تشخیص نفوذ شبکه با استفاده از رویکرد ترکیبی مدل مخفی مارکوف و یادگیری ماشین مفرط

مرضیه نجار [۱]، دانشجوی کارشناسی ارشد؛ محمد حسین معطر [۲]، استادیار

۱- دانشکده فنی و مهندسی- دانشگاه آزاد اسلامی واحد مشهد- مشهد- ایران- marziye.najjar@gmail.com
۲- دانشکده فنی و مهندسی- دانشگاه آزاد اسلامی واحد مشهد- مشهد- ایران- moattar@mshdiau.ac.ir

چکیده: با رشد فناوری اطلاعات، امنیت شبکه به‌عنوان یکی از مباحث چالش‌برانگیز مطرح است. تکنیک‌های تشخیص نفوذ مبتنی بر ناهنجاری یک فناوری ارزشمند برای حفاظت از شبکه‌ها در برابر فعالیت‌های مخرب است. در این مقاله رویکردی جدید مبتنی بر مدل مخفی مارکوف (HMM) و ماشین یادگیری مفرط (ELM) جهت تشخیص نفوذ ارائه شده است. در مدل پیشنهادی، داده‌هایی که از ترافیک شبکه جمع‌آوری شده‌اند، ابتدا پیش‌پردازش می‌شوند. سپس دنباله مشاهدات، به HMM داده می‌شود و مدل با الگوریتم بام-ولچ آموزش می‌بیند. در مرحله شناسایی نفوذ با اعمال الگوریتم ویتربی بر روی مشاهدات به‌دست‌آمده، محتمل‌ترین دنباله حالات استخراج می‌شوند. در مرحله بعد، دنباله حالات به‌عنوان ورودی برای شبکه ELM در نظر گرفته می‌شوند و دسته‌بند داده‌های جدید را با توجه به آنچه آموزش‌دیده به یکی از کلاس‌های نرمال یا حمله نسبت می‌دهد. مجموعه داده مورداستفاده Darpa98 می‌باشد که داده‌های ترافیک شبکه است. مشکلاتی همچون ناکافی‌بودن داده‌های آموزش و اثر کاهش نمونه‌های آموزشی بر صحت نهایی در این مجموعه داده مورد آزمایش قرار گرفته است، که مدل پیشنهادی نتایج بهتری نسبت به روش‌های پیشین ارائه کرده است. آزمایش‌ها نشان می‌دهد که این رویکرد توانسته نسبت به سایر روش‌ها نرخ صحت بالاتر و نرخ مثبت کاذب کمتری را حاصل نماید و کارایی تشخیص نفوذ را بهبود بخشد.

واژه‌های کلیدی: سیستم تشخیص نفوذ، مدل مخفی مارکوف، الگوریتم ویتربی، شبکه عصبی پیش‌خور، ماشین یادگیری مفرط.

Network Intrusion Detection using a Hybrid of Hidden Markov Model and Extreme Learning Machine

M. Najjar[1], M.Sc. Student; M. H. Moattar[2], Assistant professor

1- Computer Engineering Department, Mashhad branch, Islamic Azad University, Mashhad, Iran, Email: marziye.najjar@gmail.com
2- Computer Engineering Department, Mashhad branch, Islamic Azad University, Mashhad, Iran, Email: moattar@mshdiau.ac.ir

Abstract: With the growth of information technology, network security is raised as one of the most important issues and challenges. Anomaly-based intrusion detection system is a valuable technology for network protection against malicious activities. In this paper a new approach is proposed based on hidden Markov model (HMM) and extreme learning machine (ELM) for intrusion detection. In the proposed model, the data that have been collected from network traffic are preprocessed at first. Then, the sequence of observations, are fed into HMM and the model is trained using Baum-Welch algorithm. In the recognition phase, Viterbi algorithm is used and the optimal state sequences are extracted from the input observations. Then, the sequence of states is considered as the input of ELM network and classified to normal or attack classes. Darpa98 dataset which is network trafic data is used to evaluate the approach. We evaluated the approach on this data set for challenges such as insufficient training data and the effect of training samples insufficiency, for which the proposed model provided satisfacory results. Experiments show that our approach has higher accuracy and lower false positive as compared with other methods and the accuracy of the proposed intrusion detection system is 98 percent.

Keywords: Intrusion detection systems, Hidden markov model, Viterbi algorithm, Feedforward neural network, Extreme learning machine.

تاریخ ارسال مقاله: ۱۳۹۶/۰۱/۱۵
تاریخ اصلاح مقاله: ۱۳۹۶/۰۸/۲۳ و ۱۳۹۶/۱۲/۰۹
تاریخ پذیرش مقاله: ۱۳۹۷/۰۱/۰۱
نام نویسنده مسئول: سیدمحمدحسین معطر
نشانی نویسنده مسئول: مشهد، دانشگاه آزاداسلامی واحد مشهد، دانشکده فنی و مهندسی

۱- مقدمه

سیستم‌های تشخیص نفوذ به‌عنوان یکی از عناصر اصلی زیرساخت‌های[1] امنیت در بسیاری از سازمان‌ها شناخته می‌شوند. این سیستم‌ها، مجموعه‌ای از مدل‌ها و الگوهای سخت‌افزاری و نرم‌افزاری هستند که به خودکارکردن فرآیندهای پایش وقایع سیستم‌های کامپیوتری می‌پردازند. سیستم‌های تشخیص نفوذ، وقایع شبکه را برای حل مسائل امنیت سیستم‌ها و شبکه‌های کامپیوتری بررسی و تحلیل می‌کنند. این سیستم‌ها تلاش می‌کنند تا فعالیت‌های کاربر را که به دو صورت عادی و ناهنجاری است و توسط متخصصان و[2] خبرگان، با مقایسه تراکنش‌های شبکه و بر مبنای الگوهای شناخته‌شده طراحی شده است، شناسایی کنند. از قابلیت‌های دیگر سیستم تشخیص نفوذ، امکان تشخیص ترافیک غیرمتعارف از بیرون به داخل شبکه و اعلام آن به مدیر شبکه و یا بستن اتصالات مشکوک است [۱].

نفوذ، مجموعه اقدامات غیرقانونی است که صحت، محرمانگی و یا دسترسی به یک منبع را به خطر می‌اندازد [۲]. سیستم تشخیص نفوذ یک سیستم محافظتی است که ترافیک ورودی به شبکه را آنالیز کرده و خرابکاری‌های در حال وقوع در شبکه را شناسایی می‌کند. به‌منظور پیاده‌سازی روش‌های تشخیص نفوذ، سیستم‌های متعددی تحت عنوان سیستم‌های تشخیص نفوذ طراحی و ساخته شده‌اند. در حوزه امنیت کامپیوتر، سیستم‌های تشخیص نفوذ نقش هشداردهنده را ایفا می‌کنند و هر زمان که امنیت سایت در معرض خطر قرار می‌گیرد، آن‌را اعلام می‌کنند.

همه سیستم‌های تشخیص نفوذ برای تجزیه و تحلیل نیاز به یک مجموعه داده دارند. این داده‌ها می‌توانند از ترافیک شبکه جمع‌آوری شوند. استفاده از روش‌های یادگیری ماشین جهت طراحی روش‌های تشخیص نفوذ مبتنی بر ناهنجاری نیازمند این داده‌های آموزشی برای یادگیری مدل است. در برخی موارد مشکلاتی همچون ناکافی‌بودن داده‌های آموزشی باعث می‌شود که مدل پیشنهادی نتواند به‌خوبی آموزش ببیند و نتایج قابل‌قبولی را ارائه دهد. همچنین در این‌گونه مجموعه داده‌ها که به‌صورت ترتیبی و متوالی هستند، انتخاب یک روش مناسب و کارا که بتواند به‌درستی رکوردهای تراکنش‌های شبکه را مدل کند بسیار حائز اهمیت است. در طراحی سیستم‌های تشخیص نفوذ روش‌های سنتی نمی‌توانند به‌صورت کارا به کشف الگوهای ناشناخته بپردازند. لذا در این موارد از تکنولوژی‌های تصمیم‌گیر هوشمند استفاده می‌گردد تا بتوان الگو یا الگوهایی مؤثر و کارآمد در تشخیص نفوذ را شناسایی نمود [۳].

تاکنون روش‌های متعددی درزمینه طراحی و پیاده‌سازی سیستم‌های تشخیص نفوذ مطرح گردیده است. با توجه به اهمیت موضوع، در این مقاله یک روش جدید مبتنی بر مدل مخفی مارکوف و ماشین یادگیری مفرط ارائه شده است. هدف اصلی در این کار توسعه یک سیستم تشخیص نفوذ با صحت و نرخ تشخیص بالا و حداقل خطا می‌باشد، که با نظارت بر رویدادهای ترافیک شبکه و تشخیص ناهنجاری در آن‌ها، به کشف نفوذ در شبکه بپردازد. روش پیشنهادی بر اساس ساختار سیستم‌های تشخیص نفوذ مبتنی بر ناهنجاری پیکربندی شده است.

