یک روش یادگیری عمیق مبتنی بر شبکه‌های عصبی برای تشخیص نفوذ در شبکه‌های کامپیوتری

نویسندگان: غلامرضا خان عزیزی- reza.azizi499@gmail.com – کارشناسی ارشد کامپیوتر
حسین غفاریان- hossein.ghaffarian@gmail.com –دکترای کامپیوتر دانشگاه علم و صنعت
مهدی صادق زاده -sadegh_1999@yahoo.com- هیئت علمی دانشگاه قوچان

چکیده: یکی از راهکارهای مورد استفاده جهت ایجاد امنیت در شبکه‌های کامپیوتری استفاده از سیستم‌های تشخیص نفوذ می‌باشد. معماری‌های امروزی استفاده شده برای تشخیص نفوذ، طراحان این سیستم‌ها را در انتخاب نوع معماری کارایی که بتواند قابلیت اطمینان بیشتری در مورد تشخیص تهاجمات داشته باشد با دشواری‌هایی مواجه کرده است. از جمله راهکارهایی که در راستای امن‌سازی سیستم‌ها و شبکه‌های کامپیوتری به وجود آمده است، پیدایش سیستم‌های تشخیص نفوذ است. در این تحقیق روشی برای ترکیب یادگیری عمیق و یادگیری با ناظر در تشخیص نمونه‌های نفوذ با هدف افزایش امنیت شبکه‌های کامپیوتری ارائه شده است. یادگیری با ناظر به منظور آموزش پارامترهای یک الگوریتم شبکه عصبی عمیق ارائه شده که از ترکیب خطی و بازنمایی ویژگی‌های مؤثر استفاده می‌کند. این روش بر مبنای یک الگوریتم یادگیری با ناظر و یک شبکه عصبی عمیق می‌باشد که تعداد مناسب لایه‌های پنهان و تعداد نرون‌ها در هر لایه را با توجه به یک مقدار آستانه بهینه‌سازی می‌کند. نتایج آزمایش‌ها روی مجموعه داده *NSL-KDD* برتری روش پیشنهادی را با دقت ۹۷٫۶۴٪ نسبت به روش‌های *MARS* و *DLNN* نشان می‌دهد.

کلیدواژه: تشخیص نفوذ، یادگیری عمیق، شبکه عصبی، امنیت شبکه.

۱ - مقدمه

سیستم تشخیص نفوذ، در داخل خود، دارای مجموعه داده‌هایی است که حملات مختلف را دارا می‌باشند و اگر برنامه کامپیوتری یا فردی قصد نفوذ و یا ورود بدون مجوز به پایگاه داده را داشته باشد، ابتدا با داده‌های حملات از پیش تعریف شده در خود مطابقت می‌دهد. اگر به نتیجه‌ای برسد که برنامه کامپیوتری یا فردی که قصد دسترسی به پایگاه داده سیار را دارند، به صورت بدون مجوز از آن استفاده خواهند کرد، آن را به صورت یک نفوذ تلقی می‌کند و سعی بر جلوگیری از حملات خواهد داشت. اما اگر این تطابق به نتیجه‌ای نرسد، یعنی با داده‌های حملات در سیستم تشخیص نفوذ همخوانی نداشته باشد، عملیات ورود انجام می‌گیرد[۱، ۲].

با توجه به این که حملات مختلفی در سطح یک پایگاه داده سیار رخ می‌دهد، لذا ارائه یک سیستم تشخیص نفوذ در آن امری ضروری است. شناسایی حملاتی همچون محرومیت از خدمات و همین‌طور حمله محرومیت از خدمات توزیع شده به عنوان حملات اصلی جهت تشخیص، مدنظر واقع می‌گردد. در این پژوهش به ارائه یک سیستم تشخیص نفوذ چند هدفه در پایگاه داده‌های سیار بر پایه یادگیری عمیق پرداخته می‌شود. یادگیری عمیق شاخه‌ای از بحث یادگیری ماشینی و مجموعه‌ای از الگوریتم‌هایی است که تلاش می‌کنند مفاهیم انتزاعی سطح بالا را با استفاده از یادگیری در سطوح و لایه‌های مختلف مدل کنند [۳].

بیشتر مدل‌های یادگیری عمیق مدرن، بر شبکه‌های عصبی مصنوعی مبتنی هستند، گرچه ممکن است شامل فرمول‌های گزاره‌ای یا متغیرهای پنهان سازمان یافته لایه‌ای در مدل‌های مولد همچون گره‌ها در شبکه‌های باور عمیق و ماشین‌های بولتزمن عمیق نیز باشند. واژه «عمیق» در «یادگیری عمیق» به تعداد لایه‌هایی که داده‌ها از آنها منتقل می‌شوند اشاره دارد[۴]. بطور دقیق‌تر، سیستم‌های

یادگیری عمیق، دارای یک عمق اساسی برای مسیر تخصیص اعتبار می‌باشند. معماری‌های یادگیری عمیق را اغلب با یک روش لایه‌به‌لایه حریصانه می‌سازند. یادگیری عمیق به بازگشایی این انتزاعات و انتخاب ویژگی‌هایی که عملکرد را بهبود می‌دهند کمک می‌کند. ویژگی‌ها، عملکرد را بهبود می‌دهند. برای کارهای یادگیری نظارت شده، روش‌های یادگیری عمیق با انتقال داده‌ها به نمایش‌های میانی فشرده، مشابه مؤلفه‌های اصلی، نیاز به مهندسی ویژگی را برطرف می‌کنند و ساختارهایی لایه‌ای بدست می‌دهند که افزونگی را از نمایش حذف می‌کنند.

