ارائه یک سیستم تشخیص نفوذ در شبکه با کمک الگوریتم طبقه بندی اصلاح شده مبتنی بر ترکیب ماشین بردار پشتیبان و الگوریتم‌های شبیه سازی تبرید و IQPSO

رحیم اصغری[۱] ، وحید موسوی فرد [*۲] ، منوچهراسدی[۳]

۱- استادیار دانشگاه صنعتی مالک اشتر تهران

۲- دانشجوی کارشناسی ارشد مهندسی کامپیوتر-رایانش امن، دانشگاه صنعتی مالک اشتر تهران

۳- دانشجوی کارشناسی ارشد دفاع سایبری، دانشگاه صنعتی مالک اشتر تهران

چکیده

سامانه‌های تشخیص نفوذ بخشی اساسی از امنیت شبکه محسوب می‌گردد. این سامانه‌ها وقایع و فعالیت‌های شبکه را در سطحی که پیاده‌سازی شده‌اند، جهت تشخیص نفوذ، مورد بازرسی و دیده‌بانی قرار می‌دهند. یکی از انواع این سامانه‌ها، سامانه‌های تشخیص نفوذ مبتنی بر ناهنجاری است. این سامانه‌ها آموزش می‌بینند تا بتوانند ترافیک عادی و غیرعادی در شبکه را تشخیص دهند. این جستار یک مدل از سامانه‌های تشخیص نفوذ مبتنی بر ناهنجاری را با استفاده از یک الگوریتم ترکیبی به نام SA-IQPSO، ترکیب دو الگوریتم تبرید و QPSO بهبودیافته، و الگوریتم ماشین بردار پشتیبان، ارائه داده است. در الگوریتم طبقه بندی اصلاح شده از یک روش انتخاب مشخصه به نام SA-IQPSO استفاده نموده است که دقت و بهینگی بالایی را نسبت به هردو الگوریتم تبرید و IQPSO از خود نشان می‌دهد. در این مقاله کاملاً بر روی مفاهیم و روش‌های پیشنهادی بحث شده و با استدلال بر مفاهیم بهینگی و دقت به انتخاب الگوریتم‌ها پرداخته است. در پایان هم مدلی انتزاعی از سیستم تشخیص نفوذ پیشنهادی ارائه‌شده است که قابل پیاده‌سازی و عملیاتی شدن است.

کلمات کلیدی: سامانه‌های تشخیص نفوذ، ماشین‌های بردار پشتیبان، الگوریتم‌های فرا ابتکاری، انتخاب مشخصه

۱. مقدمه

با عنایت به اینکه سامانه‌های کامپیوتری نقشی اساسی و مهم در جامعه مدرن امروزی ایفا می‌کنند، آن‌ها تبدیل به هدف برای حمله‌کنندگان و سوءاستفاده شده‌اند. لذا ایجاد سامانه‌ای برای محافظت و ایجاد ضوابط مناسب ضروری است. یک سیستم کامپیوتری زمانی آسیب‌پذیر است که یک نفوذ به وقوع می‌پیوندد. نفوذ به هر عملی که به‌درستی، اطمینان یا دسترسی‌پذیری سیستم، خلال ایجاد کند، اطلاق می‌شود. در خط مقدم مقابله بانفوذ تکنیک‌های پیشگیری از وقوع نفوذ و حمله در سامانه‌های کامپیوتری همچون دیوارهای آتش قرار دارد. لکن این تکنیک‌ها از قبیل شناسایی کاربر و حفاظت اطلاعات توسط رمز کردن آن‌ها به‌تنهایی برای مقابله با حملات سایبری کافی نیستند. درحالی‌که سامانه‌های کامپیوتری روزبه‌روز پیچیده‌تر می‌شوند، به دلیل اشکالات و نقص‌هایی که در سطح برنامه‌نویسی و طراحی وجود دارد و همچنین تکنیک‌های متنوع و روزافزون نفوذ و مهندسی اجتماعی، همواره نقاط ضعفی هستند، که پتانسیل نفوذ دارند[1]. بنابراین تشخیص نفوذ یکی از اعضای جدایی‌ناپذیر در حفاظت از سامانه‌های کامپیوتری در مقابل تمامی انواع آسیب‌پذیری‌ها است. سامانه‌های تشخیص نفوذ قابلیت تشخیص نفوذ را با کمک الگوریتم‌های دسته‌بندی دارا است. سامانه‌های تشخیص نفوذ را می‌توان به الگوریتم‌های مختلفی همچون ماشین بردار پشتیبان پیاده‌سازی کرد. البته این الگوریتم به‌تنهایی بهینه نیست و

[*] **(Corresponding author) نویسنده مسئول،**
mail: vahidmousavi@alumni.iust.ac.ir
دانشجوی کارشناسی ارشد دانشگاه صنعتی مالک اشتر

نیاز است از الگوریتم‌های فرا ابتکاری جستجو برای بهینه‌سازی ماشین بردار پشتیبان و یافتن پاسخ‌های بهینه استفاده کرد. الگوریتم‌هایی نظیر تبرید، ژنتیک، بهینه‌سازی ازدحام ذرات، QPSO و IQPSO و بسیاری الگوریتم‌های بهینه‌سازی دیگر. دقت و درستی عملکرد سامانه‌های تشخیص نفوذ به الگوریتم‌هایی که با آن پیاده‌سازی می‌شود، بستگی دارد. به همین دلیل ما از الگوریتم بهبودیافته‌ی QPSO و شبیه سازی تبرید در کنار ماشین بردار پشتیبان بهره جسته‌ایم.

