
بسم الله الرحمن الرحيم

پایان‌نامه برای دریافت درجه کارشناسی ارشد«MSc»

گرایش: مهندسی نرم‌افزار

عنوان:

ارائه یک سیستم تشخیص نفوذ در شبکه‌های کامپیوتری با استفاده از الگوریتم رقابت استعماری و شبکه عصبی

گروه مهندسی کامپیوتر

پایان‌نامه برای دریافت درجه کارشناسی ارشد «M.Sc.»

گرایش : مهندسی نرم‌افزار

عنوان:

ارائه یک سیستم تشخیص نفوذ در شبکه‌های کامپیوتری با استفاده از الگوریتم رقابت استعماری و شبکه عصبی

هیئت محترم داوران:

نام و نام خانوادگی استاد راهنما	تاریخ	امضا
نام و نام خانوادگی استاد داور	تاریخ	امضا
نام و نام خانوادگی استاد داور	تاریخ	امضا

تعهدنامه اصالت رساله دکتری یا پایان‌نامه کارشناسی ارشد دانشگاه آزاد اسلامی

اینجانب .....دانش‌آموخته مقطع کارشناسی ارشد ناپیوسته در رشته مهندسی نرم افزار که در تاریخ از پایان‌نامه / رساله خود تحت عنوان" ارائه یک سیستم تشخیص نفوذ در شبکه‌های کامپیوتری با استفاده از الگوریتم رقابت استعماری و شبکه عصبی" با کسب نمره ..... و درجه..... دفاع نموده‌ام بدین‌وسیله متعهد می‌شوم :

۱ – این پایان‌نامه / رساله حاصل تحقیق و پژوهش انجام‌شده توسط این‌جانب بوده و در مواردی که از دستاوردهای علمی و پژوهشی دیگران (اعم از پایان‌نامه، کتاب، مقاله و .....) استفاده نموده‌ام، مطابق ضوابط و رویه موجود، نام منبع مورداستفاده و سایر مشخصات آن را در فهرست مربوطه ذکر و درج کرده‌ام.

۲ – این پایان‌نامه/ رساله قبلاً برای دریافت هیچ مدرک تحصیلی (هم‌سطح، پائین تر یا بالاتر) در سایر دانشگاه‌ها و مؤسسات آموزش عالی ارائه نشده است.

۳ چنانچه بعد از فراغت از تحصیل، قصد استفاده و هرگونه بهره‌برداری اعم چاپ کتاب، ثبت اختراع، و..... از این پایان‌نامه را داشته باشم، از حوزه معاونت پژوهشی واحد مجوزهای مربوطه را اخذ نمایم.

٤ – چنانچه در هر مقطع زمانی خلاف موارد فوق ثابت شود، عواقب ناشی از آن را می‌پذیرم و واحد دانشگاهی مجاز است با این‌جانب مطابق ضوابط و مقررات رفتار نموده و در صورت ابطال مدرک تحصیلی‌ام هیچ‌گونه ادعایی نخواهم داشت.

نام و نام خانوادگی

تاریخ و امضاء

به نکته زیر توجه شود

نمره از ۱۸,۱ تا ۲۰ با درجه عالی

نمره از ۱٦,۱ تا ۱۸ با درجه بسیار خوب

نمره از ۱٤,۱ تا ۱٦ با درجه خوب

معاونت پژوهش و فن آوری

به نام خدا

منشور اخلاق پژوهش

با یاری از خداوند سبحان و اعتقاد به اینکه عالم محضر خداست و همواره ناظر بر اعمال انسان و به منظور پاسداشت مقام بلند دانش و پژوهش و نظر به اهمیت جایگاه دانشگاه در اعتلای فرهنگ و تمدن بشری ما دانشجویان و اعضاء هیئت علمی واحدهای دانشگاه آزاد اسلامی متعهد می گردیم اصول زیر را در انجام فعالیت های پژوهشی مد نظر قرار داده و از آن تخطی نکنیم:

۱-اصل حقیقت جویی: تلاش در راستای پی جویی حقیقت و وفاداری به آن و دوری از هرگونه پنهان سازی حقیقت

۲-اصل رعایت حقوق: التزام به رعایت کامل حقوق پژوهشگران و پژوهیدگان (انسان، حیوان و نبات) و سایر صاحبان حق

۳-اصل مالکیت مادی و معنوی: تعهد به رعایت کامل حقوق مادی و معنوی دانشگاه و کلیه همکاران پژوهش