در روش پیشنهادی از قابلیت توالی مدل مخفی استفاده شده و ترافیک شبکه به‌صورت ترتیبی بررسی شده است. در روش پیشنهادی بعد از آماده‌سازی داده‌ها، ابتدا در مرحله اول، دنباله و توالی مشاهدات با توجه به نمونه رکوردهای به‌دست‌آمده از ترافیک شبکه و مقادیر ویژگی‌ها، وارد مدل مخفی مارکوف می‌شوند و مدل مخفی آموزش می‌بیند، سپس با اعمال الگوریتم ویتربی بهترین و محتمل‌ترین دنباله حالات، جهت استفاده در مرحله بعدی مدل‌سازی استخراج می‌شوند. در مرحله بعد با استفاده از روش ماشین یادگیری مفرط که نوعی شبکه عصبی پیش‌خور با یک‌لایه مخفی است، دنباله[6] حالاتی که در مرحله قبل به‌دست آمده‌اند به‌عنوان ورودی ELM در نظر گرفته می‌شوند. در این مرحله، شبکه ELM داده‌های ورودی جدید را با توجه به آنچه آموزش‌دیده است دسته‌بندی کرده و به یکی از کلاس‌های نرمال یا حمله نسبت می‌دهد.

دو معیار کلیدی که برای مؤثربودن مدل پیشنهادی جهت تشخیص نفوذ در نظر گرفته شده است شامل موارد زیر می‌گردد:

1. مقاومت در برابر داده‌های ناکافی
2. مدل‌سازی داده‌های ترافیک شبکه (که به‌صورت توالی هستند)

از آنجا که سرعت و تعمیم‌پذیری در تشخیص نفوذ یک مسئله بسیار مهم است، روش ماشین یادگیری مفرط برای این منظور انتخاب شده است. معیارهای دیگری که در روش پیشنهادی مدنظر قرار گرفته‌است، شامل به‌دست‌آوردن نرخ خطای طبقه‌بندی یا مثبت کاذب پایین و صحت بالا می‌باشد.

ساختار ادامه مقاله به این صورت خواهد بود که قسمت دوم شامل مرور ادبیات و پیشینه تحقیق است. در بخش سوم ابزار و روش‌های مرتبط با موضوع مطرح می‌گردد. در بخش چهارم ، روش پیشنهادی با ذکر جزئیات در دو مرحله طراحی مدل مخفی مارکوف و طراحی ماشین یادگیری مفرط معرفی و آنالیز می‌شود و دلایل استفاده از روش پیشنهادی آورده شده است. در بخش پنجم روش پیشنهادی از جنبه‌های مختلف ارزیابی می‌شود و با روش‌های دیگر در قالب نمودارها و جداول حاصل از آزمایش‌ها مقایسه می‌گردد. در پایان نیز به بیان جمع‌بندی و نتیجه‌گیری مقاله و کارهای آینده پرداخته شده است.

۲- مرور ادبیات و پیشینه تحقیق

اولین مطالعه‌ای که در رابطه با لزوم بازرسی خودکار امنیت سیستم‌ها ارائه شد، به سال ۱۹۸۰ بازمی‌گردد [۴]. در سال‌های ۱۹۸۴ تا ۱۹۸۶، Peter Neumann و Dorotty Denning تحقیقاتی در زمینه امنیت سیستم‌های کامپیوتری انجام دادند، و سیستم حاصل IDES نام‌گذاری شد [۵]. ایده مطرح‌شده در این پروژه به‌عنوان پایه خیلی از سیستم‌های نفوذ، که از آن به بعد ایجاد شدند، مورد استفاده قرار

گرفت. تشخیص نفوذ نه‌تنها در سیستم‌های کامپیوتری که در شبکه‌های کامپیوتری نیز از اهمیت بالایی برخوردار است و تحقیقات بسیاری در این زمینه و ازجمله در شبکه‌های حسگر بیسیم انجام شده است [۶]. در [۶] مطالعه از روشی مبتنی بر تئوری بازی‌ها برای این منظور استفاده شده است.

در مقاله [۷] از پرسپترون چند لایه آموزش‌دیده با الگوریتم انتشار به عقب ارتقایافته و انعطاف‌پذیر جهت تشخیص نفوذ استفاده شده است. به‌منظور افزایش سرعت همگرایی یک فاکتور یادگیری بهینه به معادله به‌روزرسانی وزن‌ها اضافه شده است. جهت حل مشکلات تشخیص نفوذ که یک مسئله با ابعاد بالا محسوب می‌شود، در [۸] روشی مبتنی بر مدل مخفی ساده بیز (HNB) معرفی شده است. مدل[۹] مخفی مارکوف (HMM) ازجمله مدل‌های پرکاربرد در حوزه‌های مختلف دسته‌بندی ازجمله شناسایی حروف است [۹]، که در تشخیص نفوذ نیز بسیار مورد استفاده قرار گرفته است. در روش پیشنهادی مقاله [۱۰]، از یک مدل HMM چند متغیره گاوسی برای تشخیص نفوذ استفاده شده است. این مدل شامل یک شبکه خودسازمانده برای خوشه‌بندی رویداد، یک طبقه‌بند مشاهدات، یک آشکارساز رانشی، یک استخراج‌کننده ویژگی، یک مدل مخلوط گوسی (GMM) و یک موتور HMM است. مقاله [۱۱] به بررسی و شناخت سرویس‌های TCP با استفاده از الگوریتم درخت تصمیم J48 می‌پردازد. نویسندگان در [۱۲]، از HMM جهت پیاده‌سازی یک سیستم تشخیص نفوذ برای مدل‌سازی رفتار کاربران استفاده کرده‌اند. با توجه به ماهیت غیرقطعی رفتار کاربر، تصمیم‌گیری در مورد رفتارهای نفوذی یا غیرنفوذی نیز لحاظ شده است و مدل احتمالاتی برای پروفایل کاربر جهت تشخیص حمله پیاده‌سازی شده است.

در مقاله [۱۳]، یک شبکه عصبی چند مرحله‌ای که شامل سه سطح شناسایی می‌باشد پیشنهاد شده است. مقاله [۱۴]، روش ELM را جهت پیاده‌سازی سیستم تشخیص نفوذ پیشنهاد کرده است. این مقاله از ELM پایه جهت استخراج ویژگی و از ELM مبتنی بر کرنل جهت دسته‌بندی استفاده کرده است. نویسندگان این مقاله معتقدند که روش‌های دیگر یادگیری ماشین مثل شبکه‌های عصبی و ماشین بردار پشتیبان معایبی همچون زمان یادگیری زیاد و نیاز به پارامترهای میزان‌سازی دارند و همچنین نمی‌توانند[۲] در مسائل طبقه‌بندی چندکلاسه کار کنند. روش پیشنهادی در [۱۵] نیز از ELM جهت تشخیص نفوذ شبکه استفاده کرده است، با این تفاوت که توانسته نسبت به روش ELM مرسوم که دارای مشکل آموزش دسته‌ای پایین‌رتبه است، بهبود حاصل نماید.[۳]

مقاله [۱۶] استفاده از الگوریتم ژنتیک جهت تشخیص نفوذ مبتنی بر ناهنجاری را پیشنهاد داده است، به‌طوری‌که با استفاده از الگوریتم ژنتیک قوانین جدیدی جهت تشخیص نفوذ یا نرمال‌بودن یک فعالیت تولید می‌شوند. پیشنهاد مقاله [۱۷]، یک سیستم تشخیص نفوذ با استفاده از ماشین بردار پشتیبان (SVM) است. در [۱۸] به‌طور[۴] همزمان انتخاب ویژگی‌ها و انتخاب مدل با روش K نزدیک‌ترین همسایه (KNN) انجام شده است. مقاله [۱۹] یک روش ترکیبی برای مدل‌سازی IDS ارائه داده است. در این روش، دسته‌بندهای C5.0 و HMM به‌عنوان یک مدل هوشمند ترکیبی و سلسله مراتبی باهم ادغام شده‌اند. مقاله [۲۰] روش HMMPayl را پیشنهاد داده است، که در آن سیستم مقدار payload که بیانگر اندازه دنباله بایت‌ها است به‌عنوان ورودی دریافت می‌کند، و تجزیه و تحلیل آن با استفاده از مدل مخفی مارکوف انجام می‌شود.