این مقاله در۷بخش، ارائه شده است. در بخش ۲ مروری بر پژوهش های قبل انجام شده است. در بخش ۳ به مطالعه ماشین های بردار پشتیبان و نحوه طبقه بندی آنها پرداختیم. در بخش ۴، نحوه تعیین طبقه‌بندی کننده‌ی بهینه ماشین‌های بردار پشتیبان بر مبنای الگوریتم فرا ابتکاری شبیه سازی تبرید و QPSO بهبودیافته را مطرح نمودیم. دربخش پنجم، بهبودی بر الگوریتم QPSO ارائه شده و در بخش ششم، بهینه‌سازی پارامترهای ماشین بردار پشتیبان با کمک الگوریتم SA-IQPSO انجام گرفت و در نهایت مدل سیستم تشخیص نفوذ پیشنهادی مبتنی بر الگوریتم مدنظر را مطرح کردیم.

۲. پیشینه‌ی پژوهش

پژوهش‌ها و مطالعات بسیاری برای بهبود عملکرد ماشین بردار پشتیبان و همچنین معرفی خود الگوریتم بردار پشتیبان ارائه‌شده است که در اینجا به‌اختصار به بخشی از مطالعات و تکنیک‌های به‌کاربرده شده خواهیم پرداخت.

در سال ۲۰۰۴ ، پینگ-فنگ پای و همکارانش یک سیستم پیش‌بینی بارگذاری جریان الکتریسیته که با الگوریتم ماشین بردار پشتیبان به همراه الگوریتم تبرید پیاده‌سازی شده بود، طراحی کردند و در آن از تکنیک ماشین بردار پشتیبان برای حل مشکل غیرخطی بودن توابع و مجموعه‌های بازگشتی و زمانی بهره بردند و از الگوریتم تبرید برای انتخاب درست و بهینه پارامترها در مدل ماشین بردار پشتیبان استفاده نمودند[2]. در سال ۲۰۱۷ موروگان و همکارانش الگوریتم دسته‌بندی بردار پشتیبان را به‌وسیله الگوریتم تبرید بهینه نمودند، جهت استفاده در تشخیص نفوذ بر پایه تشخیص ناهنجاری در شبکه‌های بی‌سیم. در این سیستم پیشنهادی الگوریتم تبرید مشخصه‌های بهینه‌ی هر گره را در شبکه بی‌سیم را جهت دسته‌بندی و ورودی الگوریتم بردار پشتیبان انتخاب می‌کند تا نفوذها را دسته‌بندی کرده و همچنین از ایجاد ترافیک در شبکه جلوگیری کند. بر پایه این انتخاب بهینه مشخصه‌ها ماشین بردار پشتیبان به‌عنوان طبقه‌بندی مشخص می‌کند یک گره در حالت حمله(غیرعادی) است یا در حالت عادی[3]. در سال ۲۰۱۲ لیل ونلینگ و همکارانش در کنفرانس بین‌المللی کنترل صنعتی و مهندسی الکترونیک مطالعه‌ای درباره بهینه‌سازی ماشین‌های بردار پشتیبان با استفاده از الگوریتم‌های تبرید و الگوریتم بهبودیافته‌ی ازدحام ذرات (QPSO) پرداختند[4].

در سال ۲۰۰۸، شیوه‌ی لینک و همکارانش روشی برای تعیین پارامتر در ماشین‌های بردار پشتیبان و تعیین مشخصه‌های بهینه توسط الگوریتم تبرید ارائه نمودند. در این جستار به مطالعه شناسایی راهی برای یافتن دسته‌بندی بهینه به کمک الگوریتم تبرید پرداخته‌شده که به‌عنوان SA-SVM از آن نام‌برده شده است. و در پایان به مقایسه خروجی‌ها و عملکرد الگوریتم بردار پشتیبان با استفاده از الگوریتم تبرید و بدون آن پرداخته‌شده است و با استدلال از یافته‌های تجربی به این مهم که استفاده از الگوریتم تبرید باعث افزایش کارایی و دقت می‌شود تأکید کرده است[5].