ادامه این تحقیق به این شرح خواهد بود: در بخش دوم به بررسی برخی از جدیدترین کارهای انجام شده در تشخیص نفوذ می‌پردازیم. روش پیشنهادی در بخش سوم مطرح شده و نتایج حاصل از آن در بخش چهارم آورده شده است. در نهایت در بخش پنجم نتیجه‌گیری و پیشنهادات بیان می‌شود.

۲- پیشینه تحقیق

یانگ و همکاران، یک سیستم تشخیص نفوذ چند بعدی برای شبکه SCADA پیشنهاد دادند[۵]. این تحقیق سیستم تشخیص نفوذی چندلایه‌ای را پیشنهاد و گسترش داده است که بر محیط فیزیکی خاص و داده‌های برنامه ایستگاه فرعی محافظت شده متمرکز است. وانگ و همکاران، حداکثر تاثیر نفوذ در شبکه‌های اجتماعی تحت شرایط یک مدل آبشاری مستقل را ارائه دادند[۶]. مدل پیشنهاد شده برای توسعه اطلاعات، به منظور توضیح دادن اینکه چگونه کاربر نظرات خودش را می‌سازد، معرفی شده است. کریمی و همکاران، موضوعات نفوذ فتوولتائیک و اثرات آن در شبکه توزیع را مورد بررسی قرار دادند[۷]. در این تحقیق اثرات PV بعنوان یک RES روی شبکه توزیع و موضوع نفوذ همراه با آن در سیستم‌های الکتریکی بطور کامل ارائه شده است. آل-جراه و ارفات، سیستم تشخیص نفوذ شبکه با استفاده از طبقه‌بندی رفتار حمله را پیشنهاد دادند[۸]. این تحقیق از شبکه‌ی عصبی TDDNN برای تشخیص رفتار زمانی حملات شبکه استفاده می‌کند.

برگوهاین، روش FuGeIDS را که به مفهوم الگوهای ژنتیکی مبهم در سیستم‌های کشف نفوذ است، پیشنهاد دادند[۹]. رویکرد استفاده از الگوریتم ژنتیک و منطق فازی برای کشف ناهنجاری‌ها استفاده می‌کند. کوک و همکاران، یک سیستم کشف نفوذ مبتنی بر دسته‌بندی ترکیبی را مبتنی بر بیزین مخفی توسعه دادند[۱۰]. هارین و ورموتا، سیستم تشخیص نفوذ مبتنی بر قانون وابستگی فازی همراه با برنامه‌نویسی شبکه ژنتیک را ارائه دادند[۱۱]. این روش با بهره‌برداری از ویژگی‌های فرعی و دسته‌بندی مبتنی بر قوانین استخراج شده پیشنهاد شده است. وانگ و همکاران، یک رویکرد نوین برای شناسایی نفوذ با استفاده از شبکه-های عصبی مصنوعی و خوشه بندی فازی توسعه داده اند[۱۲]. در تحقیقی دیگر گواش و همکاران، استفاده از برنامه پروفایل‌های رفتاری برای تشخیص نفوذ را مطرح ساختند[۱۳].

مارکام و دوبی، یک بررسی کلی در مورد کاوش قوانین فازی در سیستم تشخیص نفوذ انجام دادند[۱۴]. این تحقیق سرعت و دقت شناسایی را بهبود بخشیده و روش کاوش قوانین رابطه‌ای کارآمدتری را در چارچوب الگوریتم مقایسه‌ای برای شناسایی غیرعادی تجارب مبتنی بر شبکه پیشنهاد نموده و در نهایت به مقایسه و تحلیل نتایج تجربه می‌پردازد. زارپلو و همکاران، بررسی تشخیص نفوذ در اینترنت اشیاء را انجام داده‌اند[۱۵]. دریفوت و همکاران، یک مدل تشخیص نفوذ هوشمند برای رایانش ابری ارائه دادند[۱۶]. این روش یک مدل تشخیص نفوذ هوشمند براساس اصل همکاری است و بر پایه نفوذ مبتنی بر میزبان کار می‌کند. باروت و توشنیوال، تشخیص نفوذ در شبکه‌های هیبریدی، مبتنی بر تکنیک داده کاوی را ارائه داده‌اند[۱۷]. قلعه گلابی و رضایی پناه (۲۰۱۷)، سیستم‌های تشخیص نفوذ با کاهش ویژگی مبتنی بر الگوریتم ژنتیک و تکنیک-های داده‌کاوی را بهبود داده‌اند[۱۸]. به دلیل وجود مشخصه‌های زیاد در داده‌های مربوط به سیستم‌های تشخیص نفوذ در این تحقیق مشخصه‌های مطلوب و مؤثر با استفاده از الگوریتم ژنتیک بهبود یافته انتخاب شده و سپس با استفاده از تکنیک‌های داده کاوی استاندارد، مدلی برای طبقه‌بندی داده‌ها ارائه شده است. وارسی، انتخاب تکراری بر اساس الگوریتم بهینه‌سازی ازدحام ذرات را (IPPSO) برای سیستم‌های تشخیص نفوذ ارائه داد[۱۹]. این روش یک روش ترکیبی است که بر اساس استخراج قوانین رابطه‌ای و انتخاب تکراری آنها بر اساس الگوریتم ازدحام ذرات عمل می‌کند.