۴-اصل منافع ملی: تعهد به رعایت مصالح ملی و در نظر داشتن پیشبرد و توسعه کشور در کلیه مراحل پژوهش

۵-اصل رعایت انصاف و امانت: تعهد به اجتناب از هرگونه جانب داری غیر علمی و حفاظت از اموال، تجهیزات و منابع در اختیار

۶-اصل رازداری: تعهد به صیانت از اسرار و اطلاعات محرمانه افراد، سازمان ها و کشور و کلیه افراد و ستادهای مرتبط با تحقیق

۷-اصل احترام: تعهد به رعایت حریم ها و حرمت ها در انجام تحقیقات و رعایت جانب نقد و خودداری از هرگونه حرمت شکنی

۸-اصل ترویج: تعهد به رواج دانش و اشاعه نتایج تحقیقات و انتقال آن به همکاران علمی و دانشجویان به غیر از مواردی که منع قانونی دارد

۹-اصل برائت: التزام به برائت جویی از هرگونه رفتار غیر حرفه ای و اعلام موضع نسبت به کسانی که حوزه علم و پژوهش را به شائبه های غیر علمی می آلایند.

فهرست مطالب

عنوان | صفحه

فهرست مطالب

عنوان | صفحه

فهرست مطالب

عنوان | صفحه

فهرست شکل‌ها

عنوان صفحه

فهرست جداول

عنوان صفحه

چکیده

برای ایجاد امنیت داده‌ها در یک سیستم کامپیوتری علاوه بر دیواره‌های آتش و دیگر تجهیزات جلوگیری از نفوذ، نیاز به سیستم‌های دیگری به‌نام تشخیص نفوذ وجود دارد تا بتواند در صورتی‌که نفوذگر از دیوار آتش و دیگر تجهیزات امنیتی عبور کرد، آن‌را تشخیص دهد. سیستم‌های تشخیص نفوذ یکی از ابزارهای پرکاربرد در امنیت شبکه‌های کامپیوتری می‌باشند. از آن‌جایی که سیستم‌های تشخیص نفوذ نیازمند سرعت و دقت بالایی هستند و با حجم عظیم داده‌ها کارایی خود را از دست می‌دهند، بنابراین یادگیری ساختار داده و انتخاب ویژگی‌ها دارای اهمیت بالایی می‌باشد. هدف از انتخاب ویژگی، یافتن زیرمجموعه ویژگی‌های موثر مجموعه داده‌ی اولیه، به منظور افزایش دقت و کاهش هزینه طبقه‌بندی داده‌ها است. در این پایان‌نامه، به منظور کاهش ویژگی‌ها از الگوریتم رقابت استعماری و شبکه عصبی به عنوان مراحل جستجو و تابع هزینه الگوریتم انتخاب ویژگی استفاده شده است. نتایج شبیه‌سازی نشان می‌دهد روش پیشنهادی با استفاده از الگوریتم طبقه‌بندی ترکیبی ماشین بردار پشتیبان و شبکه عصبی دارای دقت بالایی است.

کلمات کلیدی: تشخیص نفوذ، انتخاب ویژگی، رقابت استعماری، مجموعه داده NSL-KDD

فصل اول

کلیات تحقیق

۱-۱- مقدمه

با توجه به رشد فناوری‌های اطلاعاتی در زندگی روزمره، امنیت رایانه‌ای به یکی از ضروریات تبدیل شده است. استفاده وسیع از سیستم‌های رایانه‌ای، باعث افزایش تهدیدات حیاتی مانند آسیب‌پذیری‌های روز صفر[1]، تهدیدات موبایل و غیره شده است. با وجود رشد چشمگیر تحقیقات حوزه امنیتی، هنوز شاهد کاهش این تهدیدات نبوده‌ایم. تکامل شبکه‌های کامپیوتری، نگرانی‌های امنیت کامپیوتری را تا حد زیادی افزایش داده است، مخصوصا امنیت کامپیوتری در محیط شبکه‌های امروزی و تاسیسات رایانشی پیشرفته. اگرچه پروتکل‌های اینترنتی، به گونه‌ای طراحی نشده‌اند که اولویت بالایی به مسائل امنیتی بدهند ولی مدیران شبکه، اکنون ملزم به اداره طیف وسیعی از تلاش‌های نفوذ از سوی اشخاص با نیات بدخواهانه شده‌اند (اسنپ[2] و همکاران، ۲۰۱۷). بر اساس گزارش تهدید امنیتی اینترنت سمانتیک[3]، بیش از سه میلیارد حمله بدافزاری در سال ۲۰۱۳ گزارش شده و تعداد حملات منع سرویس در سال ۲۰۱٦، رشد باورنکردنی داشته است (گزارش تهدید امنیتی اینترنتی سمانتیک، ۲۰۱۷). همانطور که در گزارش بررسی نقض داده وریزن[4] ۲۰۱٦، ذکر شده ٦۳٤۳۷ نقض امنیتی، بوسیله هکرها انجام شده است (گزارش بررسی نقض داده وریزن، ۲۰۱٦). بررسی امنیت اطلاعات جهانی ۲۰۱۵ نیز رشد بزرگ حوادث را گزارش کرده است. بنابراین تشخیص حملات شبکه، به بزرگترین اولویت امروزی تبدیل شده است. بعلاوه تخصص لازم برای ارتکاب جرائم سایبری، به‌دلیل دسترسی آسان ابزار مورد نیاز، کاهش یافته است (ابزار هک و کرک، ۲۰۱٤).