روش پیشنهادی در مقاله [۲۱]، تشخیص نفوذ با استفاده از رویکرد شبکه‌های بیزین و مدل مخفی مارکوف است. با استفاده از شبکه‌های بیزین احتمالات شرطی تخمین‌زده‌شده و وابستگی میان متغیرها به‌دست می‌آید. بر اساس اطلاعات شبکه، احتمالات انتقال حالت و ماتریس احتمال انتشار به‌دست می‌آید و این پارامترها به‌عنوان پارامترهای HMM برای ایجاد مدل استفاده می‌شود. مقاله [۲۲] روش MARK-ELM را معرفی کرده است که خروجی‌های[۱] مدل را به‌منظور بهبود تشخیص نفوذ شبکه بر روی داده‌هایی که شامل چندین کلاس از حملات هستند، ترکیب می‌کند. در این مقاله، روش ELM به‌عنوان الگوریتم یادگیری انتخاب شده است. در [۲۳] یک روش ترکیبی جدید مبتنی بر ماشین بردار پشتیبان و تحلیل مؤلفه‌های اصلی مبتنی بر هسته (KPCA) همراه با الگوریتم[۶] ژنتیک (GA) برای تشخیص نفوذ ارائه شده است. در مدل ارائه‌شده، یک طبقه‌بند SVM چندلایه برای تخمین این‌که آیا فعالیت موردنظر یک حمله است یا خیر اتخاذ شده است و KPCA به‌عنوان یک پیش پردازنده به‌منظور کاهش ابعاد بردار ویژگی و کوتاه‌شدن زمان آموزش استفاده می‌شود. همچنین، در [۲۴] یک رویکرد جدید داده‌کاوی مبتنی بر شبکه‌های عصبی فازی و ماشین بردار پشتیبان در جهت کمک به تشخیص نفوذ برای دستیابی به نرخ تشخیص[۱] بالا معرفی شده است.

ازجمله جدیدترین کارهایی که با تمرکز بر روی دادگان تشخیص نفوذ KDD Cup99 [۲۵] انجام شده است و از تکنیک‌های یادگیری ماشین بهره برده است، می‌توان به [۲۶] اشاره نمود. در این مقاله از ترکیب SVM و ELM استفاده شده است. همچنین از الگوریتم خوشه‌بندی K-means جهت کاهش مؤثر تعداد نمونه‌ها استفاده شده است. در[۵] مقاله [۲۷] از مدل نقشه خودسازمانده احتمالاتی (PSOM) برای دسته‌بندی استفاده شده است و برای استخراج ویژگی تحلیل مولفه‌های اصلی (PCA) پیشنهاد شده است. در مقاله [۲۸] از ترکیب دسته‌بند SVM و یک نسخه بهبودیافته الگوریتم شبکه کلونی مورچه‌های[۱] خودسازمان‌یافته (SOACN) استفاده شده است.[۹] در مقاله [۲۹] نیز استفاده از یادگیری شبه‌نظارتی مدل ELM برای تشخیص نفوذ پیشنهاد شده است. نقطه ضعف اکثر روش‌های فوق‌الذکر در این است که فرض را بر این گذاشته‌اند که رفتار کاربران در شبکه با استفاده از یک رکورد داده مشابه آنچه در KDD Cup مطرح است، قابل بیان است. درصورتی[۱]که در کاربرد واقعی این‌گونه نیست و

تشخیص حمله باید با توجه به ترافیک شبکه انجام گیرد. در این مقاله، یک روش ترکیبی با کمک مدل HMM به‌منظور مدل‌سازی توالی، از ترافیک شبکه به‌منظور تعیین حمله استفاده می‌شود.

۳- ابزار و روش‌ها

این بخش به معرفی ابزارها و روش‌های مورد استفاده در چارچوب پیشنهادی اختصاص دارد.

۳-۱- مدل مخفی مارکوف

مدل مخفی مارکوف یک مجموعه متناهی از حالات است، که هرکدام با یک توزیع احتمال همراه است. انتقال میان این حالت‌ها توسط مجموعه‌ای از احتمالات به نام احتمالات انتقال حالت کنترل می‌شود. در یک حالت خاص یک نتیجه و یا مشاهده را می‌توان با توجه به توزیع احتمال مربوطه تولید نمود. این خروجی تنها نتیجه است، نه حالت قابل مشاهده برای یک ناظر خارجی و در نتیجه حالات برای خارج "پنهان" هستند. مدل مخفی مارکوف به‌طور گسترده‌ای در کشف دانش، طبقه‌بندی الگو، تشخیص گفتار و مدل‌سازی توالی DNA بکار می‌رود [۹، ۳۰].

یک مدل مخفی مارکوف را می‌توان با تعیین پارامترهای زیر ایجاد نمود:

- تعداد حالات (N): تعداد حالت‌ها در موفقیت مدل نقش به‌سزایی دارد و در یک مدل مخفی مارکوف هر حالت با یک رویداد متناظر است. برای اتصال حالت‌ها روش‌های متفاوتی وجود دارد که در عمومی‌ترین شکل تمام حالت‌ها به یکدیگر متصل می‌شوند و از یکدیگر قابل دسترسی هستند (مدل ارگودیک).
- تعداد مشاهدات در هر حالت (M): تعداد مشاهدات برابر است با تعداد مشاهداتی که سیستم مدل‌شده خواهد داشت. اگر مشاهدات گسسته باشند، آنگاه M یک مقدار محدود خواهد داشت.
- توزیع احتمال انتقال حالات: یک مجموعه از احتمالات انتقال در بین حالت‌ها A=$[a_{ij}]$.

$$a_{ij} = p\{s_{t+1} = j \mid s_t = i\}, \quad 1 \le i, j \le N \qquad (۱)$$

که در آن s_t بیانگر حالت فعلی می‌باشد.

- توزیع احتمال مشاهدات: یک توزیع احتمال برای هر یک از حالت‌ها B=$\{b_j(k)\}$

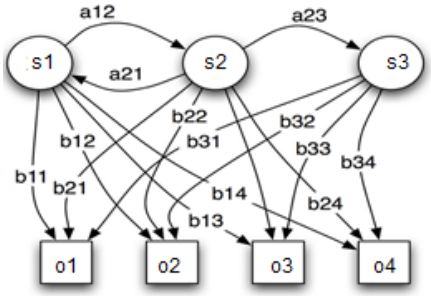
$$b_i(k) = p\{o_t = V_k \mid s_t = i\}, \quad 1 \le i \le N, \quad 1 \le k \le M \qquad (۲)$$

که در آن V_k بیانگر kامین سمبل مشاهده شده است و o_t بیانگر بردار پارامترهای ورودی می‌باشد. به‌بیان دیگر $b_i(k)$ احتمال مشاهده نماد kام در حالت i است.

- توزیع احتمال حالت آغازین $\pi = \pi_i$ که در آن:

$$\pi_i = p\{s_t = i\}, \quad 1 \le i \le N \qquad (۳)$$

مدل مخفی مارکوف می‌تواند فرایندهای پیچیده مارکوف را که حالت‌ها بر اساس یک توزیع احتمالی مشاهدات را نتیجه می‌دهند، مدل کند. در شکل ۱ نمایی از پارامترهای مدل مخفی مارکوف آورده شده است. در شکل ۱ برای مثال a_{23} احتمال انتقال از حالت ۲ به حالت ۳ و b_{23} احتمال مشاهده سمبل ۳ در حالت ۲ است.



شکل ۱: نمایی از پارامترهای احتمالاتی یک مدل مخفی مارکوف ۳ حالته

۳-۲- ماشین یادگیری مفرط

ماشین یادگیری مفرط یک الگوریتم جدید در یادگیری ماشین است که برای شبکه عصبی پیشخور با یک‌لایه مخفی (SLFN) طراحی شده است و به‌طور تحلیلی وزن‌های بهینه را محاسبه می‌کند. این مدل برای طبقه‌بندی و یا رگرسیون با یک‌لایه از گره‌های پنهان به‌کار می‌رود، که در آن وزن اتصال ورودی به گره‌های پنهان به‌صورت تصادفی است. این مدل می‌تواند عملکرد خوبی را فراهم کند و یادگیری هزار برابر سریع‌تر از شبکه‌های آموزش دیده با انتشار به عقب داشته باشد [۲۲].

به‌طور متداول، در یک شبکه پیشخور وابستگی بین لایه‌های مختلف توسط پارامترهای وزن و بایاس ایجاد می‌گردد. روش‌های مبتنی بر گرادیان نزولی معمولاً در آموزش شبکه عصبی پیشخور استفاده می‌شوند. با این حال، به‌طور کلی روش‌های یادگیری بر اساس گرادیان نزولی به‌دلیل مراحل یادگیری نادرست بسیار کند هستند، و یا ممکن است در کمینه‌های محلی گفتار شوند. علاوه بر این، ممکن است برای به‌دست‌آوردن کارایی آموزش بهتر بسیاری از مراحل یادگیری تکراری لازم باشد [۱۴].