در سال ۲۰۱۱ جواد سلیمی و همکارانش به کاربرد کامپیوتر و روشی نوین برای تشخیص بیماری هپاتیت با استفاده از الگوریتم‌های ماشین بردار پشتیبان و تبرید پرداخته‌اند و کاربردی نوین از این روش طبقه‌بندی بهینه را در علوم پزشکی ارائه داده‌اند[6] . در سال ۲۰۱۶ حسین قرائی و همکارانش یک سیستم تشخیص نفوذ بر پایه الگوریتم ژنتیک و الگوریتم ماشین بردار پشتیبان ارائه دادند در این پژوهش آن‌ها به اشکالات و معایب الگوریتم یادگیری ماشینی همچون ماشین بردار پشتیبان اشاره کردند که نیاز زیاد برای ذخیره‌سازی و دامنه جستجوی حجیم این الگوریتم است و برای حل این مشکل راه‌حلی ارائه دادند. راه‌حل آن‌ها بهره جستن از الگوریتم ژنتیک که پتانسیل بالایی را برای یافتن بهترین جواب در فضای جستجو دارا

است. اولین کار شاخص آن‌ها ارائه‌ی یک تابع برازش جدید برای الگوریتم ژنتیک برای افزایش کارایی سیستم و دومین کار شاخصشان ترکیب این الگوریتم ژنتیک با ماشین بردار پشتیبان برای تشخیص ناهنجاری‌ها بود[7].

در سال ۲۰۱۴، یو رن در مقاله‌ای در مجله‌ی مهندسی نرم‌افزار و کاربردها با ارائه‌ی یک سیستم تشخیص نفوذ مبتنی بر الگوریتم‌های ماشین بردار پشتیبان و AdaBoost الگوریتمی نوین را ارائه داد. در این مقاله الگوریتم AdaBoost برای آموزش ماژول‌ها به سیستم و تولید مدل نهایی سیستم تشخیص نفوذ با تکرار کردن یک سری عملیات به‌روزرسانی وزن نمونه‌ها و مدل تشخیص تا زمانی که به تعداد تکرار یا سطح دقت خاصی دست پیدا کند. درنهایت هم ارزیابی درستی عملکرد الگوریتم با استفاده از مجموعه داده‌ی DARPA99 صورت می‌پذیرد [8].

در سال ۲۰۱۶ این- سئونگ جئونگ و همکارانش، یک روش انتخاب مشخصه‌ی بهینه، جهت تشخیص شش نوع حمله منع سرویس به پایگاه داده‌ی NSL_KDD، ارائه دادند. در این مقاله انتخاب مشخصه به‌عنوان یک مسئله‌ی بهینه‌سازی معرفی شد و با الگوریتم فرا ابتکاری تبرید سعی در حال مسئله‌ی بهینه‌سازی مذکور صورت پذیرفت. در پایان برای ارزیابی و تشخیص درستی و دقت الگوریتم پیشنهادی، با الگوریتم‌های طبقه‌بندی MLP و ماشین بردار پشتیبان که بر روی پایگاه داده‌ی NSL_KDD اجرا شدند و الگوریتم‌های بهینه‌سازی و مقایسه‌ی آن‌ها به این نتیجه رسیدند که الگوریتم پیشنهادی‌شان یعنی تبرید بهترین عملکرد را دارد [9] .