[1] Zero day
[2] Snapp
[3] Symantecs
[4] Verizon

تشخیص ناهنجاری[1] یکی از وظایف مهم تحلیل داده می‌باشد که داده‌های غیرعادی یا ناهنجار را از مجموعه داده، تشخیص می‌دهد. این حوزه، یکی از حوزه‌های تحقیقی مهم داده کاوی بشمار می‌رود، زیرا شامل کشف و مقابله با الگوهای نادر در داده‌ها می‌باشد و در سطح وسیعی در حوزه آمار و یادگیری ماشین مطالعه شده است (احمد و همکاران، ۲۰۱٤). دیگر نام‌های این حوزه شامل تشخیص نقاط پرت، تشخیص انحراف و کاوش استثنا می‌باشد. اگرچه ناهنجاری توسط محققین به گونه‌های مختلفی بر اساس حوزه‌های کاربردی آن، تعریف شده ولی تعریف پذیرفته شده در سطح وسیع به هاوکینز (۱۹۸۰) نسبت داده شده است: ناهنجاری، به مشاهده‌ای گفته می‌شود که انحراف زیادی از دیگر مشاهدات دارد، تا حدی که این شک را بوجود می‌آورد که مشاهده بوسیله مکانیزم متفاوتی، تولید شده است". ناهنجاری، اهمیت زیادی دارد زیرا نشان‌دهنده رویدادهای مهم ولی نادر بوده و می‌تواند موجب انجام اقدامات حیاتی در طیف وسیعی از حوزه‌های کاربردی شود؛ بعنوان مثال، الگوی ترافیک غیرعادی در شبکه، می‌تواند به‌معنای آن باشد که رایانه، هک شده و داده‌ها به مقاصد ناخواسته انتقال یافته‌اند. رفتار ناهنجاری در تراکنش‌های کارت اعتباری، می‌تواند نشان‌دهنده کلاهبرداری باشد و ناهنجاری در تصویر MRI، ممکن است به‌معنای وجود تومور بدخیم باشد (احمد و همکاران، ۲۰۱۵)

۲-۱- بیان مسئله

نفوذ به مجموعه اقدامات غیرقانونی اطلاق می‌شود که صحت و محرمانگی داده‌ها را به خطر می‌اندازند و یا دسترسی به یک منبع را دچار اختلال می‌نمایند. به منظور مقابله با نفوذگران به سیستم‌ها و شبکه‌های کامپیوتری، روش‌های مختلفی وجود دارند که یکی از آن‌ها سیستم‌های تشخیص نفوذ می‌باشند.سیستم‌های تشخیص نفوذ عمل نظارت بر وقایع و اتفاقات رخ داده در یک سیستم یا شبکه کامپیوتری را بر عهده دارند.