دسته‌بند ELM دارای دو مدل مبتنی بر کرنل و پایه است. ماهیت ELM بر این فرض استوار است که برخلاف روش‌های یادگیری رایج در شبکه‌های SLFN نیاز به تنظیم چندانی ندارد. وزن ورودی‌ها و بایاس نرون‌های مخفی به‌صورت تصادفی انتخاب می‌شوند و وزن‌های خروجی لایه مخفی را برای به‌حداقل‌رساندن خطای یادگیری تعیین می‌کنند. وزن‌های ورودی، وزن اتصالات میان لایه ورودی و نرون‌های مخفی هستند. وزن‌های خروجی، وزن اتصالات میان نرون‌های مخفی و نودهای خروجی هستند. پس از این‌که وزن‌های ورودی و بایاس‌های لایه مخفی به‌صورت دلخواه انتخاب شدند، SLFN می‌تواند به‌عنوان یک

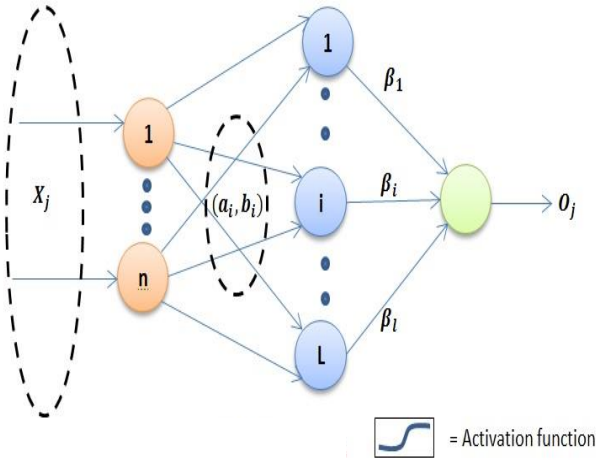
سیستم خطی در نظر گرفته شود و وزن‌های خروجی به‌صورت تحلیلی تعیین شوند [۱۲].

در ELM وزن‌های ورودی W_i و بایاس‌های لایه مخفی b_i، به‌روزرسانی نمی‌شوند، بلکه به‌صورت تصادفی انتخاب می‌شوند و ثابت هستند. این‌کار معادل با نگاشت نمونه‌ها به یک فضای ویژگی تصادفی است. بنابراین، آموزش SLFN معادل با یافتن راه‌حل خطای مربعی حداقل می‌باشد [۱۴]. با توجه به مجموعه داده یادگیری $\{x_i,t_i \mid x_i \in R^n, t_i \in R^m, i=1,....,N\}$، تابع فعال‌سازی $g(x)$ و تعداد نرون k، سه گام اصلی در الگوریتم ELM، می‌توانند به‌صورت زیر خلاصه شوند [۱۲]:

1. وزن ورودی W_i و بایاس b_i برای $i = 1, \dots, N$، با توجه به تابع چگالی احتمال پیوسته، به‌صورت تصادفی مشخص می‌شود.
2. محاسبه خروجی لایه پنهان
3. محاسبه وزن‌های خروجی β

که در آن n تعداد نرون‌های ورودی و i نرون‌های لایه مخفی می‌باشد.

بسیاری از توابع کرنل و فعال‌ساز غیرخطی می‌توانند در ELM به‌کار روند. زمانی‌که ویژگی‌های لایه مخفی برای کاربر نامشخص هستند، می‌توان یک ماتریس کرنل تعریف نمود، که (i,j) امین ورودی آن معادل $k(x_i, x_j)$ است، و $k(.,.)$ یک تابع نیمه‌قطعی مثبت است.



شکل ۲: ساختار شبکه ELM [۱۴]

۴- روش پیشنهادی

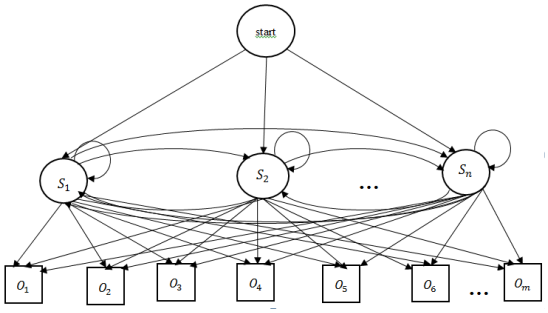
در روش ارائه‌شده، در فاز اول تحقیق، ترافیک شبکه بررسی می‌گردد و داده‌ها استخراج می‌شوند. سپس توالی مشاهدات با توجه به سری زمانی رکوردها در مجموعه‌داده به‌دست آمده و یک مدل مخفی مارکوف طراحی می‌شود. در ادامه با استفاده از الگوریتم ویتربی محتمل‌ترین دنباله حالات استخراج می‌شوند. در فاز بعدی یک بردار از حالت‌ها وجود دارد، که با استفاده از روش ELM ، بهترین حالات به‌دست‌آمده دسته‌بندی و خروجی نهایی (تشخیص نفوذ یا نرمال بودن) تولید می‌شود. روش پیشنهادی با ورود دنباله‌ای از مشاهدات o که هر یک شامل یک رکورد داده از ترافیک شبکه به‌صورت o_m=[Duration,Service,Src_Port,Dest_Port,Src_IP,Dest_IP] است، آغاز می‌شود. مثالی از مقادیر عددی هر فیلد در جدول ۱ مشاهده می‌شود.

جدول ۱: یک رکورد نمونه از مجموعه داده ورودی

Duration	Service	Src port	Dest port	Src IP	Dest IP
۰۰:۰۰:۱۳	ftp	۱۷۵۵	۲۱	192.168.1.30	192.168.0.20

شکل ۴ فلوچارت و مراحل انجام کار در روش پیشنهادی را به‌ترتیب و به‌صورت مرحله‌به‌مرحله نشان می‌دهد.

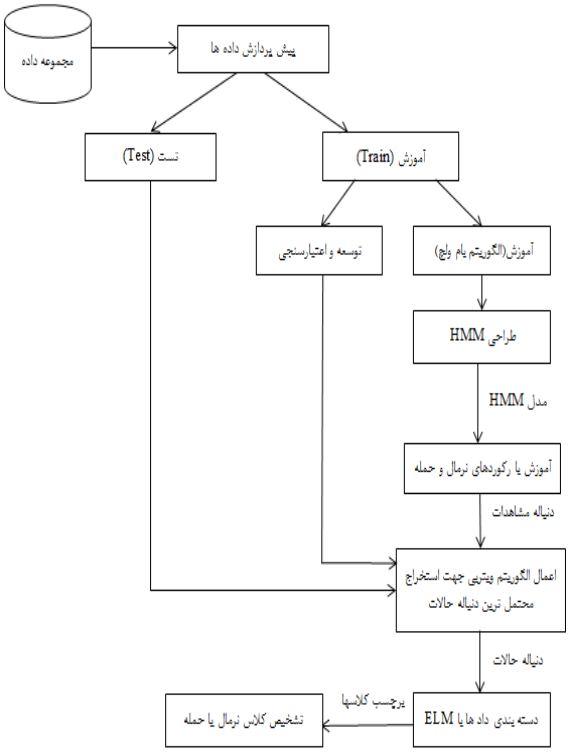
1. گسسته‌سازی داده‌ها: روند تقسیم‌بندی متغیرهای پیوسته به بازه‌های مختلف گسسته‌سازی نامیده می‌شود. این مرحله، گام اول پیش‌پردازش و آماده‌سازی داده‌ها است. برای گسسته‌سازی از روش یکنواخت با زیربازه هایی با طول برابر استفاده شده است. هر متغیر برحسب مقیاس به تعداد مشخصی زیربازه با اندازه برابر تقسیم شده است. البته همان‌طور که در جدول ۱ مشخص است، تنها فیلدduration time نیاز به گسسته‌سازی دارد. برای این منظور مقدار این فیلد که بر اساس ثانیه و دقیقه است، به ثانیه تبدیل شده و سپس به سه زیربازه گسسته‌سازی می‌شود.
2. طراحی HMM: مدل HMM با قابلیت یادگیری توالی، ابزار خوبی در راستای حل مسائل با ماهیت سری زمانی محسوب می‌گردد. در شکل ۳ مدل مخفی مارکوف پیشنهادی با سه حالت مخفی آورده شده است. در این شکل، si حالت‌های مخفی و oiها مشاهدات (مطابق جدول ۱) را نشان می‌دهند.