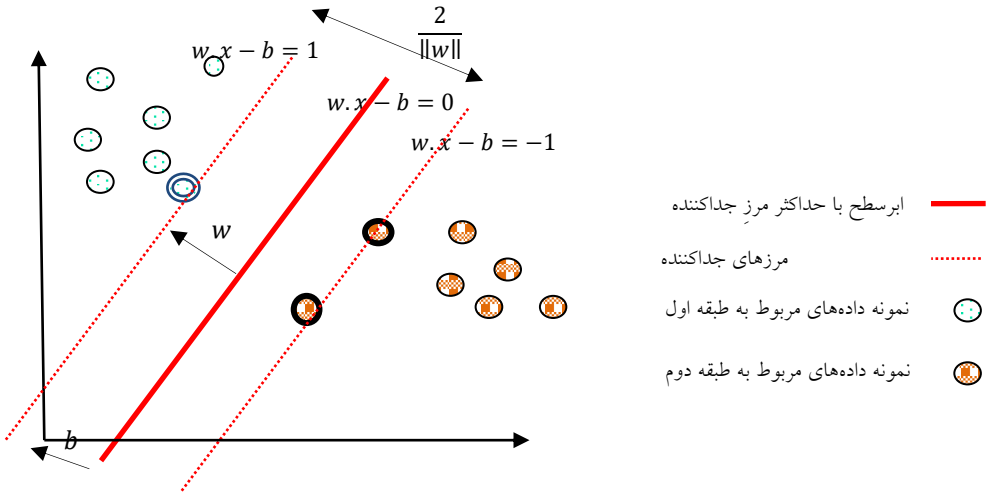
۳. ماشین‌های بردار پشتیبان

ماشین‌های بردار پشتیبان یک روش طبقه‌بندی با نظارت بر مبنای نظریه‌ی یادگیری آماری است. ایده‌ی اصلی یافتن یک ابر صفحه بهینه به‌عنوان سطح تصمیم‌گیری به‌گونه‌ای که حاشیه‌ی بین دو کلاس بیشینه باشد. درصورتی‌که داده‌ها به‌صورت خطی قابل جدا شدن نباشند، داده‌ها با کرنلی غیرخطی به فضای با ابعاد بالاتر منتقل می‌شود و ابر صفحه بهینه تعیین می‌شود. الگوریتم پایه ماشین‌های بردار پشتیبان برای طبقه‌بندی باینری توسعه داده‌شده است. ازآنجایی‌که در بیشتر کاربردها بیش از دو کلاس وجود دارد، الگوریتم‌های مختلفی برای حل مسئله چند کلاسه به کار گرفته می‌شود. یک روش مرسوم تجزیه مسئله به مسئله‌ای با چندین طبقه‌بندی باینری است. الگوریتم‌های «یک در مقابل یک» و «یک در مقابل مابقی» از این جمله است. در روش اول برای هر زوج کلاس یک بردار پشتیبان باینری در نظر گرفته می‌شود و درنهایت همه ماشین‌های بردار پشتیبان با روش رأی‌گیری حداکثر ادغام می‌شوند. روش دوم هر ماشین بردار پشتیبان، داده‌های یک کلاس را از داده‌های کلاس‌های دیگر جدا می‌کند. در این روش برای M کلاس، M طبقه‌بندی‌کننده‌ی باینری خواهیم داشت. پس از طبقه‌بندی داده جدید با M طبقه‌بندی کننده، داده به کلاسی که بیشترین نتیجه مثبت را داشته باشد، نسبت داده می‌شود [7]. مطالعات اولیه نشان می‌دهد، ماشین بردار پشتیبان کم‌ترین مربعات با تابع کرنل پایه‌ای شعاعی، عملکرد قابل قبولی دارد و به دلیل اینکه این الگوریتم مشکل بهینه محلی را حل کرده است، دقت تشخیص بالاتری نسبت به ماشین بردار پشتیبان معمولی دارد [10]. رویکرد ماشین بردار پشتیبان به این صورت است که در فاز آموزش، سعی می‌شود که مرز تصمیم‌گیری به‌گونه‌ای انتخاب گردد که حداقل فاصله آن با هر یک از دسته‌های موردنظر ماکزیمم گردد. این نوع انتخاب باعث می-شود که تصمیم‌گیری ما در عمل، شرایط نویزی را به‌خوبی تحمل کند و پاسخ‌دهی خوبی داشته باشد. این نحوه انتخاب مرز بر اساس نقاطی به نام بردارهای پشتیبان انجام می‌شود. لازم به ذکر است مفاهیمی چون تعمیم‌پذیری یک ماشین شناسایی الگو و بعد VC کاربرد زیادی در مفاهیم ماشین‌های دسته‌بندی دارند.

۳,۱ روش طبقه‌بندی ماشین بردار پشتیبان

فرض کنیم مجموعه نقاط داده $\{(x_1, c_1), (x_2, c_2), \dots, (x_n. c_n)\}$ را در اختیار داریم و می‌خواهیم آنها را به دو طبقه $c_i = \{-1,1\}$ تفکیک کنیم. هر x_i یک بردار p بعدی از اعداد حقیقی است که در واقع همان متغیرهای بیانگر رفتار

نرم افزار هستند. روشهای طبقه بندی خطی، سعی دارند که با ساختن یک ابرسطح (که عبارت است از یک معادله خطی)، داده‌ها را از هم تفکیک کنند. روش طبقه بندی ماشین بردار پشتیبان که یکی از روشهای طبقه‌بندی خطی است، بهترین ابرسطحی را پیدا می‌کند که با حداکثر فاصله داده‌های مربوط به دو طبقه را از هم تفکیک کند. به منظور درک بهتر مطلب، در شکل ۱ تصویری از یک مجموعه داده متعلق به دو کلاس نشان داده شده که روش ماشین بردار پشتیبان بهترین ابرسطح را برای جداسازی آنها انتخاب می کند.



شکل۱. ابرسطح با حداکثر مرزِ جداکننده به همراه مرزهای جداکننده برای طبقه بندی نمونه داده‌های مربوط به دو طبقه متفاوت. نمونه‌های قرار گرفته بر روی مرزها بردارهای پشتیبان نام دارند.

۴. تعیین طبقه‌بندی کننده‌ی بهینه ماشین‌های بردار پشتیبان بر مبنای الگوریتم فرا ابتکاری شبیه سازی تبرید و QPSO بهبودیافته

در طبقه‌بندی ماشین‌های بردار پشتیبان پارامتری بهترین است که پایین‌ترین درصد خطا را داراست. بنابراین لازم می‌نماید که الگوریتم‌های استاندارد بهینه‌سازی به‌صورت مستقیم به همراه الگوریتم ماشین بردار پشتیبان استفاده گردد [4]. در این مقاله ما از ترکیب دو الگوریتم بهینه‌سازی تبرید(شبیه‌سازی ذوب فلزات) و الگوریتم بهبودیافته QPSO (IQPSO) بهره برده‌ایم. در ادامه ایده موردنظرمان را شرح خواهیم داد.