[1] Anomaly detection

هدف سیستم تشخیص نفوذ تعیین این مسئله است، که آیا یک سیستم یا شبکه‌ی کامپیوتری مورد استفاده‌ی غیر مجاز قرار گرفته است یا خیر. سیستم‌های تشخیص نفوذ در مقابل حملات اقدامی انجام نمی‌دهند و صرفاً وقوع حملات را با اعلان هشدار به مدیر سیستم اطلاع می‌دهند. اعلان هشدار می‌توانند از طریق ثبت تاریخچه-ی وقایع، ارسال ایمیل، پیام کوتاه و ... انجام شود. سیستم تشخیص نفوذ اغلب از دو روش تشخیص سوء استفاده و تشخیص ناهنجاری به منظور تشخیص نفوذ استفاده می‌کنند. سیستم‌های تشخیص سوء استفاده، پایگاه دانشی از الگوی حملات دارند و چنان‌چه ترافیک شبکه با این الگو مطابقت داشته باشد، آن ترافیک را حمله، تشخیص می‌دهند. روش‌های تشخیص سوء استفاده، نرخ تشخیص بالایی دارند اما توانایی شناسایی حملات جدید را ندارند. سیستم‌های تشخیص ناهنجاری بر این فرض استوارند که هر فعالیت ناهنجار یک فعالیت بدخواهانه است. بنابراین ابتدا یک مدل از رفتارهای هنجار ایجاد می‌شود و سپس هر رفتاری که با آن مغایرت داشت به عنوان ناهنجاری شناخته می‌شود. مزیت اصلی این گونه سیستم‌ها، این است که امکان تشخیص حملات ناشناخته را دارند، اما نرخ هشدار غلط بالایی دارند. تاکنون روش‌های مختلفی به منظور افزایش نرخ تشخیص و کاهش نرخ هشدار غلط در تشخیص ناهنجاری‌ها مورد استفاده و آزمایش قرار گرفته است؛ اما هیچ یک نتوانسته‌اند به موفقیت کامل دست یابند و تحقیقات در این زمینه همچنان ادامه دارد.

با توجه به این‌که در سیستم‌های تشخیص نفوذ دو معیار دقت در تشخیص نفوذ به شبکه‌های کامپیوتری و همچنین کاهش تعداد هشدارهای نادرست از اهمیت زیادی برخوردار است، الگوریتم‌های داده کاوی می‌توانند در تحقق آن کمک شایانی کنند.

۳-۱- اهمیت موضوع

با روند رو به رشد استفاده از شبکه‌های کامپیوتری به خصوص اینترنت و مهارت رو به رشد کاربران و مهاجمان این شبکه‌ها و وجود نقاط آسیب‌پذیری مختلف در نرم‌افزارها، ایمن‌سازی سیستم‌ها و شبکه‌های کامپیوتری، نسبت به گذشته از اهمیت بیشتری برخوردار شده است. تامین امنیت در هر سیستم یا شبکه

کامپیوتری در واقع به معنای سه بعد اساسی محرمانگی، جامعیت و دسترس‌پذیری در آن می‌باشد که می‌توان به دو صورت، یکی از طریق پیشگیری امنیتی و دیگری از طریق تشخیص تخطی از سیاست‌های امنیتی انجام پذیرد. اکثر تحقیقاتی که تاکنون در زمینه امنیت کامپیوتر صورت پذیرفته است، تمرکزشان بر روی تامین امنیت با استفاده از روش‌های پیشگیرانه بوده است. در حالی که به غیر از پیشگیری باید درپی درمان نیز بود (بوتون[1] و همکاران، ۲۰۱۴). لذا تحقیق و بررسی بر روی سیستم‌های غیر پیشگیرانه همچون سیستم تشخیص نفوذ که وظیفه‌ی آن‌ها تشخیص حمله و رفتارهای ناهنجار در سیستم‌ها و شبکه‌های کامپیوتری می‌باشد، نیز از اهمیت و جایگاه خاصی برخوردار است.

۴-۱- اهداف پایان نامه

این پایان‌نامه، تشخیص ناهنجاری‌های شبکه مورد بحث فرار گرفته است. هدف اصلی، ارائه مدلی برای تشخیص ناهنجاری شبکه با استفاده از تکنیک‌های داده‌کاوی است. از آن‌جایی که از تکنیک‌های داده‌کاوی در بخش‌های مختلفی از تشخیص نفوذ استفاده شده است و هر کدام از این بخش‌ها به نوبـه خود اهمیت ویژه‌ای در موضوع تشخیص نفوذ دارند و همچنین در اکثر تحقیقات، از تکنیک‌های خاصی از داده کاوی استفاده شده و تنها بر بخش خاصی از تشخیص نفوذ متمرکز شده‌اند، لذا در این پایان‌نامه چند تکنیک داده‌کاوی، ارائه شده است، مهم‌ترین اهداف این تکنیک‌ها عبارتست از:

- کاهش ابعاد داده‌ها و انتخاب ویژگی‌های مهم و موثر در شناسایی نفوذها.
- افزایش دقت در تشخیص نفوذ، به واسطه حذف داده‌های اضافی و نامربوط و انتخاب ویژگی‌های موثر در هر حمله.
- افزایش سرعت و به‌حداقل رساندن زمان شناسایی حملات و نفوذها.

