‌تک سنتی پرداختند. این مطالعه نشان داد که موسسات بانکی اسلامی، شرکت‌های فین‌تک اسلامی را بیشتر به عنوان شرکای همکار تلقی می‌کنند تا رقبای تجاری. Zhao et al., (Zhao, Xue, Khan & Khatib, 2021) به مطالعه رفتار مشتریان و تأثیر آن بر رشد شرکت‌ها پرداختند و مدل‌های محاسباتی هوشمند ترکیبی سازگار (AHICM) را توسعه دادند. این مدل‌ها، نوآوری‌های محصول، تغییر نگرش‌ها و ذهنیت‌های جدید در جامعه و بررسی دقیق بازارها و نیازهای مشتریان را ضروری می‌دانند.

Manogaran et al., (G. Manogaran et al., 2020) سیستم امنیتی یکپارچه مبتنی بر بلاکچین (BISM) را معرفی کردند که با هدف ارائه مدیریت دسترسی امن و حفاظت از حریم خصوصی افراد و کالاها طراحی شده است. این سیستم، دسترسی کاربران را بر اساس وضعیت منابع مجازی در بازه‌های زمانی متفاوت تنظیم می‌کند، در حالی که حفاظت از حریم خصوصی بر پایه مدت زمان پاسخ‌دهی استوار است. در نهایت، پیشنهاد شده است که نوآوری‌های کسب‌وکار مبتنی بر هوش مصنوعی و فناوری بلاکچین (BI-AIBT) می‌تواند با ارتقاء روش‌های کسب‌وکار موجود، از تکنیک‌های فعلی فراتر رفته و به بهبود نسبت‌های پیش‌بینی تقاضا، کیفیت محصول، توسعه کسب‌وکار، تحلیل رفتار مشتری و رضایت مشتری کمک کند.

۳. پیشنهاد نوآورانه در حوزه کسب‌وکار متکی بر فناوری‌های پیشرفته هوش مصنوعی و بلاکچین:

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

بهره‌گیری از قابلیت‌های بلاکچین به عنوان یک پروتکل امن جمع‌آوری داده‌ها است که مقاومت بالایی در برابر تغییرات غیرمجاز، نفوذ و تقلب دارد. بلاکچین، به عنوان یک دفتر کل دیجیتال، در سراسر شبکه‌های کامپیوتری تکثیر می‌شود و به عنوان فناوری دفتر کل توزیع‌شده (DLT) شناخته می‌شود که سابقه‌ای دائمی و شفاف از تمام دارایی‌های دیجیتالی را از طریق فرآیندهای تمرکززدایی و رمزنگاری هش ارائه می‌دهد. این فناوری، زنجیره‌ای از اطلاعات توزیع‌شده و غیرمتمرکز را فراهم می‌کند که دسترسی همزمان به سوابق را ممکن می‌سازد.

در دوران اخیر، کاربردهای بلاکچین فراتر از محدوده مالی گسترش یافته و در زمینه‌هایی چون مدیریت زنجیره تأمین و هویت دیجیتال مورد استفاده قرار گرفته است. استفاده از بلاکچین در مدیریت زنجیره تأمین، مزایایی نظیر بهبود همکاری، کاهش هزینه‌ها و افزایش کارایی را به همراه دارد. این امکانات، سازمان‌ها را در مدیریت بهینه تقاضا، تحویل به‌موقع موجودی، مدیریت اختلالات، کاهش هزینه‌ها و پاسخگویی مؤثر به نیازهای مشتریان یاری می‌رساند. بلاکچین همچنین به افراد این قدرت را می‌دهد که کنترل بیشتری بر هویت دیجیتال خود داشته باشند، به گونه‌ای که سازمان‌ها تنها با رضایت مشتریان می‌توانند از داده‌ها استفاده کنند و هیچ نهاد مرکزی نمی‌تواند هویت فردی را به خطر بیندازد.

مطالعات اخیر بر اهمیت ادغام بلاکچین با سایر فناوری‌های نوین مانند اینترنت اشیاء (IoT) و هوش مصنوعی (AI) تأکید دارند. به‌عنوان مثال، بلاکچین می‌تواند در بهبود معماری سیستم‌های IoT کاربرد داشته باشد. هوش مصنوعی می‌تواند به دستگاه‌های IoTاجازه دهد تا رفتارهای هوشمندانه‌ای را شبیه‌سازی کنند و در تصمیم‌گیری‌های خودکار بدون نیاز به دخالت انسانی عمل کنند IoT .امکان ارتباط دستگاه‌ها از طریق اینترنت را فراهم می‌آورد، در حالی که AI به این دستگاه‌ها امکان می‌دهد تا از داده‌ها و تجربیات خود یاد بگیرند و به شبکه های بلاکچین ارسال کنند تا سوابق معاملاتی مشترک و غیرقابل تغییر ایجاد کنند. بلاکچین IBM به شرکای تجاری اجازه می‌دهد تا داده‌های IoT را به اشتراک بگذارند و به آن دسترسی داشته باشند بدون اینکه نیاز به یک سیستم مدیریت مرکزی باشد. در حوزه فناوری اطلاعات، تلفیق بلاکچین، اینترنت اشیاء (IoT) و هوش مصنوعی (AI) می‌تواند به ایجاد سیستم‌هایی منجر شود که با همکاری یکدیگر، کارایی و امنیت بیشتری را ارائه دهند. این سه فناوری، هر یک دارای قابلیت‌های منحصر به فردی هستند که می‌توانند در ترکیب با یکدیگر، ارزش افزوده قابل توجهی را ایجاد کنند. به عنوان مثال، IoT می‌تواند به عنوان یک سیستم جمع‌آوری داده عمل کند، در حالی که AI می‌تواند این داده‌ها را تحلیل و برای بهبود فرآیندها و تصمیم‌گیری‌ها استفاده کند. بلاکچین نیز می‌تواند به عنوان یک لایه امنیتی عمل کرده و اطمینان حاصل کند که تراکنش‌ها و داده‌ها به صورت شفاف و غیرقابل تغییر ثبت شوند. (Borah, Banerjee, Lin, Jain & Eisingerich, 2020)

در مدل کسب‌وکار مبتنی بر این تلفیق، بلاکچین می‌تواند نقش مهمی در مدیریت تراکنش‌های صنعتی ایفا کند، به طوری که نیازهای منابع و تأمین‌کنندگان منابع را به هم وصل کند. این امر می‌تواند به ایجاد یک شبکه تأمین متمرکز یا غیرمتمرکز منجر شود که در آن تراکنش‌های جدید به صورت دیجیتالی ثبت و مدیریت می‌شوند. قراردادهای هوشمند می‌توانند به عنوان پروتکل‌های تأیید عمل کنند که الگوهای داده را برای تراکنش‌ها بین تأمین‌کنندگان و مشتریان بررسی و حفاظت می‌کنند. (G. Manogaran et al., 2020)

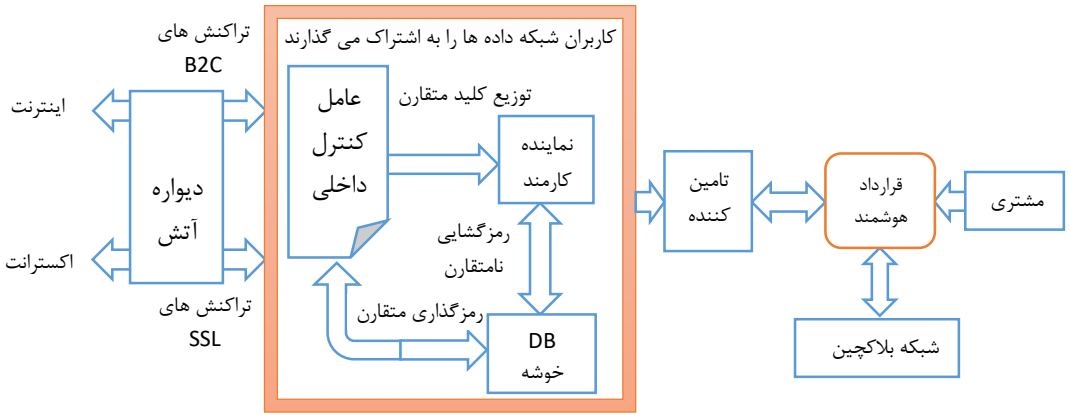
در این میان، محصولات نرم‌افزاری تجاری-آماده‌به‌کار (COTS) می‌توانند به عنوان ابزارهایی برای تسریع در توسعه و پیاده‌سازی سیستم‌ها عمل کنند، زیرا این محصولات از پیش توسعه یافته و برای استفاده در دسترس هستند. با این حال، ممکن است نیاز به سطوح بالاتری از امنیت و سفارشی‌سازی باشد که از طریق مدل‌های حفاظتی مأموریتی قابل دستیابی است. این مدل‌ها می‌توانند بر اساس نیازهای خاص هر سازمان تنظیم شوند و لایه‌های اضافی از حفاظت را برای داده‌ها در شبکه کسب‌وکار فراهم آورند.