۵. بهبود الگوریتم QPSO

جان سان، وندو ژو و همکارانشان یک مدل جدید از الگوریتم ازدحام ذرات در فیزیک کوانتوم ارائه دادند [4]. این مدل دلتا را به‌عنوان زیرساخت الگوریتم پیشنهادی خود قرار می‌دهد، با توجه به این‌که نقاط یا ذرات ما همانند کوانتوم (ذرات کوانتومی) عمل می‌کنند. در مدل QPSO محل هر ذره با رابطه زیر به دست می‌آید[4]:

$$x_{id}(t+1) = p_{id}^{*}(t) \pm a \times |C(t) - x_{id}(t)| \times \ln(1/rand)$$

$$p_{id}^{*}(t) = rand \times p_{id}(t) + (1 - rand) \times gbest(t)$$
$$C(t) = (C_1(t).C_2(t).\ldots.C_n(t) = \frac{1}{m}\sum_{i=1}^{m} p_{i1}(t).\frac{1}{m}\sum_{i=1}^{m} p_{i2}(t).\ldots.\frac{1}{m}\sum_{i=1}^{m} p_{in}(t) \qquad (1)$$

پارامتر α، نشان‌دهنده میزان انقباض و انبساط است. از این پارامتر برای کنترل درجه‌ی همگرایی الگوریتم استفاده می‌شود؛ m تعداد ذرات را نشان می‌دهد؛ $P_{id}^{*}(t)$ یک مکان تصادفی بین دو مقدار P_{id} و gbest(t) است؛ C(t) مرکز مکان بهینه تمام ذرات فعلی است؛ عملگر $\pm$ برای ایجاد عدم قطعیت و تصادفی بودن در هر بار تکرار الگوریتم است [4].

بنابراین QPSO به‌نوعی مدلی بهبودیافته از PSO است و همچنین فرمول ارزیابی QPSO نیازی به بردار سرعت همچون PSO ندارد، فرمول ارزیابی در QPSO تنها توسط بردار مکان صورت می‌گیرد؛ این شکل از الگوریتم باعث سادگی روند فرمول الگوریتم و درنتیجه خود الگوریتم می‌شود. دلیل ساده شدن الگوریتم این است که تنها به یک پارامتر کنترلی ، α ، نیاز دارد [4]. در مدل QPSO با کاهش خطی انقباض و انبساط فرمول زیر به دست می‌آید:

$$a(t) = \frac{(a_{max}-a_{min})}{t_{max}} \times (t_{max} - t) + a_{min} \qquad (2)$$

در این فرمول α_{max} و α_{min} ، بیشینه و کمینه تابع $\alpha(t)$ هستند، معمولاً مقدار آن بین ۰/۹ و ۰/۴ است؛ t_{max} بیشینه‌ی تعداد تکرار است ؛ t شماره تکرار فعلی است.

بر اساس فرمول بالا تنظیم پارامتر کنترلی الگوریتم (α) بر اساس استراتژی کاهش خطی است، بدین گونه می‌توان درجه‌ی همگرایی الگوریتم را بهتر کنترل کرد[4]، بنابراین پارامتر انقباض و انبساط بهبودیافته است. در حالت اولیه الگوریتم α نسبتاً مقدار بزرگی دارد و در ادامه روند الگوریتم به‌آرامی کاهش می‌یابد، سپس ذره موردنظر در الگوریتم قادر است جستجوی طولانی‌تری را در فضای بزرگ‌تری صورت دهد، در مراحل پایانی الگوریتم مقدار α نسبتاً کوچک می‌شود و بالطبع ذره جستجوی کوتاه‌تری در فضای کمتری را قادر به انجامش خواهد بود؛ این مورد می‌تواند از زودگذری الگوریتم جلوگیری کند. پارامترهای انقباض و انبساط بهبودیافته به‌صورت زیر تعریف می‌گردد[4]:

$$a(t) = \begin{cases} a_{max} - \frac{a_{max}-a_{min}}{u\times(t_{max})^3} \times t^3 & t \le u \times t_{max} \\ a_{min} + \frac{a_{max}-a_{min}}{(1-u)\times(t_{max})^3} \times (t_{max} - t)^3 & t > u \times t_{max} \end{cases} \qquad (3)$$

در این فرمول u یک عدد مثبت کوچک‌تر از یک است. هرچه u بزرگ‌تر باشد برای جستجوی سرسری و هرچه کوچک‌تر، برای جستجوی محلی بهتر است. لذا u به‌نوعی طراحی‌شده است که مقدارش با افزایش تعداد تکرارها، کاهش یابد [4].