۳- روش پیشنهادی

در این تحقیق برای تشخیص نفوذ از یک مدل مبتنی بر یادگیری عمیق در شبکه عصبی MLP استفاده می‌شود. به طور کلی شبکه عصبی و یادگیری عمیق زیر مجموعه‌ای از یادگیری ماشین است و خود یادگیری ماشین زیرمجموعه‌ای از هوش مصنوعی می‌باشد. یادگیری شبکه عصبی عمیق براساس یادگیری چندسطحی سعی بر استخراج اطلاعات می‌کند در حالی‌که شبکه عصبی کلاسیک حداکثر از دو لایه تشکیل می‌شود. در شبکه عصبی عمیق، ستونی از لایه‌های نرونی وجود دارد و پایین‌ترین سطح لایه، وظیفه جمع‌آوری داده‌های اولیه را برعهده دارد. هر نرون سعی در ذخیره‌سازی اطلاعات و انتقال آن به لایه‌های بعدی دارد. این فرایند به نحوی انجام می‌شود که اطلاعات درون نرون‌ها همراه با لایه پنهانی از اطلاعات نرونی استخراج می‌شود. بنابراین از این طریق داده‌ها از پایین‌ترین سطح لایه به بالاترین سطح لایه رفته و به جمع‌آوری اطلاعات می‌پردازند.

روش ارائه شده در این تحقیق بر مبنای یک الگوریتم یادگیری با ناظر حریصانه و یک شبکه عصبی عمیق مطابق شکل ۱ می‌باشد. ورودی روش پیشنهادی، داده‌های پیش‌پردازش شده سیستم تشخیص نفوذ می‌باشد. تعداد مناسب لایه‌های پنهان و تعداد نرون‌ها در هر لایه در یک فرایند تکراری با توجه به یک مقدار آستانه انجام می‌شود. همچنین به منظور تنظیم دقیق‌تر ویژگی‌های انتخابی در هر تکرار بهینه‌سازی، مقادیر تکرار قبل به عنوان ورودی شبکه عصبی عمیق در نظر گرفته می‌شود. در مرحله بعد بر مبنای پارامترهای ورودی تعیین شده الگوریتم شبکه عصبی عمیق بر روی داده‌های ورودی تشخیص نفوذ آموزش می‌بینند. در اینجا ورودی اولیه (در تکرار اول) ویژگی‌های انتخابی، پارامترهای لایه پنهان و وزن‌ها به صورت تصادفی تنظیم می‌شود.

شبکه‌های عصبی عمیق، اصطلاحی برای مجموعه‌ای از شبکه‌های عصبی با معماری چند لایه است که نشان می‌دهند چگونه شبکه‌های عصبی با تعداد زیادی از لایه‌ها می‌توانند در ایجاد ساختارهای بازنمایی مورد نیاز در یادگیری عمیق عملکرد مناسبی داشته باشند. الگوریتم‌های یادگیری می‌توانند در تنظیم وزن‌های این شبکه‌ها به کار گرفته شوند. هر چه تعداد لایه‌ها در شبکه عصبی بیشتر شود، مسئله بهینه‌سازی پیچیده‌تر می‌شود و برای حل آن می‌توان از الگوریتم‌های با ناظر و لایه‌ای حریصانه استفاده کرد. در این روش ابتدا هر لایه به صورت مجزا با فرض ثابت بودن سایر لایه‌ها آموزش داده می‌شود و در نهایت روی کل شبکه تنظیم دقیق و یکپارچه‌ای انجام می‌گیرد. به طور کلی در الگوریتم پیشنهادی وزن لایه‌ها به صورت مرحله‌ای بروزرسانی می‌شود و در هر مرحله وزن یک لایه تنظیم می‌شود. همچنین به منظور بهبود کارایی شبکه در تشخیص ویژگی‌های مؤثر ورودی، ورودی بخش انتخاب ویژگی به صورت ۰ یا ۱ مقداردهی می‌شود که ۰ به معنی عدم انتخاب ویژگی و ۱ به معنی انتخاب آن ویژگی در تشخیص نفوذ است.

پس از آموزش شبکه عصبی عمیق و بررسی نتایج پیش‌بینی، به‌روزرسانی ویژگی‌های انتخابی، تعداد لایه‌های پنهان و تعداد نرون‌ها در هر لایه پنهان براساس خروجی به دست‌آمده انجام می‌شود. بر مبنای تغییرات انجام شده یادگیری شبکه عصبی عمیق تکرار شده و این فرایند تا زمان محقق شدن دقت خروجی تکرار می‌گردد.