در نهایت، نماینده کارمندان (EA) می‌تواند نقش مهمی در کنترل پارامترهای کلیدی تولید و تعیین فرکانس بازسازی کلیدهای عمومی و خصوصی داشته باشد که این امر به افزایش امنیت و کارایی کل سیستم کمک می‌کند. این نمایندگان می‌توانند به عنوان واسطه‌هایی عمل کنند که مسئولیت‌های خاصی را برای بهبود عملکرد کلی سیستم بر عهده دارند. یک سازمان می‌تواند فردی را به

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

عنوان نماینده مجاز برای انجام معاملات مهم تعیین کند، به گونه‌ای که تنها این فرد می‌تواند به نمایندگی از سازمان عمل نماید، در حالی که سایر کارکنان این اجازه را ندارند. در زمینه حسابداری هزینه‌های غیرمستقیم(ICA) ، دسترسی به حساب‌های اجرایی (EA)به صورت محدود و تنها برای خوشه‌های مشخصی از پایگاه داده فراهم می‌شود. این سیستم از الگوریتم‌های ریاضی برای تأمین امنیت اطلاعات استفاده می‌کند که در آن اطلاعات توسط رمز عبور یا کلید خصوصی رمزگشایی می‌شوند. وب‌سایت‌هایی که داده‌های حساس مانند اطلاعات کارت‌های اعتباری یا شماره‌های حساب بانکی را منتقل می‌کنند باید این داده‌ها را قبل از ارسال به پایگاه داده رمزنگاری کنند تا از دزدیده شدن آن‌ها جلوگیری شود. در این سیستم، هر خوشه داده با استفاده از کلید منحصر به فردی رمزنگاری می‌شود تا تنها ایستگاه‌های کاری مجاز قادر به دسترسی به اطلاعات باشند ICA .در حالت دریافت، داده‌های مصرف‌کننده و شرکت را از طریق فایروال از اینترنت یا شبکه‌های خارجی دریافت و تعیین می‌کند که کدام کلید باید برای رمزنگاری داده‌ها به صورت متقارن استفاده شود. در حالت پردازش، ICAکلیدهای مخفی را تولید می‌کند که برای رمزنگاری داده‌ها استفاده می‌شوند. (G. Manogaran et al., 2020) با استفاده از رمزهای عبور، اطلاعات مشتری و شرکت در پایگاه داده ذخیره می‌شوند . ICAمسئولیت اصلی مدیریت تحویل را بر عهده دارد . EA ها می‌توانند بر اساس معیارهای مختلفی مانند موقعیت یا موجودی حساب تخصیص داده شوند. در حالت ارسال، ICA کلیدهای مخفی را به EA های مربوطه ارسال و داده‌های دریافتی از شبکه‌های خارجی را با استفاده از کلید مخفی رمزنگاری می‌کند. در پایگاه داده خوشه‌ای، ICA داده‌ها را ثبت می‌کند. (Nguyen, Leu & Liu, 2017) کارکنان می‌توانند حساب‌های کسب‌وکار خود را در ایستگاه‌های کاری خود مدیریت کنند. فرستنده داده‌ها را با استفاده از کلید عمومی گیرنده رمزنگاری می‌کند و گیرنده با استفاده از کلید خصوصی مرتبط با داده‌ها، آن‌ها را رمزگشایی می‌کند. این لایه احراز هویت پیشرفته از داده‌ها در برابر دسترسی‌های غیرمجاز محافظت می‌کند. تکنولوژی بلاکچین با ایجاد بایگانی‌های داده‌های ثابت و دسترس‌پذیر، به افزایش کیفیت داده‌ها کمک می‌کند. کیفیت اطلاعات به عنوان یکی از عوامل مهم در عملکرد سازمانی شناخته می‌شود و گزارش‌های نادرست ناشی از جمع‌آوری ناکافی اطلاعات می‌توانند به اشتباهات منجر شوند. عدم یکپارچگی در کدگذاری داده‌ها بین سیستم‌های مختلف می‌تواند مانع از شناسایی صحیح خطرات شود که در رفع مسائل مرتبط با وظایف مختلف مؤثر است.



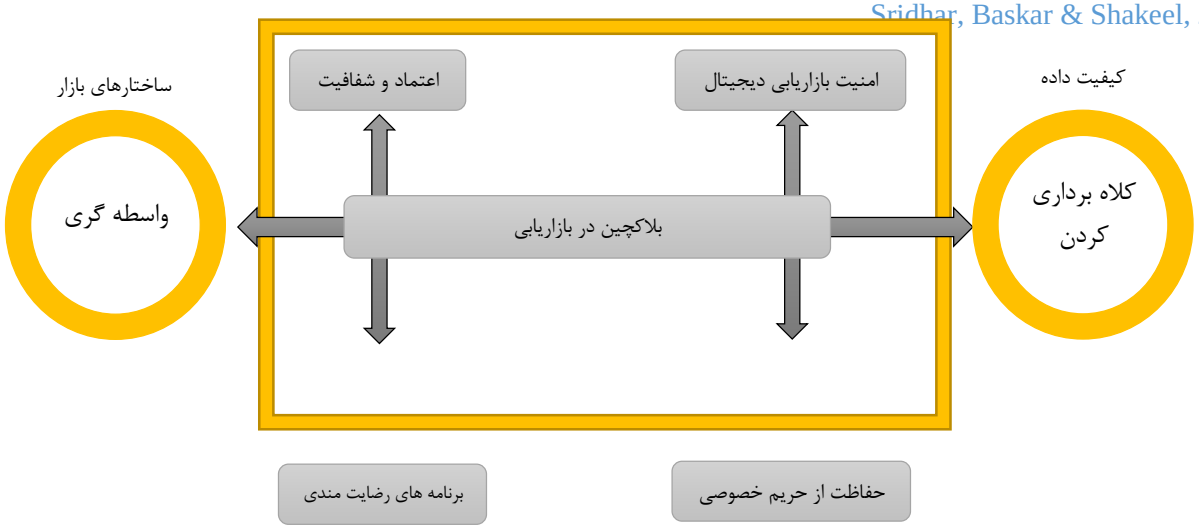
شکل ۱. نوآوری تجاری مبتنی بر هوش مصنوعی و فناوری بلاک چین (BI-AIBT)

فناوری‌های بلاکچین قادر به تحول اساسی در تجربیات تجاری از منظر مشتریان با امکان دسترسی به داده‌ها و حفاظت از حریم خصوصی هستند. این فناوری‌ها، سازوکارهای نوینی برای افزایش رضایت مشتریان فراهم می‌آورند که به ایجاد ارزش افزوده کمک می‌کنند. با پیشرفت اینترنت، روش‌های تحویل محصولات و خدمات شرکت‌ها دگرگون شده‌اند. فناوری‌های نوین، راهبردهای

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

تجارت سنتی را دگرگون ساخته، نیاز به واسطه‌های سنتی را کاهش داده و واسطه‌های الکترونیکی جدیدی را معرفی کرده‌اند. همچنین، اینترنت به عنوان بستری برای ارائه واسطه‌های برخط جدید با محصولات و خدمات نوین عمل می‌کند. (Sheron, Sridhar, Baskar & Shakeel, 2019)



شکل ۲. تاثیر بر بازاریابی بلاک چین

از سوی دیگر، گسترش شبکه‌های اجتماعی، نیاز کسب‌وکارها به حضور در این شبکه‌ها برای جلب توجه مصرف‌کنندگان را افزایش داده است. اینترنت به کانالی برای توزیع تبدیل شده که به تبلیغ‌کنندگان امکان می‌دهد محصولات و خدمات خود را به فروش برسانند و با مشتریان ارتباط برقرار کنند. اینترنت به عنوان یک رسانه ارتباطی کارآمد عمل می‌کند که به شرکت‌ها اجازه می‌دهد مستقیماً با مشتریان در تماس باشند و آن‌ها را از محصولات، خدمات و نوآوری‌های جدید خود مطلع سازند. با وجود ارزش بی‌چون و چرای حضور برخط، مواردی از کلاهبرداری، حواشی و تبلیغات ناموفق، اعتبار صنعت بازاریابی و تبلیغات را تحت تأثیر قرار داده‌اند. این مسائل به طراحی خودکار تبلیغات وب و افزایش تقاضا برای بازاریابی هدفمند منجر شده‌اند. بخشی از مصرف‌کنندگان به محصولات یا خدمات یک کسب‌وکار نیازمند یا علاقه‌مند هستند که این گروه بخشی از بازار کلی را تشکیل می‌دهند. به عنوان مثال، اسباب‌بازی‌های کودکان ممکن است برای پسران ۹ تا ۱۱ ساله طراحی شده و والدین آن‌ها به عنوان گروه هدف در نظر گرفته شوند. تبلیغات مبتنی بر اینترنت به عنوان تبلیغات وب شناخته می‌شوند که شرکت‌هایی ا که در فضای وب فعالیت دارند، شامل می‌شود. بازاریابی نه تنها در زمان پرداخت هزینه یا تلاش فعال برای جذب مشتری صورت می‌گیرد، بلکه شامل کمپین‌های ایمیلی، فعالیت‌های رسانه‌های اجتماعی، وب‌سایت و وبلاگ شرکت نیز می‌شود. کلیک‌فراگ یا تقلب کلیک، استفاده عمدی از برنامه‌های کامپیوتری یا افراد برای ایجاد کلیک‌های غیرواقعی بر روی تبلیغات برخط است که می‌تواند به منظور کسب منافع نامشروع یا خالی کردن بودجه تبلیغاتی انجام شود. تقلید غیرقانونی از تبلیغات پرداخت‌به‌ازای‌کلیک (PPC) برای افزایش درآمد سایت یا تخلیه بودجه تبلیغاتی یک سازمان به عنوان تقلب کلیک شناخته می‌شود. گاهی اوقات، صاحبان سایت‌ها ممکن است برای افزایش درآمد تبلیغاتی خود به صورت نادرست اقدام به این کار کنند. بلاکچین می‌تواند در شناسایی تقلب مؤثر باشد، با امکان به اشتراک‌گذاری داده‌ها به صورت زنده و به‌روزرسانی سوابق بر اساس توافق همه طرف‌ها، که نه تنها از تقلب جلوگیری می‌کند بلکه هزینه‌ها و زمان لازم برای فرآیندها