۶. بهینه‌سازی پارامترهای ماشین بردار پشتیبان

تنوع پارامتری توابع کلیدی بر روی پیچیدگی توزیع ساده‌ی داده‌ها در فضای مشخصه‌ها در ابعاد بالا، تأثیر می‌گذارد، همچنین بر روی تعمیم توانایی برای به دست آوردن طبقه‌بندی بهینه و بر روی توافقی که بین بیشینه‌ی جریمه مقدار پارامتر C و خطای طبقه‌بندی وجود دارد، تأثیر می‌گذارد. می‌دانیم که هر چه مقدار C بزرگ‌تر باشد، موجب هزینه‌ی ناشی از جریمه برای نمونه‌های خطای طبقه‌بندی می‌شود. برای بهینه‌سازی آسان پارامترها، فاکتور جریمه و پارامترهای توابع کلیدی، همگی را پارامترهای کلیدی می‌نامیم، بدین ترتیب توانایی بهینه‌سازی هم‌زمان را به آن‌ها می‌دهیم [4].

۶,۱ روند الگوریتم ترکیبی از شبیه‌سازی تبرید و QPSO بهبودیافته (SA-IQPSO)

در این بخش، الگوریتم ترکیبی مورد نظر را که در ۱۰ گام طراحی شده است، بیان می نماییم.

گام‌های الگوریتم به‌صورت زیر است:

گام اول: مقداردهی اولیه

این گام شامل مقداردهی اولیه ثابت‌ها، اندازه جمعیت، پارامترهای انقباض و انبساط و مقدار بیشینه تعداد تکرار در الگوریتم IQPSO است. در بخش بعدی نیاز است مقادیر زیر را به دست آوریم: شبیه‌سازی دمای اولیه الگوریتم تبرید، دمای کمینه، ثابت بولتزمان، تعداد تکرار به ازای هر دما، مقدار اولیه پارامترهای الگوریتم ماشین بردار پشتیبان داده‌ی تمرینی ورودی، پیش‌بینی و نرمال‌سازی داده‌ها.

گام دوم: محاسبه‌ی مقدار توافقی اولیه با محاسبات ریاضی در الگوریتم QPSO و به دست آوردن مقدار کمینه‌ی اولیه‌ی انرژی به نام E_old که به بهترین هدف کلی مقداردهی می‌شود. پارامتر E و پارامترهای مربوط به آن را ذخیره‌سازی کند.

گام سوم: اتمام نرمال‌سازی و ارائه‌ی نتایج اگر دما به دمای کمینه رسید یا تعداد تکرارها به بیشینه‌ی مقدار خود رسید، در غیر این صورت به گام چهارم می‌رویم.

گام چهارم: درصورتی‌که به تعداد کافی تکرار به ازای هر دما رسیدیم به گام بعدی می‌رویم، در غیر این صورت به گام دهم می‌رویم.

گام پنجم: ضریب انعطاف‌پذیری متضاد را عوض می‌کنیم و ذرات جدیدی را انتخاب می‌کنیم.

گام ششم: جمع‌آوری مقدار توافقی ذرات حدید، مقدار کمینه برای ورودی الگوریتم تبرید را برای دریافت میزان انرژی به پارامتری به نام E_new بدهیم.

گام هفتم: مقایسه‌ی E_new و بهترین هدف کلی: قبول ذرات جدید، اگر مقدار E_new از مقدار بهترین هدف کلی کم‌تر بود، مقداردهی بهترین هدف کلی با مقدار E_new ذخیره‌سازی بهترین هدف کلی و پارامترهای دیگر در غیر این صورت مقدار پارامترها را همچنان حفظ می‌کنیم.

گام هشتم: قبول حالت جدید اگر مقدار E_new از مقدار E_old کم‌تر بود، در غیر این صورت به گام بعدی می‌رویم.

گام نهم: محاسبه‌ی $P = e^{-\frac{E\_new - E\_old}{kT}}$ اگر مقدار rand از مقدار p کوچک‌تر بود، حالت جدید را بپذیر، در غیر این صورت حالت جدید را رد کرده و گام چهارم برو.

گام دهم: دما به حد پایین خود می‌رسد ، به گام سوم برو.

هر ذره در الگوریتم محاسباتی IQPSO ، برای پارامترهای الگوریتم ماشین بردار پشتیبان، دووجهی است که شامل خود ذره و تابع هسته که تابع RNF است می‌شود.

دمای اولیه بر روی ۱۰۰۰ درجه‌ی سانتی‌گراد برای دقت بالای طبقه‌بندی انتخاب می‌کنیم.