[1] Butun

- کاهش نرخ هشدار غلط در روش تشخیص ناهنجاری‌های شبکه.
- تأمین امنیت اطلاعات و سیستم‌ها در برابر نفوذها.

۵-۱- نوآوری پایان‌نامه

این پایان‌نامه ارائه یک روش نوین برای تشخیص نفوذ در سیستم‌ها و شبکه‌های کامپیوتری می‌باشد. سیستم تشخیص نفوذ پیشنهادی، یک سیستم تشخیص نفوذ مبتنی بر تشخیص ناهنجاری است. مدل پیشنهادی در این پایان نامه شامل چهار فاز می‌باشد. فاز اول شامل آماده‌سازی و پیش پردازش داده است که در این مرحله از فیلتر مقادیر از دست رفته برای جایگزینی تمام مقادیر ویژگی‌های از دست رفته در مجموعه داده استفاده می‌شود. این فیلتر تمام مقادیر گم شده را با میانگین داده‌های آموزشی جایگزین می‌کند. در مرحله دوم، با توجه به این‌که تعداد ویژگی‌ها زیاد هستند، جهت کاهش آن‌ها و انتخاب بهترین آن‌ها در تشخیص از ترکیب الگوریتم‌های بهینه‌سازی رقابت استعماری و شبکه عصبی در انتخاب ویژگی استفاده خواهد شد. در مرحله سوم از چندین طبقه‌بند استفاده می‌شود و سپس بر اساس رأی اکثریت کلاس‌بندی داده‌ها صورت می‌پذیرد.

٦-۱- ساختار پایان‌نامه

این پایان‌نامه با احتساب این فصل، در پنج فصل گردآوری شده است. خلاصه‌ای از رئوس مطالب اصلی هر فصل در ادامه ارائه‌شده است. در فصل دوم، مبانی تحقیق و هم‌چنین معرفی نفوذ را خواهیم داشت. فصل سوم پیشینه تحقیق بررسی‌شده است. فصل چهارم معماری سیستم پیشنهادی موردبحث قرارگرفته می‌شود. در فصل پنجم، شبیه‌سازی و ارزیابی روش پیشنهادی مطرح خواهد شد، درنهایت نتیجه‌گیری کلی از پایان‌نامه گرفته‌شده و پیشنهادهایی برای توسعه و بهبود روش‌های پیشنهادی برای کارهای آتی، ارائه می‌شود.

فصل دوم

مبانی تحقیق

با توجه به رشد روزافزون شبکه‌های کامپیوتری و افزایش میزان اهمیت آن‌ها در تجارت، دانش مهندسی و فنی و فن‌آوری‌های به‌روز، میزان حمله‌ها و نفوذها به شبکه نیز به طرز قابل توجهی افزایش یافته است. رشد اینترنت باعث شده تا منابع و اطلاعات به سادگی در اختیار همگان قرار گیرد. این اطلاعات می‌تواند مربوط به روش‌های حمله و نفوذ به شبکه‌ها باشد. دسترسی ساده به این اطلاعات باعث شده است تا هر فردی بتواند به سادگی حمله را تجربه نماید. به همین دلیل در یک سیستم توزیع شده‌ی کامپیوتری که مبتنی بر شبکه است، تأمین امنیت سیستم بسیار مهم است. در این فصل، ابتدا به نفوذ و سیستم‌های تشخیص نفوذ پرداخته شد، سپس داده کاوی شرح داده شد.

۱-۲- معماری سیستم‌های تشخیص نفوذ

معماری نرم‌افزار به معنی سازماندهی مؤلفه‌ها و تعیین چگونگی ارتباط میان آن‌هاست. معماری سیستم‌های تشخیص نفوذ ممکن است متمرکز، توزیع‌شده، ترکیبی و یا بر اساس مؤلفه باشد. در سیستم‌های متمرکز جمع‌آوری و تحلیل داده‌ها در تعداد محدودی از سیستم‌ها انجام می‌شود. در مقابل، در سیستم‌های توزیع شده وظایف میان سیستم‌های مختلف تقسیم می‌شود. توزیع شدگی در درجات مختلفی قابل بکارگیری است. می‌توان تنها جمع‌آوری اطلاعات را به صورت توزیع شده انجام داد. در این حالت سیستم‌های جمع‌آورنده‌ی اطلاعات، حسگر نیز گفته می‌شود. حسگرها اطلاعات دریافتی خود را به یک سیستم مرکزی برای تحلیل