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

را کاهش می‌دهد. یک چارچوب معتبر می‌تواند خطرات ناشی از کلیک‌های فریبنده را با فراهم آوردن یک محیط بازاریابی دیجیتالی امن‌تر برای مصرف‌کنندگان و نام‌های تجاری مرتفع سازد. امنیت دیجیتالی که به عنوان بیمه شخصیت پیشرفته شناخته می‌شود، به حفاظت از هویت واقعی افراد در سازمان‌ها و دسترسی‌های شبکه‌ای که استفاده می‌کنند، می‌پردازد. این حفاظت شامل دستگاه‌هایی است که افراد برای تأمین هویت، منابع و نوآوری خود در فضای برخط و قابل حمل به کار می‌برند. بازاریابی دیجیتالی قادر است به کسب‌وکارها در جذب ترافیک فوری، ایجاد سرنخ‌ها و تنظیم معاملات از طریق دستیابی به افرادی که در جست‌وجوی محصولات یا خدمات آن‌ها هستند، کمک کند. تبلیغات وب، بدون نیاز به مداخله دیگران، روشی برای معرفی آنلاین نام تجاری به سرنخ‌های بالقوه و خریداران ارزشمند است. اعتبار یک نام تجاری به اعتماد و شفافیت آن بستگی دارد. فناوری بلاکچین به بازاریابان و مشتریان امکان می‌دهد تا در یک اکوسیستم امن و شفاف برای تقویت اعتماد و شفافیت در بازاریابی دیجیتالی فعالیت کنند. بلاکچین که به عنوان دفتر کل دیجیتالی تعریف می‌شود، معاملات یا سایر رویدادها را به صورت دائمی ثبت می‌کند. اعتماد و شفافیت از طریق دفتر کلی که ورودی‌های آن قابل حذف یا تغییر نیستند، تضمین می‌شوند. همچنین، این فناوری به افراد اجازه می‌دهد تا معاملات انجام‌شده را مشاهده و بررسی کنند. شفافیت فراهم‌شده توسط بلاکچین اعتماد را افزایش می‌دهد، زیرا مشتریان قادر به مشاهده و تأیید ادعاهای نام‌های تجاری هستند. این شفافیت می‌تواند شامل تأیید گواهینامه‌ها توسط نهادهای ثالث، رویه‌های تجاری و مسئولیت‌های اجتماعی شرکت‌ها باشد. نام‌های تجاری می‌توانند چندین هویت الهام‌بخش را به نمایش بگذارند و اطمینان حاصل کنند که آن‌ها در حفظ شفافیت و تأکید بر انگیزه‌های خیرخواهانه خود برای دنبال کردن منافع برتر مشتریان پایبند هستند. حفاظت از داده‌های شخصی در عصر دیجیتال، چالشی است که به طور مداوم در حال تغییر و تحول است و می‌تواند تأثیر قابل توجهی بر اعتماد مصرف‌کنندگان به استفاده از منابع دیجیتال داشته باشد. مطالعات نشان می‌دهند که مصرف‌کنندگان در انجام معاملات خود تمایل به حفظ ناشناسی و احتیاط دارند. این امر به دلیل افزایش خطرات ناشی از دسترسی‌های غیرمجاز و سوءاستفاده‌های احتمالی است. امنیت در شبکه‌های بلاکچین که شامل استفاده از ساختارهای دفاعی برخط، سیستم‌های تأیید هویت و بهترین روش‌های مقابله با حملات و تقلب است، به عنوان یک سیستم جامع مدیریت ریسک شناخته می‌شود. مسائل مربوط به حفظ حریم خصوصی با جمع‌آوری و ذخیره‌سازی اطلاعات شخصی توسط کوکی‌های وب‌سایت‌ها تشدید می‌شود. با استفاده از روش‌های پیشرفته داده‌کاوی و تکنولوژی‌های نوین در جمع‌آوری داده‌ها، بازاریابان قادر به تجزیه و تحلیل سریع و دقیق اطلاعات مشتریان هستند. این تحولات، مسائل جدیدی را در زمینه حریم خصوصی مصرف‌کنندگان برخط به وجود آورده است. برندها باید زیرساخت‌های فناوری مستحکمی را توسعه دهند تا تحلیل‌های کیفی را بهبود بخشیده و توجه مصرف‌کنندگان را در فضای بازاریابی دیجیتال افزایش دهند، قبل از اینکه امنیت هویتی را در استراتژی‌های بازاریابی خود گنجانند. بازاریابی دیجیتال از دو بعد میکرو و ماکرو تشکیل شده است که به ترتیب به محیط کاری و محیط کلان اشاره دارند. این دو بعد بر جوامع، مشاغل، خریدها و رفتارهای مصرف‌کنندگان تأثیر می‌گذارند و باید در استفاده از فناوری‌های پیشرفته که تأثیر و شناخت برند را افزایش می‌دهند، نقش مهمی ایفا کنند. زیرساخت‌های فناوری پایدار به این معنی است که برندها و مصرف‌کنندگان از مزایای فناوری بلاکچین بهره‌مند می‌شوند که سطح بالایی از امنیت را تضمین می‌کند. از دیدگاه مشتری‌مداری، بلاکچین می‌تواند تأثیر قابل توجهی بر بهبود ارتباطات و داده‌های مشتریان داشته باشد.

تمرکز اولیه بر روی شفافیت است و در ادامه، تقویت مکانیزم‌های حفاظتی و امنیتی مورد توجه قرار می‌گیرد. این رویکرد شامل ارزیابی روش‌های نوآورانه در برنامه‌های وفاداری مشتری است که می‌تواند به خلق ارزش افزوده منجر شود. امنیت در فناوری بلاکچین به طور چشمگیری تحت تأثیر ساختار ذخیره‌سازی داده‌های پراکنده و غیرمتمرکز است. پایگاه‌های داده متمرکز نیز با محدودیت‌های خاصی مواجه هستند که می‌توان به موارد زیر اشاره کرد: این داده‌ها به شدت به ساختار شبکه وابسته‌اند. هرچه سرعت اینترنت کمتر باشد، زمان لازم برای دسترسی به پایگاه داده افزایش می‌یابد. همچنین، ممکن است به دلیل حجم بالای ترافیک، انسداد ایجاد شود. استفاده از پروتکل‌های امنیتی متفاوت، از جمله رمزنگاری نامتقارن، امضاهای دیجیتال و مدیریت دسترسی، کیفیت، تحویل و بازیابی داده‌ها را برای تعداد زیادی از مشتریان تضمین می‌کند. رمزنگاری نامتقارن که به عنوان رمزنگاری کلید عمومی نیز شناخته می‌شود،

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

از دو کلید رمزنگاری مجزا اما به لحاظ ریاضی مرتبط برای رمزگذاری و رمزگشایی اطلاعات استفاده می‌کند. امضاهای دیجیتال که به‌مانند "اثرانگشت‌های" الکترونیکی عمل می‌کنند، یک امضاکننده را به صورت ایمن به یک سند در یک معامله ثبت‌شده متصل می‌کنند. این امضاها، پیاده‌سازی خاصی از فناوری امضای الکترونیکی (eSignature) هستند. مدیریت دسترسی که سیستمی برای کنترل دسترسی منابع توسط نمایندگان، شرکا، پیمانکاران و مشتریان است، به طور فزاینده‌ای برای حفظ امنیت اهمیت پیدا کرده است، چه این کنترل به صورت دستی، در برنامه‌ها کدگذاری شده یا به صورت خودکار در یک پلتفرم سازمانی انجام شود. در زمینه تحلیل نظریه؛ با استفاده از هوش مصنوعی و عملکرد کمکی سیستم‌های کامپیوتری، یک طرح توسعه یافته است. این روش که برای توضیح عملکردهای یک سیستم پیچیده به کار می‌رود، به عنوان عملکردهای تحلیل نظریه شناخته می‌شود. اصل اساسی این است که سیستم به عنوان محاسبه‌کننده یک تابع یا به طور کلی‌تر، به عنوان حل‌کننده یک مشکل پردازش اطلاعات در نظر گرفته می‌شود. کاری که باید توضیح داده شود به مجموعه‌ای منظم از توابع ساده‌تر تقسیم می‌شود. برای آماده‌سازی زمینه‌های لازم برای توسعه کارهای بعدی، از کد خاص کامپیوتر برای طبقه‌بندی داده‌های استراتژی تحلیل نظریه استفاده می‌شود. ترکیب داده‌ها از سیستم بهینه‌سازی در بازاریابی در معادله (۱) به دست می‌آید:

$$Q(n)=\frac{inv_{me}(y)}{\vartheta^{2}}\sqrt{\rho_{1}}y \qquad Q(m)=\frac{\rho_{1}-inv_{me}(y)}{\vartheta^{0.5}}\sqrt{inv_{me}(y)\rho_{1}} \tag{۱}$$