شکل ۳: ساختار مدل مخفی مارکوف

در مسئله تشخیص نفوذ مورد بررسی، وضعیت هر جلسه به‌صورت مشاهداتی دارای زمان مهر از ترافیک شبکه ثبت می‌شود. مطابق آنچه گفته شد، هر مشاهده خود رکوردی با ۶ ویژگی است. همچنین چون مسئله یادگیری مذکور یک مسئله با ناظر است، هر جلسه دارای برچسب عادی یا حمله (نفوذ) می‌باشد. این بردارهای مشاهدات متغیرهای o_i در شکل ۳ را تشکیل می‌دهند. برای هریک از کلاس‌های حمله یا عادی یک مدل HMM ایجاد می‌گردد. در این ساختار s_i نشان‌دهنده وضعیت دسترسی‌ها می‌باشد، که در دادگان ورودی وجود ندارند و اصطلاحاً مخفی هستند. دنباله‌ای از این دسترسی‌ها

مشخص‌کننده رخداد نفوذ یا نرمال است. همان‌طور که گفته شد، به‌دلیل ماهیت مخفی‌بودن مدل HMM و از طرفی مخفی‌بودن رفتار کاربران سیستم در یک جلسه (اعم از حمله یا عادی)، وضعیت سیستم در هرلحظه از زمان مشخص نیست (به‌عبارتی مخفی است) و تنها تعداد آن‌ها مهم است.



شکل ۴: فلوچارت روش پیشنهادی

در روش پیشنهادی تعداد حالات مخفی در بازه ۳ تا ۵ مورد بررسی قرارگرفته‌اند، و در نهایت حالات ۴ در نظر گرفته شده‌اند (S_1 تا S_4). بعد از آموزش مدل، در گام بعدی از الگوریتم ویتربی جهت استخراج بهترین و محتمل‌ترین دنباله یا توالی حالت‌ها (وضعیت دسترسی‌ها) استفاده شده است.

3. آموزش مدل مخفی مارکوف: نمونه‌های مربوط به کلاس‌های نرمال و حمله برای آموزش HMM در نظر گرفته شده است و سپس HMM با استفاده از الگوریتم مشهور بام-ولچ آموزش دیده است. خروجی الگوریتم بام-ولچ یک ماتریس احتمال انتقال و ماتریس‌های احتمال مشاهدات (یک ماتریس به‌ازای هر حالت) برای هریک از کلاس‌های حمله و نرمال است.
4. استخراج محتمل‌ترین دنباله‌ها: در گام بعدی با استفاده از مدل HMM و به کمک الگوریتم ویتربی محتمل‌ترین دنباله حالات استخراج می‌شود. خروجی این گام دنباله‌ای از حالات بهینه طی‌شده توسط الگوریتم ویتربی به‌صورت s(t) , t=1,2,.. است، که برای سهولت تنها با شماره حالت نشان داده می‌شود. این بردار به لایه ورودی شبکه ELM وارد می‌شود. برای مثال یک خروجی احتمالی این گام به‌صورت {s_2,s_1,s_2,s_3,s_4} است. دنباله فوق به این معناست که ترافیک شبکه ابتدا در حالت (مخفی) s_2 قرار گرفته و سپس به‌شکل ذکرشده حالات مختلف را طی کرده است.
5. طراحی ELM: از این دنباله حالات جهت آموزش دسته‌بند ELM استفاده کرده و شبکه عصبی خود را با نرونهای ورودی که همان دنباله حالات الگوریتم ویتربی در مرحله قبل هستند و تعداد ۵ نرون لایه مخفی طراحی شده است.
6. دسته‌بندی داده‌ها: برای طبقه‌بندی چندکلاسه ترافیک شبکه از ELM استفاده شده است. دنباله‌های ورودی جدید که در مرحله قبل با الگوریتم ویتربی استخراج شدند، به مدل ELM داده می‌شود و بسته ورودی در یکی از حالت‌های حمله که خود شامل سه کلاس حمله می‌شود و یا رفتار عادی قرار می‌گیرد (طبقه‌بندی چندکلاسه).

۴-۱- دلایل انتخاب روش پیشنهادی

یکی از دلایل استفاده از رویکرد ترکیبی HMM-ELM این است که در صورت کافی‌نبودن داده‌های آموزش، تخمین پارامترهای مدل مخفی شاید به‌درستی انجام نشود. زیرا HMM برای محاسبه و تخمین تابع چگالی احتمال، نیازمند داده‌های آموزش زیادی است و با تعداد نمونه‌های کم ممکن است ماتریس مشاهدات تنک شود و تخمین تابع چگالی به‌درستی انجام نشود. از طرفی ELM چنین مشکلی ندارد. ازاین‌رو جهت رفع این مشکل در روش پیشنهادی خود دو رویکرد HMM وELM برای رسیدن به نتایج بهتر ادغام شده‌اند. در واقع از قابلیت توالی مدل مخفی مارکوف برای مدل‌سازی ترافیک شبکه و استخراج ویژگی (محتمل‌ترین دنباله حالات) و از ماشین یادگیری مفرط جهت بهبود مدل مخفی مارکوف و مقاومت در برابر داده‌های کم استفاده شده است. عملاً ELM به‌تنهایی توانایی مدل کردن توالی را ندارد و HMM هم به‌تنهایی نمی‌تواند با داده‌های کم به‌درستی آموزش ببیند ولی این دو روش در کنار هم می‌توانند نواقص یکدیگر را پوشش داده و نتایج خوبی را ارائه دهند.

نوآوری کار در استفاده از یک روش ترکیبی مبتنی بر مدل مخفی مارکوف و ماشین یادگیری مفرط جهت طراحی سیستم تشخیص نفوذ می‌باشد. به‌طوری‌که توسط مدل مخفی مارکوف ویژگی‌ها یا همان دنباله حالات را استخراج کرده و در مرحله بعد از این دنباله حالات جهت آموزش دسته‌بند استفاده شده است.

۵- آزمایش و ارزیابی

برای پیاده‌سازی و ارزیابی مدل پیشنهادی از مجموعه داده darpa98 استفاده شده است که یکی از معمول‌ترین مجموعه داده‌های در دسترس در حوزه تشخیص نفوذ است. این مجموعه داده برای شبیه‌سازی سیستم‌های تشخیص نفوذ به‌کار می‌رود [۳۱].

در این مجموعه داده یک جلسه دنباله‌ای از بسته‌ها است که در[۸] زمان‌های مشخص آغاز و پایان می‌یابد. در طول مدت برقراری جلسه،

تحت یک قرارداد تعریف‌شده، داده‌ها از یک آدرس IP مبدأ به یک آدرس IP مقصد و برعکس، جریان دارند. هر اتصال در این مجموعه داده دارای یک زمان‌مهر است که مشخص‌کننده وقایع ترافیک شبکه است، هر رکورد در این مجموعه دارای برچسب است.

مجموعه داده darpa98 شامل ۵ هفته ترافیک کاری شبکه می‌باشد که:

- هفته ۱ و ۳ حمله‌ای ندارد و جهت آموزش به‌کار می‌رود.
- هفته ۲، ۴۳ نمونه دارد که ۱۸ تا به‌عنوان حمله برچسب‌گذاری شده و جهت توسعه سیستم به‌کار می‌رود.
- هفته ۴ و ۵، ۲۰۱ نمونه دارند که شامل ۵۸ حمله می‌شود (۴۰ تا جدید) و جهت ارزیابی سیستم به‌کار می‌رود.

پیاده‌سازی روش پیشنهادی و مقایسه با سایر روش‌ها در محیط متلب و روی سیستم‌عامل ویندوز ۷، با مقدار رم ۸ گیگابایت و پردازنده core i7 انجام شده است.

۵-۱- معیارهای ارزیابی

در این بخش معیارهای ارزیابی برای دسته‌بندی مورد بررسی قرار خواهند گرفت. برای ارزیابی باید برچسبی که دسته‌بند به حمله نسبت داده با برچسب مرجع مقایسه شود. وقوع حالات مختلف برای دسته‌ها با توجه به مجموعه داده‌های ورودی برای دسته‌بندی با مقادیر TP, FP, TN, FN برای دو دسته مثبت و منفی در جدول ۲ نشان داده شده است.

جدول ۲: ماتریس درهم‌ریختگی برای یک مسئله دسته‌بندی دو دسته‌ای

رکوردهای تخمینی			
رکوردهای واقعی	نوع دسته	منفی	مثبت
	مثبت	TN	FP
	منفی	FN	TP

با توجه به پارامترهای مطرح‌شده معیارهای ارزیابی مختلفی ارائه شده است که ازجمله مهم‌ترین آن‌ها می‌توان به معیارهای خاصیت ، حساسیت و صحت اشاره کرد. معیار صحت مهم‌ترین معیار برای تعیین کارایی یک الگوریتم دسته‌بندی است. این معیار صحت کل یک دسته‌بند را محاسبه می‌کند. این معیار نشان‌دهنده این موضوع است که چند درصد از کل مجموعه داده به‌درستی دسته‌بندی شده است .