شکل ۱: فلوچارت روش پیشنهادی

۳-۱ پیش‌پردازش داده‌های تشخیص نفوذ

پیش‌پردازش به منظور آماده‌سازی داده‌ها جهت پردازش و همچنین بهبود کیفیت داده‌های واقعی انجام می‌شود. در این مرحله به ویژگی‌های کیفی، مقادیر کمی انتساب داده می‌شود. سپس نمونه‌هایی با مقادیر از دست رفته از پایگاه داده حذف می‌شوند. پس از آن از روش Z-score برای نرمال‌سازی داده‌ها مطابق رابطه (۱) استفاده می‌شود.

$$x_{i,j}^{Z-score} = \frac{x_{i,j} - \mu_j}{\sigma_j} \qquad (۱)$$

جائیکه $x_{i,j}$ و $x_{i,j}^{Z-score}$ به ترتیب مقدار واقعی و مقدار نرمال شده j-مین ویژگی در نمونه i-ام است. μ_j و σ_j نیز به ترتیب میانگین و انحراف معیار مقادیر تمام نمونه‌ها برای j-مین ویژگی می‌باشد.

۳-۲ مقداردهی اولیه پارامترهای شبکه عصبی عمیق

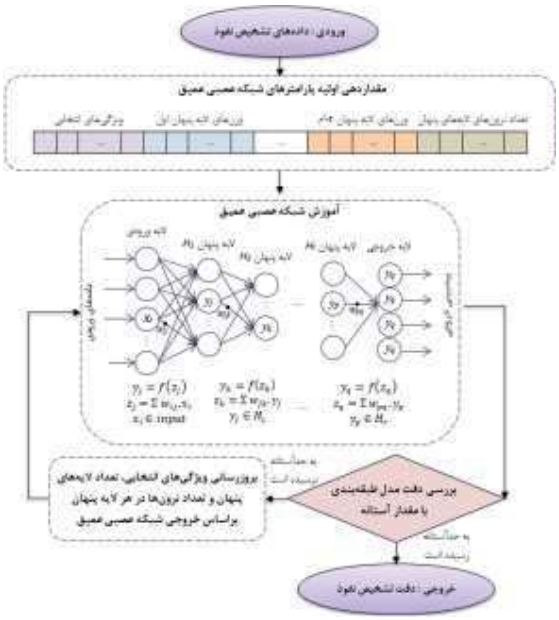
مطابق شکل ۲ پارامترهای شبکه عصبی مقداردهی می‌شوند. این پارامترها شامل وزن‌های هر لایه از شبکه و همچنین مشخصات لایه‌های پنهان می‌باشد. در ابتدا این پارامترها در بازه‌های خاصی به صورت تصادفی مقداردهی شده و پس از آن در هر مرحله براساس خروجی شبکه به روزرسانی می‌شوند. به منظور بهبود کارایی شبکه در تشخیص ویژگی‌های مؤثر ورودی، وزن لایه ورودی نیز در نظر گرفته شده و به صورت ۰ یا ۱ مقداردهی می‌شود که ۰ به معنی عدم انتخاب ویژگی و ۱ به معنی انتخاب آن ویژگی در تشخیص نفوذ است.

تعداد ویژگی‌های ورودی m، تعداد لایه‌های پنهان r و H_j تعداد نرون‌های j-مین لایه پنهان است. تعداد وزن‌های لایه پنهان اول برابر $m \times H_1$ و به طور مشابه تعداد وزن‌های لایه پنهان دوم برابر $H_1 \times H_2$ خواهد بود. در نهایت تعداد وزن‌های r-مین لایه پنهان برابر H_r می‌باشد. در این تحقیق خروجی چهار نرون (تعداد حملات) در نظر گرفته شده است. بنابراین تعداد کل پارامترهای وزنی جهت مقداردهی به صورت رابطه (۲) می‌باشد.

$$\#W = [m \times H_1] + [H_1 \times H_2] + \cdots + [H_{r-1} \times H_r] \qquad (۲)$$

پارامتر تعداد نرون‌ها برای هر لایه پنهان متناسب با تعداد لایه‌های پنهان (r) مشخص می‌شود. همچنین تعداد پارامترهای ویژگی‌های انتخابی برابر m است؛ بنابراین تعداد آیتم‌های نهایی جهت مقداردهی برای شبکه عصبی به صورت رابطه (۳) می‌باشد.

$$\#Items = \mathrm{m} + \#W + r \qquad (۳)$$



شکل ۲: روند جریان به‌روزرسانی وزن‌ها در شبکه عصبی عمیق

۳-۳ آموزش شبکه عصبی عمیق

در این تحقیق برای یادگیری وزن‌ها از روش BP[۱] استفاده شده است. در این روش با استفاده از چندین تکنیک مختلف نظیر GD[۲]، RB[۳] و LM[۴] سعی شده تا مربع خطای بین خروجی‌های شبکه و تابع هدف کمینه شود. در اینجا از تابع فعال‌سازی غیرخطی sigmoid استفاده شده که نگاشت ورودی‌ها به خروجی‌ها را مدیریت می‌کند.