در معادله (1) ، Q را می‌توان به عنوان بخشی از تجزیه و تحلیل نظری ضرایب پروژه در نظر گرفت، ρ1 را می‌توان به عنوان شاخصی برای نوآوری در زمینه بهینه‌سازی وب‌سایت‌های تجاری تعریف کرد، invme به عنوان یک راه‌حل قابل اجرا برای نوآوری مطرح است و ϑ به عنوان شاخصی برای بهینه‌سازی در نظر گرفته می‌شود. همچنین، ضریب بهینه‌سازی تقاضا، در تحلیل‌های نظری مربوط به عملکرد و قابلیت اجرای گره‌ها مورد بررسی قرار می‌گیرد. این تحلیل‌ها شامل مدلی است که پیش‌تر توصیف شده و با هدف افزایش کارایی عملیاتی مدل‌های نوآورانه و تدوین معیارهای امنیتی مطابق با فرآیندهای بازرسی طراحی شده‌اند. علاوه بر این، S و Rبه عنوان شاخص‌های تحلیل نظری مربوط به عملکرد و قابلیت اجرای گره‌ها مطرح هستند و b به عنوان شاخص بهینه‌سازی تجاری معرفی می‌شود. این تحلیل‌ها نیز شامل مدلی است که پیش‌تر توصیف شده و با هدف افزایش کارایی عملیاتی مدل‌های نوآورانه و تدوین معیارهای امنیتی مطابق با فرآیندهای بازرسی طراحی شده‌اند.

$$S_{b}A=\frac{R_{A}b_{A}}{\left(\frac{R_{A}}{b_{A}}\right)+\left(\frac{S_{B}}{b_{B}}\right)}\times S_{b} \tag{2}$$

به منظور دستیابی به دقت مطلوب در راه‌حل، روش آزمایشی مطلوب تا زمانی که ضریب گره بعدی در حالت بهینه قرار گیرد، به کار گرفته می‌شود. تنظیم ضریب تغییرپذیری جهت تعدیل ضریب ارزش و بهینه‌سازی تابع اطلاعاتی، جزئی از تحلیل نظری پیوستگی است. با توجه به اینکه دامنه تنش ماکزیمم مطلوب‌ترین حالت را نشان می‌دهد، تنظیم مطلوب موقعیت نقطه در این حالت صورت می‌گیرد و این امر منجر به پیشرفت در روش تحلیل نظری می‌شود. مطالعه فرضی سیستم توزیع سرمایه فیلتر Gt در معادله زیر آورده شده است:

$$G_{t}=\frac{\sum_{n}^{m-1}y_{m}+R_{m}}{\sum R_{m}^{0.5}}+(A.v_{1}) \tag{3}$$

مطابق با معادله (۳)، استخراج مقدار (y_m) از طریق یک روش پیشنهادی مبتنی بر بهینه‌سازی، منجر به ارتقاء(R_m) ، که نشان‌دهنده قابلیت اطمینان مدل (A) است، می‌شود. این مقاله یک روش فنی نوآورانه را معرفی کرده که به عنوان یک فرایند آزمایشی مرحله‌ای (v_1) عمل می‌کند، به این معنا که روش‌های پیشین نمی‌توانند دقت عملیات‌های بعدی را پیش‌بینی کنند.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

در حوزه خطاها، سه دسته‌بندی اصلی وجود دارد: خطاهای نحوی، خطاهای منطقی، و خطاهای اجرایی. خطاها به طور کلی به دسته‌های سیستماتیک، تصادفی و دائمی تقسیم می‌شوند. خطاهای ناخالص ناشی از اشتباهات تصادفی در استفاده از ابزارها، اندازه‌گیری‌ها و ثبت داده‌ها هستند. برای اطمینان از دقت نتایج، باید خطاهای الگوریتمی کاهش یابند. این بخش از تحقیق برای محاسبه احتمالات مرتبط با ورودی‌ها و استخراج خروجی‌های بهینه طراحی شده است. فرمول‌بندی مورد استفاده در مدل در معادله (۴) به شرح زیر است:

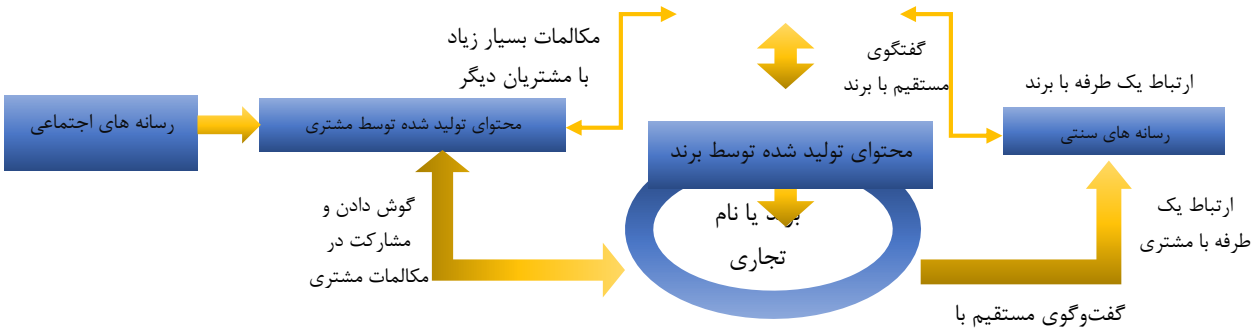
$$g = \sqrt{-1 + a'}\left(\frac{2-a}{22^2-1}\right) + \sum b_n \quad (۴)$$

مطابق با معادله (۴) که ارائه شده، عملیات تمرکزی (g) به عنوان مبدأ برای مرحله نهایی عمل می‌کند که در نتیجه، به (a') که نشان‌دهنده عملکرد دستورالعمل‌های متینگ و (bn) که ارزش تولید شده توسط هوش مصنوعی است، منجر می‌شود. نوع بعدی ضریب بهینه‌سازی در پلتفرم کسب‌وکار، استراتژی بهینه‌سازی برای ساختار را نشان می‌دهد و نشان‌دهنده ارزش نوآوری در توسعه الگوریتم‌ها، برچسب مدل کوتاه الکتریکی و ضریب برای مدل‌های نوآوری است. (\beta_{mx}) به عنوان برچسب مدل کوتاه الکتریکی شناخته می‌شود. شکل ۳ طراحی برای همکاری در استراتژی‌های بازاریابی را نمایش می‌دهد، که این همکاری‌های چندوجهی از برند تا مشتری و از مشتری تا مصرف‌کننده، بخشی از توسعه چارچوب است. رسانه‌های اجتماعی به عنوان پلتفرمی برای تبلیغات هم برای مشتری و هم برای برند عمل می‌کنند. به طور خاص‌تر، مدار تماس به عنوان جزء اصلی در جذب مشتریان مورد توجه قرار می‌گیرد. برندها می‌توانند از طریق رسانه‌های سنتی و چندین کانال دیگر، از جمله تبادلات رسانه‌های اجتماعی و تعاملات مشتری به صورت برخط یا دیجیتالی، با مشتریان خود ارتباط برقرار کنند. شرکت‌ها می‌توانند از وب‌سایت(وب‌گاه) خود به عنوان مرکزی برای رسانه‌های اجتماعی استفاده کنند یا مصرف‌کنندگان را با پلتفرم‌های رسانه‌های اجتماعی طرف‌های ثالث مانند فیسبوک و توییتر مرتبط سازند. توسعه ارزش برند به عنوان جزء مهمی از برندینگ در نظر گرفته می‌شود و به طور کلی به معنای "پول تأمین شده توسط برند برای محصول" است. وفاداری به برند، ویژگی‌های مصرف‌کننده‌ای را شناسایی می‌کند که به شدت به یک برند متصل هستند و زمانی که یک برند نسبت به دیگران موفق‌تر است و اغلب استفاده می‌شود، وابستگی مثبت خریداران به یک مورد خاص یا برند، به عنوان وفاداری به برند شناخته می‌شود. مشتریانی که وفاداری به برند را نشان می‌دهند، به محصول یا خدماتی پایبند هستند که از طریق خریدهای مکرر خود، علی‌رغم تلاش‌های رقبا برای جذب آن‌ها، این وفاداری را نشان می‌دهند. این یک بخش حیاتی از بازاریابی است که به سازمان‌ها کمک می‌کند تا برند قوی بسازند و مشتریان را دوباره جذب کنند. این فرآیند تنها به خرید مجدد محصولات محدود نمی‌شود، بلکه شامل ایجاد یک تصویر برند مثبت در ذهن مشتری و تبدیل شدن به یک طرفدار مثبت برند نیز است. وفاداری به برند، مقایسه محصولات با سایر برندهایی که همان مزایا را ارائه می‌دهند شامل می‌شود و یکی از اجزای ارزش برند است که در آن اندازه‌گیری‌های اعتماد به برند صورت می‌گیرد، بنابراین، جریان‌های نقدی آینده را تضمین می‌کند. بازدیدکنندگان می‌توانند به صورت فردی به سایت(پایگاه اینترنتی) فیسبوک مراجعه، صفحات مربوط به برندها را مرور کنند، به دنبال نام تجاری مورد نظر باشند، به اطلاعات شبکه‌های اجتماعی دوستان دست یابند یا اطلاعاتی را با دوستان خود به اشتراک بگذارند. آن‌ها در جست‌وجوی آخرین اخبار کسب‌وکار و تخفیف‌های موجود بر روی سایت فیسبوک هستند. این اطلاعات قابلیت به اشتراک‌گذاری با دیگر کاربران در شبکه را دارند. ارتباط مثبت با برند، تأثیر به‌سزایی بر ارزش آن دارد؛ تعاملات و تجربیات مشتریان می‌تواند به شکل‌گیری، تغییر یا تقویت روابط مثبت یا منفی منجر شود که این امر می‌تواند تأثیرات متفاوتی داشته باشد. مواردی که به برند مرتبط می‌شوند، شامل محصولات ویژه، نام‌های تجاری و موارد مشابه است. این ارتباطات در سه دسته اصلی زیر مجموعه نام تجاری قرار می‌گیرند: قابلیت‌ها، منافع و عملکردها. اعتماد به برند، به معنای اطمینان به پایداری و قدرت آن در بازار است. اعتماد، زیربنای توسعه پایدار کسب‌وکار و روابط است. اعتماد به عنوان امیدی عمومی تعریف می‌شود که فرد می‌تواند بر کلام دیگری حساب کند. اعتماد به برند، تما[illegible]اء بر توانایی برند در ایفای نقش مورد انتظار خود است.