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} * 100 \quad (۴)$$

معیار حساسیت تعیین می‌کند که نمونه‌های حمله، تا چه میزان در کلاس مثبت قرار داده شده‌اند، یا به‌عبارت دیگر چه تعدادی از حملات به‌درستی تشخیص داده شده‌اند.

$$Sensitivity = \frac{TP}{TP+FN} * 100 \quad (۵)$$

معیار خاصیت تعیین می‌کند که چه تعدادی از نمونه‌های نرمال به‌درستی تشخیص داده شده‌اند.

$$Specificity = \frac{TN}{FP+TN} * 100 \quad (۶)$$

۵-۲- مقایسه روش پیشنهادی با سایر روش‌ها

میزان معیارهای ارزیابی در روش ترکیبی پیشنهادی برای مسئله تشخیص دوکلاسه حملات و رفتار عادی در جدول ۳ نشان داده شده است، که حاکی از عملکرد مطلوب روش پیشنهادی می‌باشد.

جدول ۳: نتایج ارزیابی کارایی روش پیشنهادی

روش	*Accuracy*	*Specificity*	*Sensitivity*
HMM-ELM	۹۸٪/۰۰	۹۴٪/۵۶	٪۹۸/۸۴

در ادامه برای بررسی کارایی روش ارائه‌شده، مقادیر هر سه معیار با چند روش دیگر مقایسه شده است. اولین مقایسه با مدل مخفی مارکوف به‌تنهایی انجام شده است، تا مشخص شود عملکرد روش پیشنهادی چقدر توانسته است نسبت به HMM [۱۲] بهبود حاصل نماید. به‌علاوه، صحت مدل پیشنهادی با مدلی مبتنی برHMM-C5.0 که در [۱۹] ارائه شده مقایسه شده است. همچنین در کلیه آزمایش‌ها دو روش مبتنی بر KNN [۱۸] و ELM [۱۴] نیز مورد مقایسه قرار گرفته است. در عین حال روش پیشنهادی از دیدگاه صحت و نرخ هشدار نادرست با نتایج گزارش‌شده در [۲۶]، که از روش ترکیبی SVM و ELM استفاده نموده است، نیز مقایسه شده است.

با توجه به مقایسات انجام‌گرفته با روش‌های فوق جدول ۴ به‌دست آمده است. این نتایج حاکی از آن است که عملکرد روش پیشنهادی به‌مراتب بهتر از سایر روش‌ها بوده است.

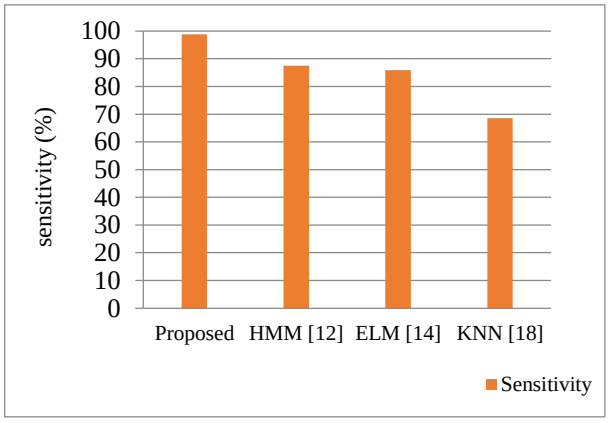
جدول ۴: نتایج مقایسه روش پیشنهادی با سایر روش‌ها بر اساس تعداد رکورد تشخیص داده‌شده

روش	FN	FP	TP	TN
Proposed HMM-ELM	۱۰۶	۱۲۰	۸۹۹۸	۲۰۸۵
HMM [12]	۵۰۰	۱۲۱۴	۸۵۲۸	۵۲۷
KNN [18]	۲۰۶۲	۲۵۲۸	۴۵۰۷	۲۱۷۹
ELM [14]	۱۳۱۴	۱۰۲۸	۸۰۰۰	۴۲۷

در شکل ۵ صحت روش پیشنهادی با روش‌های دیگر مقایسه و ارزیابی گردیده است. در سیستم‌های تشخیص نفوذ معیار صحت به‌عنوان مهم‌ترین عامل در نظر گرفته می‌شود.

شکل ۵: نمودار مقایسه صحت روش پیشنهادی با سایر روش‌ها برروی مجموعه داده Darpa98

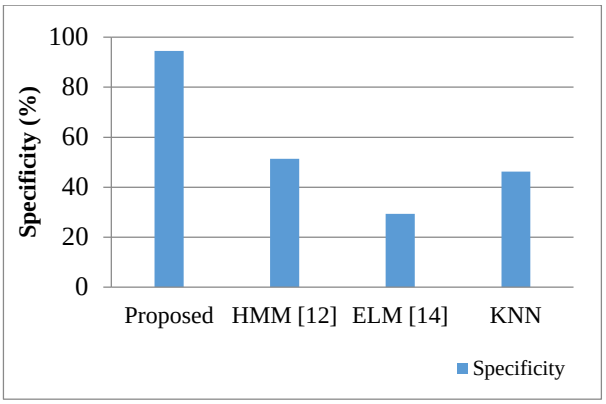
شکل ۶ نشان می‌دهد که در هر روش چند درصد از حملات به‌درستی طبقه‌بندی شده‌اند. نمودار بیانگر این است که روش پیشنهادی توانسته است نرخ تشخیص بهتری را نسبت به سایر روش‌ها فراهم آورد.



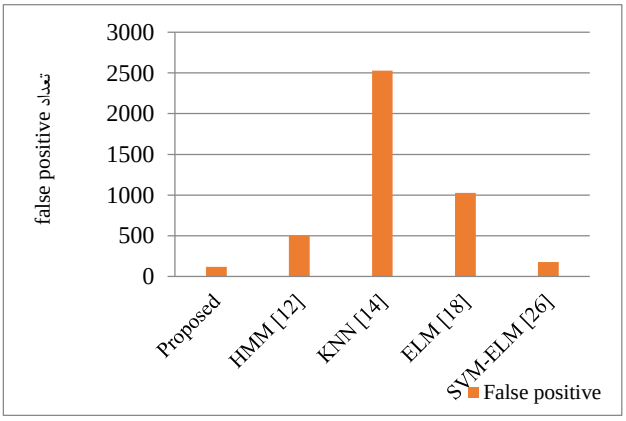
شکل ۶: نمودار مقایسه حساسیت روش پیشنهادی با سایر روش‌ها روی مجموعه داده Darpa98

شکل ۷ تعیین می‌کند که در هر روش چند درصد از فعالیت‌های نرمال به‌درستی طبقه‌بندی شده‌اند. نمودار بیانگر این است که روش پیشنهادی توانسته است نرخ تشخیص بهتری را نسبت به سایر روش‌ها فراهم نماید.

در سیستم‌های تشخیص نفوذ مسئله میزان نرخ مثبت کاذب پایین اهمیت به‌سزایی دارد، زیرا میزان نرخ مثبت کاذب بالا در میزان تشخیص سیستم تأثیر منفی مستقیم خواهد داشت. مثبت کاذب در واقع تعداد رکوردهایی است که به‌اشتباه به‌عنوان حمله شناسایی شده‌اند، درحالی‌که در واقع آن‌ها فعالیت نرمال بوده‌اند. نتایج به‌دست‌آمده از پیاده‌سازی مدل پیشنهادی و سایر روش‌ها در جهت کاهش مثبت کاذب به‌صورت نمودار شکل ۸ نشان داده شده است. شکل ۸ حاکی از آن است که روش پیشنهادی توانسته مثبت کاذب کمتری را نسبت به سایر روش‌ها فراهم کند.



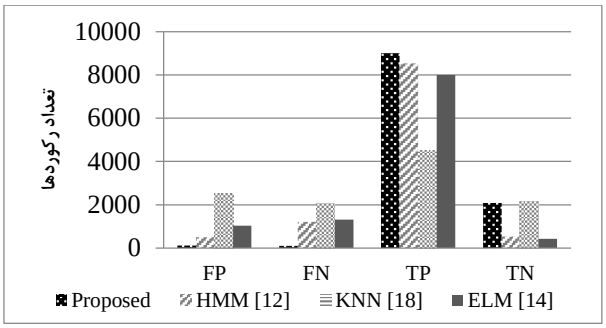
شکل ۷: نمودار مقایسه خاصیت روش پیشنهادی با سایر روش‌ها روی مجموعه داده Darpa98



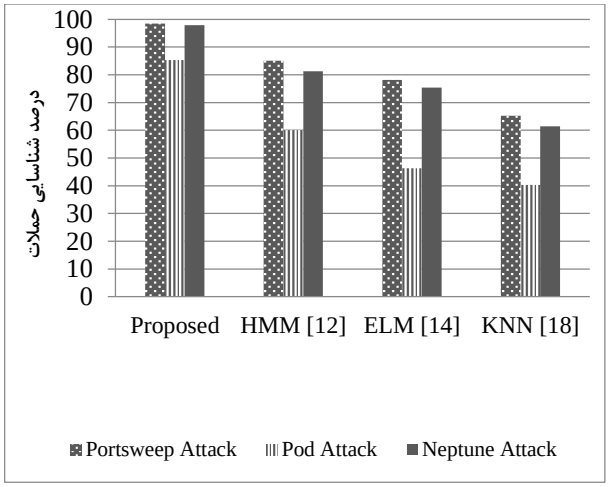
شکل ۸: نمودار مقایسه مثبت کاذب در روش‌های مورد ارزیابی

نمودار شکل ۹ نمایانگر تعداد رکوردهای تخمینی در روش‌های مورد مقایسه بر روی مجموعه داده موردنظر می‌باشد. به‌طور مثال تعداد رکورد حملاتی که به‌درستی طبقه‌بندی شده‌اند (TP) در روش پیشنهادی برابر ۸۹۹۸ است، یعنی روش پیشنهادی توانسته از این دیدگاه بهبود حاصل نماید. همچنین تعداد رکوردهایی که در روش پیشنهادی به‌اشتباه به‌عنوان حمله قرار گرفتند، در حالی که در واقع فعالیت نرمال بوده‌اند ۱۲۰ است، که این مقدار از بقیه روش‌ها پایین‌تر است، درحالی‌که روشی مثل HMM، مثبت کاذب ۵۰۰ دارد.