- تعداد مرتبه‌ی تکرار به ازای هر درجه از دما باید تغییر کند و مقدار به رابطه عکس با دمای فعلی و رابطه مستقیم با تعداد کل تکرار الگوریتم دارد که فرمول $a\frac{i}{Tj}$ به دست می‌آید (a یک ثابت عددی، i تعداد کل تکرارها و T_j هم‌دمای فعلی است). این فرمول تعداد تکرارها در دمای مشخصی را تعیین می‌کند. طبق فرمول می‌توان دریافت که در دماهای بالا با تعداد تکرار پایین به حالت تعادل می‌رسد و در دماهای پایین زمان تعداد تکرار و رد جواب افزایش می‌یابد.
- دوره تناوب افت دما با رابطه $T_{i+1} = \mu T_i$ به دست می‌آید. برای مقدار بهینه کلی بهتر است که در ابتدا دما سریع‌تر کاهش پیدا کند، همچنین برای بهینه‌سازی لازم است در ادامه این روند افت دما آرام‌تر صورت پذیرد [4].

برای آزمایش بهینگی استفاده توأم از دو الگوریتم تبرید و IQPSO لی ونلینگ و همکارانش در سال ۲۰۱۲ در مقاله خودآزمایشی را بر روی یک مجموعه داده‌ی معینی از پایگاه داده‌ی UCI انجام دادند تا بهینگی الگوریتم‌های تبرید، IQPSO

و ترکیب این دو را باهم مقایسه کند. آن‌ها به نتایج جالبی دست پیدا یافتند که در زیر به‌اختصار به نتیجه ارزیابی آن‌ها می‌پردازیم. انتخاب داده‌ها از یک سری بیماری‌ها صورت گرفته است که مشخصه‌های هرکدام در جدول۱ به نمایش درمی آید.

Data	Classes	Attribute	Length
Wine	3	13	**178**
Breast-Cancer	2	11	**286**
Iris	3	5	**153**
Hepatitis	2	19	**155**

جدول ۱: داده‌های مورد آزمایش و مشخصه‌های آن‌ها

در مرحله‌ی بعد آن‌ها بر روی این داده‌ها الگوریتم SA-IQPSO برای بهینه‌سازی پارامترهای ماشین بردار پشتیبان، شبیه‌سازی و ارزیابی شد، همچنین بهینه‌سازی را هرکدام از الگوریتم‌های شبیه‌سازی تبرید و IQPSO هم صورت گرفت که نتایج در جدول۲ ارائه شده است.

		شبیه‌سازی تبرید	IQPSO	SA-IQPSO
Wine	C	1.2018	2.1057	**1.5683**
	ϒ	0.7062	1.0639	**0.5814**
	Accuracy	0.9213	0.9326	**0.9326**
Breast-Cancer	C	4.3267	1.9843	**3.6302**
	ϒ	0.5321	0.2874	**0.2593**
	Accuracy	0.9142	0.9406	**0.9615**
Iris	C	63.5913	9.6782	**11.2903**
	ϒ	0.8625	0.6301	**0.7534**
	Accuracy	0.8889	0.9281	**0.9542**
Hepatitis	C	10.2349	3.8346	**5.9284**
	ϒ	0.9816	1.3647	**1.6918**
	Accuracy	0.8903	0.9542	**0.9542**

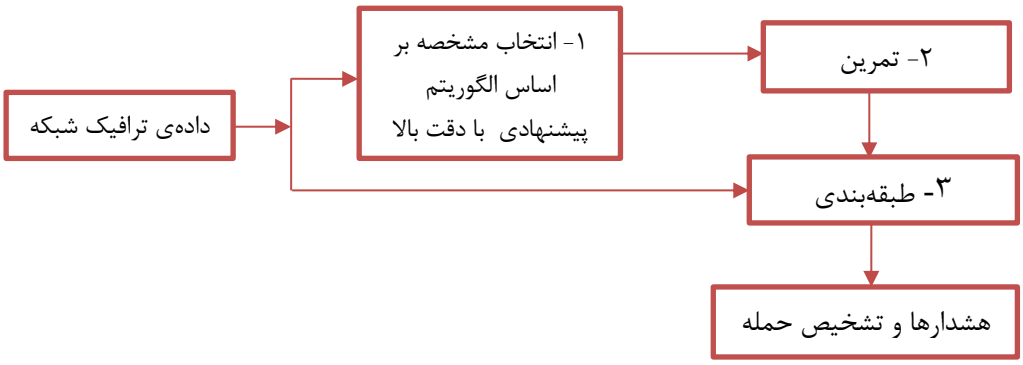
جدول ۲: نتایج شبیه‌سازی و ارزیابی بر روی‌داده‌های آزمایشی

همان‌طور که از نتایج این آزمایش برمی‌آید، بهتر است برای افزایش درستی از ترکیب دو الگوریتم شبیه‌سازی تبرید و QPSO استفاده نماییم. [4]

۷. مدل پیشنهادی سیستم تشخیص نفوذ

سیستم تشخیص نفوذ پیشنهادی قرار است داده‌های جمع‌آوری‌شده در شبکه را، در هر سطحی که سیستم تشخیص نفوذ پیاده‌سازی شده باشد، به الگوریتم ترکیبی پیشنهادی ما بدهد. در الگوریتم ترکیبی پیشنهادی که SA-IQPSO نامیده می شود، برای بهینه‌سازی پارامترها با کمک ماشین بردار پشتیبان مقادیر جمع‌آوری‌شده به الگوریتم بردار پشتیبان داده می‌شود تا ماشین بردار پشتیبان، مقادیر بهینه پارامترهای C و گاما را تعیین کند، بدین گونه که با الگوریتم اصلاح شده SA-IQPSO