رفتار مشتری

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279



شکل ۳. طراحی برای همکاری در راهبردهای بازاریابی

ثبات درک شده به عنوان یک معیار ارزیابی در ارتباط با ارزش‌گذاری بر برتری یک محصول یا نام تجاری مطرح است. وجود ثبات درک شده پایدار، اعتماد مصرف‌کنندگان را به نام تجاری انتخابی تقویت می‌کند، نام تجاری را از سایر رقبا در بازار متمایز سازی می‌کند، امکان دریافت قیمت بیشتری را برای شرکت فراهم و به توسعه تجارت کمک می‌کند. کیفیتی که در محیط ارزش نام تجاری مبتنی بر مشتری به نمایش گذاشته می‌شود، به عنوان یک عنصر مهم شناخته می‌شود. این کیفیت، تأثیری مهم بر ارزش نام تجاری دارد و به کاهش ریسک کمک می‌کند که یکی از عوامل اصلی برای ثبات درک شده به شمار می‌رود. در صورتی که اصول پیش‌بینی رفتار مشتریان مشخص شده‌اند، در صورتی که میانگین نرخ S وب در کل وب‌سایت به شدت بالا یا پایین باشد، ممکن است نتایج نامطلوبی حاصل شود که این امر می‌تواند برای مصرف‌کنندگانی که نیازمند ارتقاء عملکرد هستند، مشکل‌ساز باشد BI-AIBT. از یک سیستم برای بهینه‌سازی فرآیند واقعی بر اساس پیش‌بینی‌های سیستم X(m) و محاسبات استفاده می‌کند تا اطمینان حاصل شود که وزن‌های هر مصرف‌کننده به طور متوازن تنظیم شوند، مطابق با فرمول‌های زیر:

$$X(m) = P^{(m)} + \left(z^{(n)} - x^T h^{(m)}\right).Int_n(a)S_n(a) \quad (5)$$

اکنون مقدار P(m) در معادله نشان داده شده است. (۶)

$$P^{(m)} = \exp\left(-\frac{(H^{(n)} - h)^2}{2n^2}\right) \qquad \text{Where } x = (X^T PH)^{-1} h^T PM \quad (6)$$

در معادله شماره (5) ، P(m) به عنوان شاخصی برای ارزیابی نرخ خطا در عملکرد کلی یک سازمان تعریف می‌شود. همچنین، معادله شماره (۶) تابعی به نام z(n) را معرفی می‌کند که قابلیت تنظیم مقدار n2را دارد تا بتواند مقدار PM را محاسبه کند.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

شکل ۴. معماری اشتراک داده حفاظت از داده های کنترل شده توسط کاربر، (الف): نرخ نگهداری غیر ثابت

این ماتریس خاص می‌تواند در تعیین ضرایب مرتبط با وب به کار رود و Intn(a) را به عنوان یک مدل پیش‌بینی منحصر به فرد تولید کند. xTبه عنوان نماینده‌ای از گزینه‌های خرید آنلاین، h(m)به عنوان نماینده‌ای از فعالیت‌های تبلیغاتی در شبکه‌های اجتماعی، و Sn(a) به عنوان نماینده‌ای از توزیع سریع‌تر محصولات به مشتریان عمل می‌کند. شکل ۴ معماری اشتراک‌گذاری داده‌های کنترل‌شده توسط کاربر را برای حفاظت از حریم خصوصی نشان می‌دهد. در این فرآیند، اطلاعات مشتریان فعلی هرگز در بلاکچین فاش نمی‌شوند، بلکه ابتدا رمزنگاری شده و در ذخیره‌سازی خارج از زنجیره نگهداری می‌شوند. صاحبان داده می‌توانند اطلاعات را مستقیماً از برنامه‌های کاربردی مشتریان در این ذخیره‌سازی‌ها ثبت کنند. الزامات دسترسی به داده‌ها در قراردادهای هوشمند کدگذاری شده و به همراه اطلاعات پایگاه داده و هش‌ها در شبکه بلاکچین منتشر می‌شوند که این امر از داده‌ها در برابر دستکاری توسط نهادهای مرکزی محافظت می‌کند. در روش آدرس‌دهی مبتنی بر محتوا، محتوا با استفاده از نام‌ها آدرس‌دهی می‌شود که معمولاً شامل نام تأمین‌کننده، نام محتوا و یا نسخه است. این روش به عنوان شناسه‌ای برای بازیابی هش‌های داده به کار می‌رود و تنها با فراخوانی‌های موثر قرارداد، کلید رمزگشایی اطلاعات کاربر آزاد می‌شود، در حالی که قراردادهای هوشمند برای دسترسی به اطلاعات کاربر توسط داده‌های مورد نیاز فراخوانی می‌شوند. بلاکچین، به عنوان یک دفتر کل توزیع‌شده، امکان ذخیره‌سازی داده‌ها در یک شبکه غیرمتمرکز را فراهم می‌آورد. این فناوری از الگوریتم‌های رمزنگاری پیشرفته برای تضمین امنیت و حفظ حریم خصوصی استفاده می‌کند. هر بلاک در زنجیره حاوی یک هش منحصربه‌فرد است که به بلاک قبلی متصل می‌شود، این امر باعث می‌شود تغییر یا دستکاری داده‌ها بدون اطلاع سایر شرکت‌کنندگان در شبکه غیرممکن باشد. قراردادهای هوشمند که بر روی بلاکچین اجرا می‌شوند، به کاربران اجازه می‌دهند تا تراکنش‌ها و توافق‌های خود را بدون نیاز به واسطه‌ها انجام دهند. این قراردادها می‌توانند شرایط خاصی را تعریف کنند که باید قبل از اجرای تراکنش‌ها برآورده شوند و این امر به افزایش شفافیت و کاهش احتمال تقلب کمک می‌کند. در مورد ذخیره‌سازی داده‌ها، بلاکچین می‌تواند با سیستم‌های ذخیره‌سازی آف‌چین ترکیب شود تا حجم بزرگی از داده‌ها را به صورت امن ذخیره کند. داده‌های آف‌چین می‌توانند شامل اطلاعات حساس یا حجم‌های بزرگی از داده‌هایی باشند که نیاز به دسترسی سریع دارند. این داده‌ها می‌توانند با استفاده از کلیدهای رمزنگاری که در بلاکچین ذخیره شده‌اند، رمزگشایی شوند. بلاکچین‌ها معمولاً از الگوریتم‌های هش مانند SHA-256 برای ایجاد هش‌های منحصربه‌فرد استفاده می‌کنند. این الگوریتم‌ها از ویژگی‌هایی مانند اثر بهمنی برخوردار هستند که در آن تغییرات کوچک در ورودی‌ها می‌تواند منجر به تغییرات بزرگ در خروجی‌ها شود و این امر به افزایش امنیت کمک می‌کند. در نهایت، بلاکچین امکان ایجاد یک سیستم مولتی‌چین را فراهم می‌آورد که در آن داده‌های مختلف می‌توانند در بلاکچین‌های متعدد ذخیره شوند تا امنیت و کارایی را بهبود بخشند. این سیستم‌ها می‌توانند به صورت موازی کار کنند و از توانایی‌های مکمل یکدیگر بهره ببرند.

در راستای بهینه‌سازی مدیریت داده‌ها، داده‌های جمع‌آوری شده به صورت انتخابی در مکان‌هایی خارج از بلاکچین ذخیره می‌شوند تا از فضای ذخیره‌سازی و پهنای باند به نحو اثربخشی استفاده شود. سیستمBI-AIBT ، راهکارهایی برای دسترسی به داده‌های غیرمتمرکز از طریق پلتفرم MultiChain ارائه می‌دهد. این فرآیندها شامل رمزنگاری داده‌های کاربران، ذخیره‌سازی اطلاعات رمزنگاری شده به صورت محلی، ثبت تعهدات هش فایل در بلاکچین، جست‌وجوی داده‌های حیاتی، تأیید صحت داده‌ها و تأمین اطلاعات است. دفتر کل به عنوان سابقه‌ای برای سازماندهی و خلاصه‌سازی معاملات به کار می‌رود. بلاکچین MultiChain امکان

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

به اشتراک‌گذاری داده‌های مشتریان میان شرکت‌های مختلف را فراهم می‌کند که این امر به توزیع جانبی اطلاعات و اشتراک‌گذاری آن‌ها با مشتریان از طریق قراردادهای هوشمند منجر می‌شود. سرورهای مدیریت کلیدی، امنیت کلیدهای رمزنگاری شده را در تمام چرخه عمر آن‌ها تضمین می‌کنند و بر توسعه، استفاده، بازیابی، بایگانی و حذف فناوری‌های مرتبط با کلیدهای رمزنگاری، نظارتی دقیق دارند. تصمیم‌گیری اجماعی به عنوان روشی نوآورانه برای توسعه راهبردهای سازمانی مطرح است که در آن کارکنان با ارائه بازخورد و پیشنهادها به توافق می‌رسند و هدف از این روش، دستیابی به راه‌حل‌هایی است که منافع تمامی کارکنان را تأمین کند.
(7)