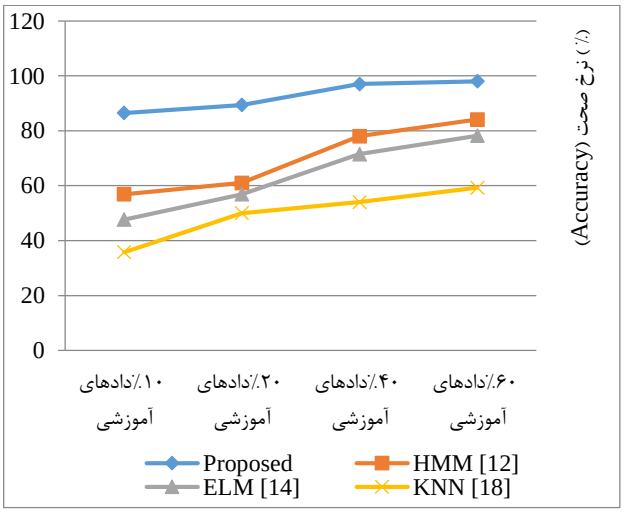
شکل ۱۰ عملکرد هر یک از روش‌ها در تشخیص کلاس‌های حمله از نوع Neptune, pod, portsweep را نشان می‌دهد. همان‌طور که این نمودار نشان می‌دهد روش پیشنهادی توانسته است ۹۸/۴۸٪ حملات portsweep را تشخیص دهد، درحالی‌که به‌طور مثال در روش ELM به‌تنهایی ۷۸/۱۴٪ از حملات portsweep تشخیص داده شده است. به‌همین ترتیب در مورد حملات pod روش پیشنهادی با تشخیص ۸۵/۳۶٪ نسبت به بقیه روش‌ها بهبود داشته است. در مورد حملات Neptune نیز روش پیشنهادی توانسته ۹۷/۹۵٪ حملات را تشخیص دهد.



شکل ۹: نمودار نتایج رکوردهای تخمینی در روش‌های مورد مقایسه



شکل ۱۰: نمودار تشخیص انواع حملات در روش‌های مورد مقایسه



شکل ۱۱: تأثیر کاهش نمونه‌های آموزشی روی مدل پیشنهادی HMM-ELM در مقایسه با سایر روش‌ها

با توجه به این‌که یکی از دلایلی که این رویکرد جدید مبتنی بر مدل مخفی مارکوف و ماشین یادگیری مفرط مطرح شده است، مقاومت در برابر داده‌های آموزشی ناکافی بود، تأثیر افزایش و کاهش این داده‌ها در روش پیشنهادی و سایر روش‌ها، مورد ارزیابی و مقایسه قرار گرفت. نتایج به‌دست‌آمده از این ارزیابی‌ها در نمودار شکل ۱۱ مشهود است. همان‌طور که مشاهده می‌شود، روش پیشنهادی توانسته است در برابر داده‌های کم نیز نسبت به سایر روش‌ها نتایج و نرخ صحت قابل قبولی را ارائه نماید.

به‌عنوان آخرین آزمایش روش پیشنهادی از نظر زمان اجرا با سایر روش‌ها مقایسه شده است. نتایج این آزمایش در جدول ۵ مشاهده می‌شود. مقادیر ذکرشده در این جدول همگی برحسب میلی‌ثانیه است و به‌طور متوسط برای هر جلسه محاسبه شده است.

جدول ۵: زمان اجرای روش‌های مورد مقایسه بر روی دادگان Darpa98

روش	ELM [14]	KNN [18]	HMM [12]	Proposed HMM-ELM
زمان تشخیص (میلی‌ثانیه)	۰/۰۳۵	۷/۵۴	۲/۹۵	۳/۳۴

جدول فوق نشان می‌دهد که الگوریتم پیشنهادی نسبت به KNN سرعت بهتری دارد، اما به‌دلیل اینکه یک رویکرد ترکیبی است، نسبت به هریک از دو روش HMM و ELM کندتر است. با این حال زمان لازم برای تشخیص بسیار پایین و در مقیاس میلی‌ثانیه ناچیز است، و به این معناست که روش پیشنهادی علی‌رغم ترکیبی‌بودن، برای کاربردهای بلادرنگ مناسب است.

یک سؤال احتمالی، مقایسه این روش با روش‌هایی است که برای حل مسئله در شرایط عدم وجود اطلاعات دنباله ترافیک شبکه طراحی شده‌اند. از این موارد برای مثال می‌توان به دادگان KDD Cup اشاره نمود، که در این دادگان هر جلسه نه با یک دنباله از مشاهدات که با یک رکورد توصیف شده است. این دادگان یک نسخه از دادگان Darpa98 است، با این تفاوت که پردازش شده است و استخراج ویژگی از آن صورت گرفته است [۳۱]. از این‌رو اگرچه مقایسه مستقیم روش‌ها امکان‌پذیر نیست، اما می‌توان جهت یک بررسی مقایسه‌ای به برخی از آخرین دستاورد در این زمینه اشاره کرد. برای مثال مقاله [۲۶] که روش ترکیبی SVM-ELM بوده است، در مجموع به صحت ۹۵/۷۵ دست یافته است. یا مقاله [۲۹] که در بهترین حالت بر روی دادگان KDD Cup به صحت ۸۴/۱۲ رسیده است. این نتایج حاکی از صحت بیشتر روش پیشنهادی در مقایسه با کارهای اخیر در زمینه تشخیص نفوذ است. البته همان‌طور که ذکر شد، روش پیشنهادی نسبت به دو روش مذکور از اطلاعات بیشتری که همان توالی است استفاده می‌کند، و بخش عمده برتری آن نیز به‌همین جهت است. هرچند آزمایشات، مطابق آنچه در نمودارهای قبلی گزارش شده است، نشان می‌دهد که روش پیشنهادی نسبت به سایر روش‌های مبتنی بر توالی نیز موفق‌تر است.

۶- نتیجه‌گیری

در این پژوهش یک سیستم تشخیص نفوذ ارائه شده است که از رویکرد ترکیبی مدل مخفی مارکوف و ماشین یادگیری مفرط استفاده می‌کند. بدین‌صورت که بعد از پیش‌پردازش داده‌ها و گسسته‌سازی

مقادیر آن‌ها، داده‌ها به‌عنوان مشاهدات وارد مدل مخفی مارکوف می‌شوند و در این مدل وضعیت دسترسی‌ها همان حالات مدل مخفی هستند. مدل مخفی مارکوف با استفاده از الگوریتم یادگیری آموزش دیده و سپس با اعمال الگوریتم ویتربی توالی بهترین و محتمل‌ترین حالات استخراج می‌شود. در مرحله بعد از این دنباله حالات جهت ورودی و آموزش دسته‌بند ELM استفاده شده است.

در این روش از قابلیت توالی و ترتیبی مدل مخفی مارکوف بهره گرفته شده است و در مرحله اول، ازآن‌جهت استخراج دنباله مشاهدات در مجموعه داده برحسب رکوردهای متوالی ترافیک شبکه استفاده شده است. سپس در مرحله دوم، برای حل مشکلاتی همچون ناکافی‌بودن داده‌های آموزشی از روش ماشین یادگیری مفرط جهت دسته‌بندی نتایج استفاده شده است، زیرا ناکافی‌بودن داده‌های آموزش باعث خواهد شد که مدل مخفی مارکوف نتواند به‌درستی آموزش ببیند. بنابراین HMM با ماشین یادگیری مفرط ادغام شده است تا در برابر داده‌های کم مقاومت کرده و نتایج خوبی حاصل شود.