در مرحله یادگیری شبکه عصبی عمیق از معیار «خطای میانگین مربعات (MSE)» برای ارزیابی وزن‌های تخصیص داده شده استفاده می‌شود. در فرایند یادگیری شبکه عصبی عمیق وزن لایه ورودی را به صورت ۰ یا ۱ درنظرگرفته تا ویژگی‌های ورودی به سیستم یادگیری مشخص شوند. از

[۱] BackPropagation
[۲] Gradient Descent
[۳] Resilient Backpropagation
[۴] Levenberg- Marquardt

این تکنیک برای شناسایی ویژگی‌های مؤثر به صورت خودکار استفاده شده است.

۳-۴ به‌روزرسانی پارامترهای شبکه عصبی عمیق

شبکه عصبی عمیق طراحی شده دارای سه پارامتر ویژگی-های انتخابی، وزن‌های شبکه عصبی و مشخصات لایه‌های پنهان می‌باشد. این پارامترها در ابتدای کار به صورت تصادفی در بازه‌های خاصی مقداردهی شد. ویژگی‌های انتخابی مقدار ۰ یا ۱، وزن‌ها بین [+1 ,-1] و تعداد لایه‌های پنهان بین $[1, r_{max}]$ مقداردهی می‌شوند (اندازه نرون‌ها در هر لایه مخفی نیز بین $[0, H_{max}]$). پس از آموزش شبکه عصبی بر مبنای این پارامترها، وزن‌های شبکه در طی فرایند یادگیری براساس داده‌های آموزشی (روش BP) تنظیم شده و وزن‌های جدید بر مبنای وزن‌های آموزش دیده شده، مقداردهی می‌شود. با توجه به ثابت بودن بخش مشخصات لایه‌های پنهان در فرایند آموزش شبکه عصبی، باید بهینه‌ترین تنظیمات (ابعاد شبکه) برای شبکه عصبی جستجو شود. برای این کار از یک فرایند بهینه‌سازی محلی مطابق شکل ۲ استفاده شده است.

دقت خروجی شبکه عصبی روی داده‌های آموزشی برازندگی راه‌حل‌ها در فرایند بهینه‌سازی است. هر راه‌حل شامل ترکیبی از ویژگی‌های انتخابی، وزن‌های شبکه عصبی و مشخصات لایه‌های پنهان است. به منظور به‌روزرسانی مشخصات لایه‌های پنهان در هر تکرار به احتمال Γ مقدار هر بخش یک واحد کم یا زیاد می‌شود. این بخش‌ها شامل «تعداد لایه‌های پنهان r» و «اندازه نرون‌های لایه پنهان H_j، جائی که $j = 1,2, \dots, r$» می‌باشند. روابط (۴ و ۵) بروزرسانی هر بخش را نشان می‌دهد.

$$r_{new} = \begin{cases} r & \Gamma < RND \\ r \pm 1 & otherwise \end{cases} \quad (۴)$$

$$H_{new,j} = \begin{cases} H_j & \Gamma < RND \\ H_j \pm 1 & otherwise \end{cases} \quad (۵)$$

در نهایت به‌روزرسانی ویژگی‌های انتخاب شده انجام می-شود. در این مرحله نیز یک راه‌حل جدید تولید شده و به عنوان ورودی شبکه عصبی در نظر گرفته می‌شود. ویژگی-های انتخابی جدید از یک روش ابتکاری و از طریق ویژگی-های انتخاب شده جاری تولید می‌شود (همسایه راه‌حل جاری). تعیین ویژگی‌های انتخابی جدید از طریق میزان سودمندی ویژگی‌ها انجام می‌شود. مراحل این کار به صورت زیر می‌باشد:

۱. مجموعه داده اصلی را تنها با ویژگی‌های انتخابی (از راه‌حل جاری) در S قرار بده.
۲. آنتروپی مجموعه داده S را محاسبه کن.
۳. میزان سودمندی (بهره) اطلاعات همه ویژگی-های مجموعه داده S محاسبه شود.
۴. مجموعه داده اصلی تنها با ویژگی‌های انتخاب نشده در S' کپی شود.
۵. برای ویژگی‌های انتخاب شده جدید؛ به احتمال Γ ویژگی با کمترین سودمندی از بخش ویژگی-های انتخاب شده حذف می‌شود و به احتمال Γ ویژگی‌ای که دارای بیشترین سودمندی در بخش ویژگی‌های انتخاب نشده است، به بخش ویژگی-های انتخاب اضافه می‌شود.

۴- نتایج و بحث

برای شبیه‌سازی از نرم‌افزار MATLAB ورژن 2016a روی یک PC با پیکربندی: پردازنده اینتل ۷ هسته‌ای با فرکانس ۳,۰ گیگاهرتز، حافظه ۸ گیگابایت و سیستم عامل ویندوز ۱۰ نسخه ۶۴ بیتی استفاده شده است.