مدل‌سازی طول عمر مشتریان Tracus در یک محیط تجویزی که نشان‌دهنده یک نرخ نگهداری متغیر است در معادله (۷) تعریف شده است. مطابق با معادله (۷) و شکل ۴(a)، نرخ نگهداری متغیر به عنوان O(N)، تابع بقای مشتریان، به عنوان یک متغیر مهم در تحلیل داده‌های زندگی مشتریان محسوب می‌شود.n Ln نمایانگر ارتباط مشتری با سازمان، Qn برای پیش‌بینی الگوهای رفتاری مشتری در آینده، و Vn برای یکپارچه‌سازی ارتباطات با ارائه‌دهندگان خدمات در چارچوب مدیریت ارتباط با مشتری (CRM) است.WSt نمایانگر میزان احتیاط در تحلیل‌های تجاری است و (1+h) نشان‌دهنده تأثیرات جهانی حاصل از بازارهای مختلف است. مدل بازاریابی با هدف ارزیابی ارزش سهام در آینده طراحی شده است، به گونه‌ای که پیش‌بینی می‌شود سود سهام پس از دوره‌ای از رشد فوق‌العاده به حالت پایدار بازگردد. شکل ۵ چارچوب مفهومی کسب‌وکار را به تصویر می‌کشد با رشد فناوری‌های نوین و مدل‌های کسب‌وکار پیشرفته، تضمین می‌شود که این مدل‌ها قابلیت اجرا در هر مکانی - چه فیزیکی، دیجیتالی یا مجازی - و با هر فردی در هر نقطه‌ای و در هر اکوسیستمی از مدل‌های کسب‌وکار داشته باشند.

شکل ۵. چارچوب مفهومی

محله مالی به عنوان یک بستر از روابط هماهنگ‌کننده عمل می‌کند و شامل عناصری مانند منابع طبیعی در اکوسیستم کسب‌وکار است. این محله مالی علاوه بر ارائه محصولات با ارزش برای مشتریان، که خود بخشی از اکوسیستم هستند، شامل تأمین‌کنندگان، رقبا، سهامداران و تولیدکنندگان نیز می‌شود. استراتژی اطلاعاتی یک سازمان، بازار هدف، نیازهای آن بازار و نقش محصولات یا خدمات در برآورده ساختن این نیازها را توصیف می‌کند. توسعه استراتژی فرایندی است که در آن سازمان طرح‌های خود را تغییر می‌دهد. بنابراین، گروه‌های مدل کسب‌وکار باید به سرعت دیدگاه‌های خود را تغییر دهند و با رشد فناوری‌های مدل‌سازی کسب‌وکار، تنظیم استراتژیک انجام دهند. شرکت‌ها باید ساختاری را توسعه دهند که به آن‌ها امکان یادگیری نحوه کارکرد و اجرای مدل‌های کسب‌وکار را بدهد .کسب‌وکارها باید همواره آگاه باشند که مدل‌های کسب‌وکارشان چگونه عمل می‌کنند و چگونه می‌توان آن‌ها را طراحی کرد. در نهایت، کسب‌وکارها باید درک کنند که مدل‌های کسب‌وکار می‌توانند به صورت عملی اجرا شوند، چه بر پایه فناوری‌های پیشرفته باشند یا نه. در این راستا، فناوری‌های جدید به عنوان یک ابزار کمکی عمل می‌کنند.

هوش مصنوعی (AI) و واقعیت مجازی (VR) امکانات جدیدی را برای تحلیل و ارزیابی شرکت‌های پیشرو و پیشگام فراهم می‌آورند. برای استفاده از این فناوری‌ها، شرکت‌ها باید ابتدا با مفاهیم اساسی مدل‌های کسب‌وکار (BM) و نوآوری در مدل‌های کسب‌وکار (BMI)آشنا شوند. در غیر این صورت، درک و به‌کارگیری این مفاهیم در محیط‌های نوآورانه دشوار خواهد بود. مدل محاسباتی ما شامل چهار بخش اصلی است: واسط کاربری انسانی، واسط کاربری ماشینی، تحلیل الگوهای مدل‌های کسب‌وکار و اکوسیستم مدل‌های کسب‌وکار. این چهار جزء به صورت جداگانه تعریف و با مثال‌های مرتبط برای هر بخش و مطالعات موردی ارائه شده‌اند.

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

این رویکرد مفهومی با هدف ارائه دیدگاهی کلی در ترکیب با درک عمیق از فرصت‌های دیجیتالی کردن مدل‌های کسب‌وکار طراحی شده است. برای این منظور، از نمایش دیجیتالیBM ها استفاده می‌شود. در حالی که برخی از اجزای مدل مفهومی می‌توانند در BMهای غیر دیجیتالی نیز مورد استفاده قرار گیرند، هدف اصلی مدل نظری، دیجیتالی کردنBM ها است. واقعیت کسب‌وکار به ما امکان می‌دهد تا محدوده‌ای جدید و هیجان‌انگیز از محصولات را کشف کنیم. این شامل یک شبیه‌سازی یادگیری مبتنی بر بازی است که جریان‌های مالی یک شرکت را شبیه‌سازی می‌کند. ارزش مشترک، مفهومی است که در آن شرکت‌ها باید ارزش اقتصادی را از طریق فرآیندهایی که به حل نیازها و چالش‌های جامعه کمک می‌کنند، ایجاد نمایند. این رویکرد، دیدگاه جدیدی را به ارزش متقابل در میان مدل‌های کسب‌وکار ارائه می‌دهد. واسطه‌ها، خریداران و فروشندگان محصولات، خدمات و دارایی‌ها را بدون نیاز به مالکیت، ترکیب می‌کنند. آن‌ها به جای خرید و توزیع محصولات به عنوان عمده‌فروشان یا توزیع‌کنندگان، معمولاً درصدی از کل معامله را به عنوان پاداش دریافت می‌کنند. روش BI-AIBT پیشنهادی، بهبود در نسبت‌های پیش‌بینی تقاضا، کیفیت محصول، توسعه کسب‌وکار، تحلیل رفتار مشتری و رضایت مشتری را هدف قرار می‌دهد.

۴. بحث و نتیجه‌گیری:

این تحقیق، اجرای آزمایش BI-AIBT را بر اساس داده‌های عددی مورد بررسی قرار داده است. مطالعه حاضر، متغیرهایی چون نسبت پیش‌بینی تقاضا، سطح کیفیت محصول، رشد کسب‌وکار، تحلیل رفتار مشتریان و میزان رضایت آن‌ها را مورد تجزیه و تحلیل قرار داده است. شکل ۶، نسبت پیش‌بینی تقاضا را به تصویر می‌کشد. پیش‌بینی تقاضای بازار، فرآیندی است که از طریق آن می‌توان تقاضای آتی یک محصول را ارزیابی کرد. استفاده از داده‌های تاریخی به عنوان معتبرترین روش برای پیش‌بینی تقاضا شناخته شده است. با بررسی سیستم‌های کنترل سفارشات و داده‌های فروش، مشتریان می‌توانند الگوها و روندهای موجود را شناسایی کنند. این امکان وجود دارد که محصولات جدید برای مدت‌های طولانی به فروش برسند. پیش‌بینی‌های دقیق، توانایی شرکت‌ها را در پیش‌بینی و مدیریت نوسانات درآمدی تقویت می‌کند. داشتن پیش‌بینی‌های دقیق تقاضا برای بازاریابی، برای شرکت‌ها در حوزه‌های تخصصی، از اهمیت بالایی برخوردار است. تنها از طریق تحقیقات دقیق و تصمیم‌گیری‌های مبتنی بر داده، می‌توان سازمان‌ها را در حفظ موجودی مناسب یاری کرد. این موضوع می‌تواند در آینده توسط کاربران مورد بحث قرار گیرد. اکنون به بررسی دقیق‌تر روش‌های مختلف پیش‌بینی بازار می‌پردازیم. پیش‌بینی تقاضا، روشی مبتنی بر مشاهده است که برای ارزیابی تقاضای محصول در طول زمان به کار می‌رود. به طور خلاصه، این روش به مشتریان امکان می‌دهد تا درآمد بالقوه را پیش‌بینی کرده و برنامه‌ریزی لازم برای مدیریت موجودی و تقاضا را انجام دهند ضمن اینکه اهمیت پیش‌بینی تقاضا با گذشت هر سال افزایش می‌یابد. جدول ۱، سطح کیفیت محصول را نمایش می‌دهد. فرآیند تدارکات، با توجه به شرایط متفاوت مشتریان برای محصولی معین دنبال می‌شود. در برخی کسب‌وکارها، کارایی، ثبات کیفیت محصولات و ایمنی، از قیمت مهم‌تر هستند. یک ویژگی مشترک در مفاهیم کسب‌وکار، هماهنگی ثبات کالا یا خدمات با نیازهای مصرف‌کنندگان است - ویژگی‌هایی که نیازهای مشخص یا ضمنی را برآورده می‌سازند. محصولات با کیفیت بالا، رضایت مشتری را به طور مؤثر ترویج می‌کنند و به افزایش تمایل به خرید مجدد منجر می‌شوند. اگر مشتریان کیفیت یک کالا را مطلوب ارزیابی کنند، آن را به طور مکرر خریداری و به دیگران توصیه می‌نمایند.