عملیات انتخاب بهینه مشخصه‌ها صورت پذیرد و به الگوریتم ماشین بردار پشتیبان برای طبقه‌بندی بهینه بدهد برای دقت بیشتر و درنهایت حالت‌های حمله از عادی با دقت خوبی معین شوند. قابل‌ذکر است با توجه به آزمایشی که ویتهال مانکار و همکارش در سال ۲۰۱۴ انجام دادند، الگوریتم ماشین بردار پشتیبان با استفاده از تابع کرنل RBF عملیات خود را صورت می‌دهد که در آن آزمایش نشان داده می‌شود که این تابع کرنل دقت و بهینگی بالاتری نسبت به توابع دیگر ازجمله تابع خطی و گاوسی، دارد. مدل ارائه شده در شکل۲ مدل پیشنهادی ما برای سیستم تشخیص نفوذ است.



شکل ۲: مدل پیشنهادی سیستم تشخیص نفوذ

۸. نتیجه‌گیری

در این مقاله، هدف ما طراحی یک سیستم تشخیص نفوذ کارآمد مبتنی بر الگوریتم طبقه‌بندی اصلاح شده ماشین‌های بردار پشتیبان با کمک روش پیشنهادی بهینه‌سازی پارامتر و مشخصه که از آن به SA-IQPSO نام بردیم بوده است. الگوریتم کارآمد IQPSO الگوریتم ترکیبی از الگوریتم فرا ابتکاری تبرید و الگوریتم بهبودیافته QPSO است. در استفاده از الگوریتم‌های مذکور با سنجش تمامی جوانب به مدلی بهینه دست‌یافتیم که کاملاً قابل پیاده‌سازی و طراحی و قرار گرفتن در فاز آزمایش است.

۹. مراجع

[1] Manekar and Vitthal; Waghmare, "Intrusion Detection System Using Support Vector Machine(SVM) and Particle Swarm Optimization(PSO)", Kalyani. International Journal of Advanced Computer Research; Bhopal Vol. 4, Iss. 3, (Sep 2014): 808-812.

[2] Ping-Feng Pai and Wei-ChiangHong, "Support vector machines with simulated annealing algorithms in electricity load forecasting", Energy Conversion and Management 46 (2005) 2669–2688.

[3] K. Murugan and Dr. P. Suresh, *"Optimized Simulated Annealing Svm Classifier For Anomaly Intrusion Detection In Wireless Adhoc Network"*, Australian Journal Of Basic And Applied Sciences (2017)

[4] Li Wanling, Wang Zhensheng, Song Xiangjun, *"Parameters Optimization of Support Vector Machine Based on Simulated Annealing and Improved QPSO"*, International Conference on Industrial Control and Electronics Engineering (2012).

[5] Shih-Wei Lin , Zne-Jung Lee, Shih-Chieh Chen and Tsung-Yuan Tseng, *"Parameter determination of support vector machine and feature selection using simulated annealing approach"*, Applied Soft Computing 8 (2008) 1505–1512.

[6] Javad Salimi Sartakhti, Mohammad Hossein Zangooei and Kourosh Mozafari, *"Hepatitis disease diagnosis using a novel hybrid method based on support vector machine and simulated annealing (SVM-SA)"*, computer methods and programs in biomedicine, Volume108, Issue 2 (2012) 570–579

[7] فرهاد صمدزادگان و حدیثه سادات حسنی ، " تعیین ماشین‌های بردار پشتیبان بهینه در طبقه‌بندی تصاویر فرا طیفی بر مبنای الگوریتم ژنتیک" ، فصلنامه فناوری اطلاعات و ارتباطات ایران، سال چهارم، شماره‌های ۱۳، ۱۴ پاییز و زمستان ۱۳۹۱.

[8] Ren, Y. (2014) *"An Integrated Intrusion Detection System by Combining SVM with AdaBoost."*, Journal of Software Engineering and Applications, 7, 1031-1038. doi: 10.4236/jsea.2014.712090.

[9] In-Seon Jeong, Hong-Ki Kim et al. , *"A Feature Selection Approach Based on Simulated Annealing for Detecting Various Denial of Service Attacks"*, (2016) Convergence Security, Vol. 1, 1–18

[10] Hussein Gharaee and Hamid Hosseinvand, "A New Feature Selection IDS based on Genetic Algorithm and SVM", 2016 8th International Symposium on Telecommunications (lST'2016).

[11]ZHENG Shui-bo, TANG Hou-jun, HAN Zheng-zhi etc. *"Sensor Fault Diagnosis Method in ESP System with Support Vector Machines"*, JOURNAL OF SYSTEM SIMULATION, 2005,17(3): 682-684.