فرایند اعتبارسنجی 10-Fold است که روی مجموعه داده NSL-KDD اعمال شده است. اطلاعات این مجموعه داده از مرجع [۲۰] دردسترس است. مجموعه داده NSL-KDD از سه بخش Test Set، Train set و Percent Training Set مطابق جدول ۱ تشکیل شده است. جدول ۲، برچسب کلاس‌های این مجموعه داده را نشان می‌دهد.

جدول ۱: اطلاعات بخش‌های مجموعه داده NSL-KDD

بخش‌ها	تعداد ویژگی‌ها	تعداد نمونه‌ها	تعداد کلاس
Test Set	۴۱	۲۲۵۴۳	۵ کلاس
Train set	۴۱	۱۲۵۹۷۳	۵ کلاس
Percent Training Set	۴۱	۲۵۱۹۳	۲ کلاس

جدول ۲: کلاس‌های مجموعه داده NSL-KDD

نوع حمله	تعداد نمونه آزمایش	تعداد نمونه آموزش
Normal	۹۷۱۰	۶۷۳۴۳
DOS	۷۴۵۸	۴۵۹۲۷
R2L	۲۷۵۴	۹۹۵
U2R	۲۰۰	۵۲
Probe	۲۴۲۱	۱۱۶۵۶

پارامترهای الگوریتم پیشنهادی مطابق زیر به صورت سعی و خطا در بهترین حالت تنظیم شده است.

$$\Gamma = 0.4, epoch = 1000, r_{max} = 10, H_{max} = 10$$

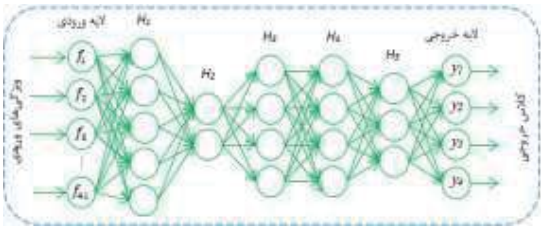
در این تحقیق از معیارهای مختلفی برای ارزیابی روش پیشنهادی مطابق جدول ۳ استفاده شده است.

جدول ۳ : معیارهای ارزیابی

معیارها	رابطه
دقت	$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$
درستی	$Precision = \frac{TP}{TP + FP}$
یاداور	$Recall = \frac{TP}{TP + FN}$
اندازه‌گیری F	$F - Measure = \frac{2 \times Precision \times Recall}{Precision + Recall}$
اتصالات	$Connection = \alpha \times \beta + \beta \times Y + \beta + Y$ α تعداد ویژگی‌ها، β اندازه لایه مخفی و Y اندازه لایه خروجی
زمان اجرا	زمان اجرای الگوریتم بر حسب ثانیه

در این تحقیق از دو مدل طبقه‌بندی کلاسیک شبکه عصبی BPNN و SVM و همچنین روش‌های MARS [۲۱] و DLNN [۲۲] که به تازگی ارائه شده‌اند، جهت مقایسه با روش پیشنهادی استفاده شده است. MARS توسط موباراک (۲۰۱۶) ارائه شد که در آن از الگوریتم‌های متفاوتی برای تشخیص نفوذ استفاده کرد که از جمله میتوان به ترکیب SVM و شبکه عصبی SOM اشاره کرد. DLNN نیز توسط روی و همکاران (۲۰۱۷) ارائه شد. این الگوریتم یک شبکه عصبی مصنوعی مبتنی بر یادگیری عمیق برای تشخیص نفوذ است که از ماشین‌های بولتزمن محدود و یک الگوریتم حریصانه برای یادگیری استفاده می‌کند.

نتایج شبیه‌سازی نشان می‌دهد که شبکه عصبی نهایی مطابق شکل ۳ دارای ۵ لایه مخفی به ترتیب با تعداد نرون‌های ۵، ۲، ۴، ۴ و ۳ می‌باشد.



شکل ۳: ساختار شبکه عصبی نهایی

در این تحقیق مقایسه‌ای بین تکنیک‌های پس انتشار خطا RP، LM و GD ارائه شده است. جدول ۴ نتایج الگوریتم پیشنهادی با تکنیک‌های مختلف پس انتشار خطا در معیارهای دقت، درستی، یاداور و اندازه‌گیری F را نشان می‌دهد.

جدول ۴: نتایج الگوریتم پیشنهادی با روش‌های مختلف یادگیری پس انتشار خطا

تکنیک‌های یادگیری	دقت	درستی	یاداور	اندازه‌گیری F
RP	۹۷,۶۴	۹۸,۲۳	۹۵,۸۹	۹۷,۰۵
LM	۹۶,۸۹	۹۷,۵۰	۹۵,۰۶	۹۶,۲۶
GD	۹۷,۱۱	۹۸,۰۰	۹۵,۶۴	۹۶,۸۱

با توجه به این آزمایش تکنیک پس انتشار خطا RP بهترین نتایج را نسبت به دو تکنیک LM و GD بدست آورده است. در اینجا دقت روش پیشنهادی برمبنای تکنیک RP به طور متوسط ۹۷,۶۴٪ است. این درحالی است که روش‌های LM و GD به ترتیب به دقت‌های ۹۶,۸۹٪ و ۹۷,۱۱٪ رسیده

است. نتایج مربوط به معیارهای اتصالات و زمان اجرا در جدول ۵ گزارش شده است.