جدول ۱. نسبت کیفیت محصول

تعداد محصولات	BDA	SMEs	BI-AIBT
۱۰	۶۳	۶۷	۷۹
۲۰	۵۸	۷۲,۶	۷۸
۳۰	۶۲	۷۴	۸۲

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

۴۰	۶۹,۱	۷۸	۸۶
۵۰	۴۴	۷۳	۸۰,۸
۶۰	۶۷	۷۰	۷۸
۷۰	۶۵	۷۵	۸۷
۸۰	۵۳	۷۱,۲	۷۹
۹۰	۷۲,۵	۷۶	۷۸
۱۰۰	۵۵	۷۷	۹۸,۳

شکل ۶. نسبت پیش بینی تقاضا

تولید محصولاتی با استانداردهای بالا از اهمیت ویژه‌ای برخوردار است. استاندارد محصول در بازار برای جلب رضایت مشتری، کاهش خطر و کاستن از هزینه‌های ناشی از حذف محصولات نامرغوب ضروری است. تولید محصول نامطلوب توسط یک شرکت ممکن است مانع دسترسی مصرف‌کنندگان به محصولات مناسب شود. برای اطمینان از رشد و موفقیت پایدار در تجارت الکترونیک، رسیدگی به مسائل مربوط به استاندارد کالاهای عرضه شده برخط امری حیاتی است. فناوری بلاکچین به عنوان یک ابزار مهم در توسعه کسب‌وکارها مطرح شده است. این فناوری نه تنها در معاملات مالی بلکه در کاربردهای تجاری متنوعی مورد استفاده قرار می‌گیرد. بلاکچین امنیت را تقویت و به تسهیل اشتراک‌گذاری دانش کمک می‌کند، در حالی که پاسخگویی را حفظ می‌نماید. این فناوری در حال تقویت حضور خود در بازار تجاری است تا به شرکت‌ها در جذب مشتریان بیشتر یاری رساند. بلاکچین در افزایش امنیت نقش مهمی دارد و برای پذیرش کامل توسط جامعه تجاری به زمان بیشتری نیاز دارد. بلاکچین هزینه‌ها و زمان نیازمند برای واسطه‌گری را کاهش داده و به تقویت اعتماد در اکوسیستم تجاری کمک می‌کند. تأثیر این فناوری در تحلیل‌های آتی و شرکت‌هایی که به کاربرد بلاکچین می‌پردازند، منعکس می‌شود. در حالی که کاربرد فناوری‌های نوین امیدبخش است، هنوز تجزیه و تحلیل‌ها و تجربیاتی در بهینه‌سازی مدل‌های کسب‌وکار موجود و ایجاد مدل‌های تازه وجود دارد. تحلیل رفتار مشتریان شامل الگوهای اجتماعی، فرکانس استفاده و تأثیر آن‌ها بر تصمیمات خرید می‌شود. شرکت‌ها با ارزیابی رفتار مشتریان، به دنبال توسعه کالاها و خدماتی هستند که به درک بهتر اهداف جمعیتی کمک کند. رفتار خرید مصرف‌کنندگان به فعالیت‌هایی اشاره دارد که مشتریان قبل از خرید یک محصول یا خدمات انجام می‌دهند. درک و تحلیل رفتار مصرف‌کننده، اساس تدوین راهبردهای بازاریابی در عصر دیجیتال است. استراتژی BI-AIBT، که شامل استفاده از موتورهای جست‌وجو و پوشش رسانه‌های اجتماعی است، به شرکت‌ها امکان می‌دهد تا بازاریابی خود را با رویکردهای موفق گذشته همسو سازند. با پیشرفت فناوری‌های دیجیتالی، مصرف‌کنندگان با تبلیغات متنوع برندهای برخط مواجه می‌شوند که این امر بر الگوهای خرید آن‌ها تأثیر می‌گذارد تا جایی که به یکی از کانال‌های بازاریابی کلیدی تبدیل شده است. در این دوران دیجیتالی، کسب‌وکارها به اهمیت درک نیازها و ترجیحات مشتریان پی برده‌اند.

شکل ۷. نسبت توسعه کسب و کار

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

جدول ۲. نسبت تحلیل رفتار مشتری

تعداد محصولات	BDA	SMEs	BI-AIBT
۱۰	۴۵,۸	۶۱	۸۷
۲۰	۵۸	۷۲	۷۲,۶
۳۰	۵۷	۶۸,۱	۸۰
۴۰	۶۰	۷۲	۸۴
۵۰	۵۱,۴	۶۷	۸۹
۶۰	۵۹	۶۲	۹۱
۷۰	۵۵	۶۹	۷۹,۵
۸۰	۵۷	۷۲,۷	۸۹
۹۰	۶۵	۷۰	۸۱
۱۰۰	۵۲,۷	۶۴	۹۶,۳

شکل ۸. نسبت رضایت مشتری

مشتریان از طریق فروشگاه‌های برخط به خرید اقلام مورد نظر خود می‌پردازند و این امکان را دارند که محصولات را مستقیماً از تأمین‌کنندگان دریافت کنند. رضایت مشتری هرچند که تضمین‌کننده خرید مجدد نیست، نقش مهمی در جذب و حفظ مشتریان ایفا می‌کند. نظرسنجی‌های وفاداری مشتری با هدف جمع‌آوری بازخوردهای دقیق و مستمر از مشتریان طراحی می‌شوند تا به شرکت‌ها در اجرای کمپین‌هایی که مشتریان را جذب کرده و اطمینان از حفظ مشتریان وفادار کمک کنند. مصرف‌کنندگان به دنبال کیفیت و قیمت مناسب هستند و شرکت‌ها با ارائه قیمت‌های رقابتی سعی در جلب وفاداری آن‌ها دارند. در نهایت، BI-AIBTنشان دهنده این است که مفاهیم رضایت و وفاداری مشتری در مرکز بحث‌های شرکت‌ها قرار دارند و این وفاداری به طور مستقیم بر ترجیحات خرید مصرف‌کنندگان تأثیر می‌گذارد.

۵. نتیجه‌گیری:

در عرصه پیشرفت‌های تجاری، تجزیه و تحلیل فناوری دیجیتال اهمیت بیشتری می‌یابد و محافظت از اطلاعات از طریق فناوری بلاکچین تضمین می‌شود. این مقاله، فرآیندهای بازاریابی را تقویت کرده و ارتباطات مطمئن بین مشتریان متنوع را با استفاده از فناوری‌های پیشرفته هوش مصنوعی و بلاکچین حفظ می‌کند. مجموعه‌ای از مشارکت‌کنندگان از دو حوزه بازار متفاوت برای جمع‌آوری داده‌های مشاهداتی کیفی فراهم شده‌اند. BI-AIBT بر استفاده از پلتفرم‌های شبکه‌های اجتماعی برای ایجاد ارتباط بین برندها و مصرف‌کنندگان تمرکز دارد و از تجربیات دیجیتالی و فیزیکی برای تأثیرگذاری بر نگرش‌های مشتریان بهره می‌برد.BI-AIBTتوسط شرکت‌ها مورد ارزیابی قرار گرفته و تفاوت‌ها و شباهت‌ها در ایجاد ارزش، راهبردها و تأثیرات کسب‌وکار دیجیتال مورد بررسی قرار گرفته‌اند. همچنین، BT می‌تواند به تقویت ارتباط بین ظرفیت‌های سازمانی و مهارت‌های کارکنان کمک کند. یافته‌های تجربی نشان می‌دهند که دگرگونی دیجیتال اغلب به عنوان یک عنصر مهم در نظر گرفته شده و به افزایش استراتژی‌های نوآورانه

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

تجاری منجر می‌شود. داده‌های عددی پیشنهادیBI-AIBT ، نسبت پیش‌بینی تقاضا (۹۷/۱%)، نسبت کیفیت محصول (۹۸/۳%)، نسبت رشد تجاری (۹۸/۹%)، نسبت تجزیه و تحلیل رفتار مشتری (۹۶/۳%)و نسبت رضایتمندی مشتری (۹۷/۲%) را نشان می‌دهد.

Reference:

[1] Amin, M., Faragallah, O. S., & El-Latif, A. A. (2010). A chaotic block cipher algorithm for image cryptosystems. Communications in Nonlinear Science and Numerical Simulation, 15(11), 3484–3497.