آزمایشات نشان می‌دهد که کارایی روش پیشنهادی در مقایسه با روش‌های پیشین به‌صورت قابل‌توجهی افزایش یافته است. نتایج به‌این‌صورت است که نرخ صحت روش پیشنهادی ۹۸/۰۰٪، نرخ حساسیت ۹۸/۸۴٪ و نرخ خاصیت روش پیشنهادی ۹۴/۵۶٪ به‌دست آمده است. همچنین روش پیشنهادی توانسته است نرخ مثبت کاذب بهتر و پایین‌تری نیز در مقایسه با سایر روش‌های مورد بررسی فراهم کند.

یکی از کارهایی که می‌توان در آینده به آن پرداخت بهینه‌سازی ساختار مدل مخفی مارکوف است. یعنی با کم‌کردن تعداد مشاهدات و یا کاهش ویژگی‌هایی که لزومی به استفاده از آن‌ها نیست و اهمیت چندانی ندارند، از پیچیدگی مدل مخفی مارکوف کاسته شود. همچنین لازم به ذکر است که ساختار پیشنهادی، به‌جز در برخی تنظیمات، وابسته به مسئله نیست و می‌توان از این چارچوب در سایر کاربردهایی که با مدل‌سازی توالی سروکار دارند و در عین حال با مشکل کمبود داده‌های آموزشی مواجه است استفاده نمود.

۷- مراجع

[1] J. Cannady, “Artificial neural networks for misuse detection, ” *National information systems security conference*, pp. 368-81, 1998.

[2] S. E. Smaha, “Haystack: An intrusion detection system,” Aerospace Computer Security Applications Conference, pp. 37-44, 1988.

[3] M. Panda and M. R. Patra, “Mining association rules for constructing a network intrusion detection model,” International journal of applied engineering research, vol. 4, pp. 381-98, 2009.

[4] J. P. Anderson, *Computer Security Threat Monitoring and Surveillance*, Technical report, James P. Anderson Company, Fort Washington, Pennsylvania, 1980.

[5] E.D. Denning, “An intrusion detection model,” Seventh IEEE Symposium on Security and Privacy, pp. 119–131, 1986.

[۶] رحیم بجانی، محمد کلانتری و امیرمسعود افتخاری مقدم، «ارائه چهارچوبی مبتنی بر نظریه بازی‌ها برای جلب مشارکت گره‌ها در فرآیند شناسایی گره‌های مخرب در شبکه‌های حسگر بی‌سیم»، مجله مهندسی برق دانشگاه تبریز، مقاله آماده انتشار، انتشار آنلاین از تاریخ ۳ شهریور ۱۳۹۶.

[7] R. S. Naoum, N. A. Abid and Z. N. Al-Sultani, "An enhanced resilient backpropagation artificial neural network for intrusion detection system," International Journal of Computer Science and Network Security (IJCSNS), vol. 12, pp. 11-16, 2012.

[8] L. Koc, T. A. Mazzuchi, and S. Sarkani, "A network intrusion detection system based on a hidden naïve bayes multiclass classifier," Expert Systems with Applications, vol. 39, pp.13492-13500, 2012.

[۹] مسعود فرکی و مازیار پالهنگ. «بازشناسی برخط حروف فارسی بر پایه مدل مخفی مارکوف»، مجله مهندسی برق دانشگاه تبریز، ۴۰(۱)، ۲۳-۳۴، ۱۳۸۹.

[10] R. Khanna and H. Liu, "System approach to intrusion detection using hidden markov model,” International conference on Wireless communications and mobile computing, pp. 349-354,2006.

[11] R. Jain and N. S. Abouzakhar, “Hidden markov model based anomaly intrusion detection,” International Conference of Internet Technology And Secured Transactions, pp. 528-533, 2012.

[12] J. C. Badajena and C. Rout, “Incorporating hidden markov model into anomaly detection technique for network intrusion detection,” International Journal of Computer Applications, vol. 53, No. 11, pp. 42-47, 2012.

[13] S. Selim, M. Hashem and T. M. Nazmy, “Intrusion detection using multi-stage neural network,” International Journal of Computer Science and Information Security, vol. 8, No. 4, pp. 14-20, 2010.

[14] C. Cheng, W. P. Tay and G.-B. Huang, “Extreme learning machines for intrusion detection,” International Joint Conference on Neural Networks (IJCNN), pp. 1-8, 2012.

[15] G. Creech and F. Jiang, “The application of extreme learning machines to the network intrusion detection problem,” International Conference of Numerical Analysis and Applied Mathematics, pp. 1506-1511, 2012.

[16] S. Dhopte and M. Chaudhari, “Genetic algorithm for intrusion detection system,” IJRIT International Journal of Research in Information Technology, vol. 2, pp. 503-509, 2014.

[17] Y. B. Bhavsar and K. C. Waghmare, "Intrusion detection system using data mining technique: support vector machine," International Journal of Emerging Technology and Advanced Engineering, vol. 3, pp. 581-586, 2013.

[18] M. Govindarajan and R. Chandrasekaran, “Intrusion detection using k-Nearest Neighbor,” International Conference of Advanced Computing, pp. 13-20, 2009.

[19] M. Khosronejad, E. Sharififar, H. A. Torshizi and M. Jalali, "Developing a hybrid method of hidden markov models and c5.0 as a intrusion detection system," International Journal of Database Theory and Application, vol. 6, pp. 165-174, 2013.

[20] D. Ariu, R. Tronci and G. Giacinto, "HMMPayl: An intrusion detection system based on hidden markov models," computers & security, vol. 30, pp. 221-241, 2011.

[21] N. Devarakonda, S. Pamidi, V. V. Kumari and A. Govardhan, "Intrusion detection system using bayesian

network and hidden markov model," Procedia Technology, vol. 4, pp. 506-514, 2012.

[22] J. M. Fossaceca, T. A. Mazzuchi and S. Sarkani, "MARK-ELM: application of a novel multiple kernel learning framework for improving the robustness of network intrusion detection," Expert Systems with Applications, vol. 42, pp. 4062-4080, 2015.

[23] F. Kuang, W. Xu and S. Zhang, "A novel hybrid kpca and svm with ga model for intrusion detection," Applied oft Computing, vol. 18, pp. 178-184, 2014.

[24] A. Chandrasekhar and K. Raghuveer, "Intrusion detection technique by using k-means, fuzzy neural network and SVM classifiers," International Conference Computer Communication and Informatics, pp. 1-7, 2013.

[25] "KDD Cup Dataset." http://kdd.ics.uci.edu/databases/kddcup99/

[26] W. L. Al-Yaseen, Z. A. Othman and M.A.A. Nazri, "Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system," Expert Syst. Appl, vol. 67, pp. 296-303, 2017.

[27] E. De la Hoz, E. De la Hoz, A. Ortiz, J. Ortega and B. Prieto, "PCA filtering and probabilistic SOM for network intrusion detection" Neurocomputing, vol. 164, pp. 71–81, 2015.

[28] W. Feng, Q. Zhang, G. Hu and J.X. Huang, "Mining network data for intrusion detection through combining SVMs with ant colony networks," Future Generation Computer Systems, vol. 37, pp. 127–140, 2014.

[29] R.A.R. Ashfaq, X.Z. Wang, J.Z. Huang, H. Abbas and Y.L. He, "Fuzziness based semi-supervised learning approach for intrusion detection system," Information Sciences, vol. 378, pp. 484–497, 2017.

[30] P. Saini and S. Godara, "Modelling intrusion detection system using hidden markov model: a review," International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 4, pp. 542-547, 2014.

[31] "Darpa1998 Dataset." Available: http://www.ll.mit.edu/ ideval/ data/

زیرنویس‌ها

۱ Intrusion Detection Systems(IDS)

۲ Normal

۳ Anomaly

۴ Hidden Markov Model

۵ Extreme Learning Machine

۶ Feedforward

۷ Introusion Detection Expert System

۸ Backpropagation

۹ Hidden Naïve Bayes

۱۰ Gaussian Mixture Model

۱۱ Extreme Learning Machine

۱۲ Tuning

۱۳ Inferior

۱۴ Support Vector Machine

۱۵ Multiple Adaptive Reduced Kernel ELM

۱۶ Kernel Principal Component Analysis

۱۷ Genetic Algorithm

۱۸ Probabilistic Self-Organizing Map (PSOM)

۱۹ Self-Organizing Ant Colony Network (SOACN)

۲۰ Ergodic Model

۲۱ Single Layer Feedforward Network

۲۲ backpropagation

۲۳ Kernel based

۲۴ Basis Elm

۲۵ positive semi-definite

۲۶ bin

۲۷ time stamp

۲۸ session

۲۹ specificity

۳۰ sensitivity

۳۱ Accuracy