جدول ۵: نتایج معیارهای اتصالات و زمان اجرا با روش‌های مختلف یادگیری پس انتشار خطا

تکنیک‌های یادگیری	اتصالات	زمان اجرا (ثانیه)
RP	۱۳۱	۸۸۳
LM	۱۳۵	۹۷۵
GD	۱۲۴	۹۱۳

نتایج این آزمایش نشان می‌دهد که تکنیک یادگیری RP که دارای بهترین دقت تشخیص نفوذ برای روش پیشنهادی بود، دارای پیچیدگی اتصالات ۱۳۱ می‌باشد و این در حالیست که بهترین شبکه تولید شده از طریق GD تنها دارای ۱۲۴ اتصال است. همچنین نتایج زمان اجرا نشان می‌دهد که استفاده از تکنیک RP برای ساخت مدل طبقه-بندی با متوسط زمان اجرا ۸۸۳ ثانیه نسبت به دو تکنیک دیگر سریع‌تر است. به طور کلی با توجه به نتایج حاصل شده، تکنیک RP نسبت به دو تکنیک دیگر برای تشخیص نفوذ مناسب‌تر است و از این رو برای کار مقایسه از این تکنیک استفاده می‌شود. نتایج مقایسه روش پیشنهادی با مدل‌های طبقه‌بندی SVM و BPNN و همچنین دو الگوریتم MARS و DLNN در جدول ۶ ارائه شده است. نتایج برای روش‌های مختلف در شرایط یکسانی روی مجموعه داده NSL-KDD محاسبه شده و برمبنای کلاس‌های مختلف مجموعه داده گزارش شده است.

جدول ۶: مقایسه نتایج روش پیشنهادی با سایر روش‌ها به تفکیک نوع کلاس

الگوریتم‌ها	Normal	DOS	U2R	R2L	Probe	اتصالات	دقت
SVM	۷۴,۱۲	۶۱,۲۰	۶۴,۹۱	۷۱,۰۶	۷۳,۱۸	-	۷۳,۸۰
BPNN	۷۱,۴۵	۶۰,۲۳	۶۲,۸۷	۶۹,۶۵	۷۱,۹۰	-	۷۱,۰۵
MARS	۹۰,۷۵	۹۶,۸۵	۹۵,۷۵	۷۳,۰۰	۹۶,۹۷	-	۹۶,۷۱
DLNN	۹۷,۳۲	۹۹,۱۷	۷۴,۶۵	۸۱,۵۱	۹۹,۷۸	۱۴۰	۹۷,۴۱
روش پیشنهادی	۹۸,۰۱	۹۹,۲۴	۶۶,۵۰	۹۴,۶۳	۹۶,۷۰	۱۳۱	۹۷,۶۴

نتایج نشان دهنده برتری، مدل طبقه‌بندی کلاسیک SVM نسبت به BPNN است. همچنین از بین دو الگوریتم MARS و DLNN، الگوریتم DLNN با دقت ۹۷,۴۱٪ عملکرد بهتری برای تشخیص نفوذ داشته است. هرچند نتایج در حالت میانگین برتری روش پیشنهادی را نسبت به مدل‌های طبقه‌بندی SVM و BPNN به ترتیب با ۳۲,۳۰٪ و ۳۷,۴۲٪ در معیار دقت نشان می‌دهد. به طور متوسط این برتری نسبت به روش‌های MARS و DLNN به ترتیب ۰,۹۶٪ و ۰,۲۴٪ است. دلیل برتری روش پیشنهادی بهینه-سازی همزمان پارامترهای شبکه عصبی (اندازه لایه‌های مخفی، تعداد نرون‌های هر لایه و همچنین مقادیر وزن‌ها) می‌باشد که از طریق یادگیری عمیق حاصل شده است. نتایج این مقایسه بر مبنای معیار تعداد اتصالات شبکه تنها بین روش پیشنهادی و DLNN گزارش شده است، چون تنها این دو روش به دنبال بهترین ساختار شبکه عصبی برای ایجاد مدل طبقه‌بندی می‌باشد. در هر حال روش پیشنهادی شبکه فشرده‌تری نسبت به روش DLNN تولید می‌نماید که این مورد از نظر محاسباتی برتری قابل توجهی برای روش پیشنهای است. روش پیشنهادی با ایجاد یک شبکه عصبی با تعداد اتصالات ۱۳۱ نسبت به روش DLNN حدود ۶,۸۷٪ برتری داشته است.

۵- نتیجه‌گیری

وظیفه سیستم‌های تشخیص نفوذ، شناسایی حملات به شبکه‌های کامپیوتری، تشخیص ضعف‌های کلی سیستم و گزارش آن به مدیر سیستم است. در این تحقیق برای افزایش دقت سیستم تشخیص نفوذ از یک تکنیک یادگیری عمیق مبتنی بر شبکه‌های عصبی استفاده شد که نتایج قابل قبولی را ارائه می‌دهد. رویکرد یادگیری با ناظر ارائه شده به منظور آموزش پارامترهای شبکه عصبی عمیق از ترکیب خطی و بازنمایی ویژگی‌های مؤثر استفاده می‌کند.