[2] Andoni, M., Robu, V., Flynn, D., Abram, S., Geach, D., Jenkins, D., et al (2019). Blockchain technology in the energy sector: A systematic review of challenges and opportunities. Renewable and Sustainable Energy .Reviews, 100, 143–174

[3] Arjun, R., & Suprabha, K. R. (2020). Innovation and Challenges of Blockchain in Banking: A Scientometric .(۳)View. International Journal of Interactive Multimedia & Artificial Intelligence, 6

[4] Asghar, M. Z., Subhan, F., Ahmad, H., Khan, W. Z., Hakak, S., Gadekallu, T. R., et al. (2021). Senti-eSystem: A sentiment-based eSystem-using hybridized fuzzy and deep neural network for measuring customer .satisfaction. Software: Practice and Experience, 51(3), 571–594

[5] Belazi, A., Khan, M., El-Latif, A. A., & Belghith, S. (2016). Efficient cryptosystem approaches: S-boxes and .permutation–substitutionbased encryption. Nonlinear Dynamics, 87(1), 337–361

[6] Borah, A., Banerjee, S., Lin, Y. T., Jain, A., & Eisingerich, A. B. (2020). Improvised marketing interventions .in social media. Journal of Marketing, 84(2), 69–91

[7] Feng, Q., He, D., Zeadally, S., Khan, M. K., & Kumar, N. (2019). A survey on privacy protection in .blockchain system. Journal of Network and Computer Applications, 126, 45–58

[8] Filimonau, V., & Naumova, E. (2020). The blockchain technology and the scope of its application in .hospitality operations. International Journal of Hospitality Management, 87, Article 102383

[9] Fu, H., Manogaran, G., Wu, K., Cao, M., Jiang, S., & Yang, A. (2020). Intelligent decision-making of online shopping behavior based on internet of things. International Journal of Information Management, 50, 515–525

[10] Gao, J., Wang, H., & Shen, H. (2020a). Machine learning based workload prediction in cloud computing. In .29th International Conference on Computer Communications and Networks (ICCCN)

[11] Gao, J., Wang, H., & Shen, H. (2020b). Smartly handling renewable energy instability in supporting a cloud .datacenter. In IEEE International Parallel and Distributed Processing Symposium (IPDPS)

[12] Hakala, H., O'Shea, G., Farny, S., & Luoto, S. (2020). Re-storying the business, innovation and entrepreneurial ecosystem concepts: The model-narrative review method. International Journal of .Management Reviews, 22(1), 10–32

[13] Hu, L., Nguyen, N. T., Tao, W., Leu, M. C., Liu, X. F., Shahriar, M. R., et al (2018). Modeling of cloud-.based digital twins for smart manufacturing with MT connect. Procedia manufacturing, 26, 1193–1203

[14] Jan, M. A., Cai, J., Gao, X. C., Khan, F., Mastorakis, S., Usman, M., et al (2020). Security and blockchain convergence with Internet of Multimedia Things: Current trends, research challenges and future directions. .Journal of Network and Computer Applications, Article 102918

[15] Kaur, K., Garg, S., Kaddoum, G., Ahmed, S. H., & Atiquzzaman, M. (2019). Keids: Kubernetes-based energy and interference driven scheduler for industrial iot in edgecloud ecosystem. IEEE Internet of Things Journal, .7(5), 4228–4237

[16] Khelifi, H., Luo, S., Nour, B., Moungla, H., Ahmed, S. H., & Guizani, M. (2020). A blockchain-based architecture for secure vehicular Named Data Networks. Computers & Electrical Engineering, 86, Article .106715

[17] Kumar, G., Saha, R., Buchanan, W. J., Geetha, G., Thomas, R., Rai, M. K., et al. (2020). Decentralized accessibility of e-commerce products through blockchain technology. Sustainable Cities and Society, 62, .Article 102361

[18] Manogaran, G., Alazab, M., Shakeel, P. M., & Hsu, C. H. (2021). Blockchain Assisted Secure Data Sharing .Model for Internet of Things Based Smart Industries. IEEE Transactions on Reliability

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

[19] Manogaran, G., Baskar, S., Hsu, C. H., Kadry, S. N., Sundarasekar, R., Kumar, P. M., et al. (2020a). FDM: Fuzzy-optimized Data Management Technique for Improving Big Data Analytics. IEEE Transactions on .Fuzzy Systems

[20] Manogaran, G., Rawal, B. S., Saravanan, V., Kumar, P. M., Martínez, O. S., Crespo, R. G., et al. (2020b). Blockchain based integrated security measure for reliable service delegation in 6G communication environment. Computer Communications, 161, 248–256. https://doi.org/10.1016/j.comcom.2020.07.020

[21] Manogaran, G., Rawal, B. S., Saravanan, V., Kumar, P. M., Martínez, O. S., Crespo, R. G., et al. (2020c). Blockchain based integrated security measure for reliable service delegation in 6G communication .environment. Computer Communications, 161, 248–256

[22] Manogaran, G., Srivastava, G., Muthu, B. A., Baskar, S., Shakeel, P. M., Hsu, C. H., et al. (2020d). A Response-aware Traffic Offloading Scheme using Regression Machine Learning for User-Centric Large- .Scale Internet of Things. IEEE Internet of Things Journal

[23] Mistry, I., Tanwar, S., Tyagi, S., & Kumar, N. (2020). Blockchain for 5G enabled IoT for industrial .automation: A systematic review, solutions, and challenges

[24] Mechanical Systems and Signal Processing, 135, Article 106382. Morkunas, V. J., Paschen, J., & Boon, E. .(2019). How blockchain technologies impact your business model. Business Horizons, 62(3), 295–306

[25] Mustafa, & Khan, S. (2020). FinTech, Blockchain and Islamic Finance: An Extensive Literature Review. International Journal of Economics and Business Administration, 65–86. https://doi.org/10.35808/ijeba/444. .VIII (Issue 2)

[26] Nguyen, N. T., Liu, B. H., Chu, S. I., & Weng, H. Z. (2018a). Challenges, designs, and performances of a distributed algorithm for minimum-latency of data-aggregation in multi-channel WSNs. IEEE Transactions .on Network and Service Management, 16(1), 192–205

[27] Nguyen, N., Leu, M. C., & Liu, X. F. (2017). Real-time communication for manufacturing cyber-physical systems. In IEEE 16th International Symposium on Network Computing and Applications (NCA) (pp. 1–4). .Cambridge, MA, USA

[28] Nguyen, N. T., Leu, M. C., Zeadally, S., Liu, B. H., & Chu, S. I. (2018b). Optimal solution for data collision .avoidance in radio frequency identification networks. Internet Technology Letters 2018, 1, E49

[29] P, A. K., G, S. S., Maddikunta, P. K., Gadekallu, T. R., Al-Ahmari, A., & Abidi, M. H. (2020). Location .Based Business Recommendation Using Spatial Demand.Sustainability, 12(10), 4124

[30] Pham, D. V., Nguyen, G. L., Nguyen, T. N., Pham, C. V., & Nguyen, A. V. (2020). Multi-Topic Misinformation Blocking With Budget Constraint on Online Social Networks. IEEE access : practical .innovations, open solutions, 8, 78879–78889

[31] Ruan, J., Hu, X., Huo, X., Shi, Y., Chan, F. T., Wang, X., et al. (2019). An IoT-based E-business model of intelligent vegetable greenhouses and its key operations management issues. Neural Computing and .Applications, 32(19), 15341–15356

[32] Sheron, P. F., Sridhar, K. P., Baskar, S., & Shakeel, P. M. (2019). A decentralized scalable security framework for end-to-end authentication of future IoT communication. Transactions on Emerging .Telecommunications Technologies, e3815. https://doi.org/10.1002/ett.3815

[33] Stratan, A., Novac, A., & Vinogradova, N. (2020). Cooperation for Innovation: Opportunities and Challenges .for SMEs (The Case of the Republic of Moldova). LUMEN Proceedings, 14, 01–20

[34] Sun, J., Yan, J., & Zhang, K. Z. (2016). Blockchain-based sharing services: What blockchain technology can .contribute to smart cities. Financial Innovation, 2(1), 1–9

[35] Thuethongchai, N., Taiphapoon, T., Chandrachai, A., & Triukose, S. (2020). Adopt big-data analytics to .explore and exploit the new value for service innovation. Social Sciences, 9(3), 29

[36] Trad, A. (2021). The business transformation framework and enterprise architecture framework for managers
in business innovation: An applied holistic mathematical model. International Journal of Service Science, .Management, Engineering, and Technology (IJSSMET), 12(1), 142–181

[37] Ur-Rehman, A., Gondal, I., Kamruzzaman, J., & Jolfaei, A. (2020). Vulnerability modelling for hybrid .industrial control system networks. Journal of Grid Computing, 18 (4), 863–878

[38] Wang, S., Huang, L., Hsu, C. H., & Yang, F. (2016). Collaboration reputation for trustworthy Web service .selection in social networks. Journal of Computer and System Sciences, 82(1), 130–143

[39] Zhao, J., Xue, F., Khan, S., & Khatib, S. F. (2021). Consumer behaviour analysis for business development. Aggression and Violent Behavior. Article 101591. https://doi.org/10.1016/j.avb.2021.101591

دانشگاه آزاد اسلامی واحد الکترونیکی
مجله مدیریت اطلاعات، امنیت و سیستم ها
DOI:

Print ISSN: 2251-9335
Online ISSN: 2252-0279

Title:

Strategic Development and Innovation in Business Leveraging Artificial Intelligence and Blockchain Technology

Abstract:

In an era where business transformations are occurring at an unprecedented pace, advanced technologies such as Artificial Intelligence (AI) are providing new capabilities to enhance commercial performance. These advancements are revolutionizing corporate interactions with customers and employees through information technology-based services. With the expanding use of AI, businesses must re-evaluate their current strategies and actively seek to discover new market opportunities. With increased focus on research in the field of commercial innovations, blockchain has been proposed as a solution for ensuring data security. This article introduces the AI and Blockchain-based Business Innovation Model (BI-AIBT) to strengthen business processes and ensure secure interactions among diverse customers. The model has been examined using qualitative empirical data from participants in two business sectors. BI-AIBT, by analyzing the impact of information technology usage on value creation, proposals, and business attraction, has demonstrated that blockchain can be effective in enhancing interactions between organizational capacities and employee skills. Experimental results of this model indicate that the transformation brought about by information technology is recognized as a significant element in bolstering business innovation strategies, and the BI-AIBT model enhances ratios of demand forecasting (97.1%), product quality (98.3%), business development (98.9%), customer behavior analysis (96.3%), and customer satisfaction (97.2%).

Keywords: Advanced technologies, Artificial intelligence, Blockchain, Innovation