این روش بر مبنای یک الگوریتم یادگیری با ناظر حریصانه و یک شبکه عصبی عمیق می‌باشد. ساختار شبکه عصبی نهایی برای مجموعه داده NSL-KDD با ۵ لایه مخفی به ترتیب با تعداد نرون‌های ۵، ۲، ۴، ۴ و ۳ ایجاد شده که به نسبت الگوریتم‌های MARS و DLNN شبکه کوچک‌تر با دقت بالاتری ارائه شده است.

مراجع

[1] M. A. Faisal, Z. Aung, J. R. Williams, and A. Sanchez, "Data-Stream-Based Intrusion Detection System for Advanced Metering Infrastructure in Smart Grid: A Feasibility Study". IEEE Systems journal, Vol. 9, No. (1), pp. 31-44, 2015.

[2] R. A. R. Ashfaq, Z. 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.

[3] P. Garcia-Teodoro, J. Diaz-Verdejo, G. Maciá-Fernández, and E. Vázquez, "Anomaly-based network intrusion detection: Techniques, systems and challenges". computers & security, Vol. 28, No. 1, pp. 18-28, 2009.

[4] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning. Nature", Vol. 521, no. 7553, pp. 436, 2005.

[5] Y. Yang, H. Q. Xu, L. Gao, Y. B. Yuan, K. McLaughlin, anf S. Sezer, "Multidimensional intrusion detection system for IEC 61850-based SCADA networks". IEEE Transactions on Power Delivery, Vol. 32, No. 2, pp. 1068-1078, 2017.

[6] Q. Wang, Y. Jin, Z. Lin, S. Cheng, and T. Yang, "Influence maximization in social networks under an independent cascade-based model". Physica A: Statistical Mechanics and its Applications, Vol. 444, pp. 20-34, 2016.

[7] M. Karimi, H. Mokhlis, K. Naidu, S. Uddin, and A. H. A. Bakar, "Photovoltaic penetration issues and impacts in distribution network–A review". Renewable and Sustainable Energy Reviews, Vol. 53, pp. 594-605, 2016

[8] O. Al-Jarrah, and A. Arafat, "Network Intrusion Detection System using attack behavior classification". In Information and Communication Systems (ICICS), 2014 5th International Conference on (pp. 1-6). IEEE, 2014.

[9] R. Borgohain, "Fugeids: Fuzzy genetic paradigms in intrusion detection systems". arXiv preprint arXiv:1204.6416, 2012.

[10] 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, No. 18, pp. 13492-13500, 2012.

[11] K. Harinee, and A. Veeramuthu, "Intrusion Detection System Based on Fuzzy Association Rule with Genetic Network Programming", International Journal of Innovative Research in Science, Engineering and Technology, Vol. 3, No. 4, pp. 11077-11083, 2014.

[12] G. Wang, J. Hao, J. Ma, and L. Huang, "A new approach to intrusion detection using Artificial Neural Networks and fuzzy clustering". Expert systems with applications, Vol. 37, No. 9, pp. 6225-6232, 2010.

[13] A. K. Ghosh, A. Schwartzbard, and M. Schatz, "Using Program Behavior Pro les for Intrusion Detection". 2014

[14] V. Markam, and L. S. M. Dubey, "A general study of associations rule mining in intrusion detection system". International Journal of Emerging Technology and Advanced Engineering, Vol. 2, No. 1, pp. 347-356, 2012.

[15] B. B. Zarpelão, R. S. Miani, C. T. Kawakani, and S. C. Alvarenga, "A survey of intrusion detection in Internet of Things". Journal of Network and Computer Applications, Vol. 84, pp. 25-37, 2017.

[16] M. Derfouf, M. Eleuldj, S. Enniari, and O. Diouri, "Smart intrusion detection model for the cloud computing". In Europe and MENA Cooperation Advances in Information and Communication Technologies (pp. 411-421). Springer, Cham, 2017.

[17] V. Barot, and D. Toshniwal, "A new data mining based hybrid network Intrusion Detection model". In Data Science & Engineering (ICDSE), 2012 International Conference on (pp. 52-57). IEEE, 2012.

[18] M. Ghalehgolabi, and A. Rezaeipanah, "Intrusion Detection System Using Genetic Algorithm and Data Mining Techniques Based on the Reduction". International Journal of Computer Applications Technology and Research, Vol. 6, no. 12, pp. 1-6, 2017.

[19] S. Warsi, R. Yogesh, and S. Kushwaha, "Selective Iteration based Particle Swarm Optimization (SIPSO) for Intrusion Detection

System". International Journal of Computer Applications, Vol. 124, No. 17, 2015.
[20] http://nsl.cs.unb.ca/KDD/NSL-KDD.html
[21] S. L. Mubarak, "Intrusion Detection System using SVM, SOM& NN", 2016.
[22] S. S. Roy, A. Mallik, R. Gulati, M. S. Obaidat, and P. V. Krishna, "A deep learning based artificial neural network approach for intrusion detection". In International Conference on Mathematics and Computing (pp. 44-53). Springer, Singapore, 2017.