3rd INTERNATIONAL CONFERENCE ON RESEARCH IN SCIENCE AND TECHNOLOGY
BERLIN . GERMANY 9 July, 2016
RSTCONF
سومین کنفرانس بین المللی پژوهش در علوم و تکنولوژی
برلین - آلمان ۱۹ تیرماه ۱۳۹۵

بررسی وچگونگی راهکارهای تشخیص نفوذ با استفاده از تکنیک های داده کاوی

سمیه میمندی

دانشجوی کارشناسی ارشد مهندسی نرم افزار کامپیوتر،دانشگاه آزاد اسلامی واحد لامرد

Email:meymandi62@gmail.com

حسین محمدی شیوه کش

دانشجوی کارشناسی ارشد مهندسی نرم افزار کامپیوتر،دانشگاه آزاد اسلامی واحد لامرد

Email:shivehkesh@yahoo.com

چکیده:

با توجه به رشد روزافزون فناوری اطلاعات ، مسئله امنیت به عنوان یکی از مباحث مهم و بزرگ مطرح می باشد. سیستم تشخیص نفوذ نقش مهمی در شبکه ها ایفا می کنند بنا بر این سیستم های تشخیص نفوذ قدیمی و فایروال ها که توسط بسیاری از سازمانها به منظور حفاظت از امنیت سیستم های اطلاعاتی به کار گرفته می شوند به تنهایی نمی توانند خود را با حملات جدید تطبیق دهند از این رو امروزه سیستم های تشخیص نفوذ مبتنی بر داده کاوی مطرح می شود. به منظور مقابله با نفوذگران به سیستم‌ها و شبکه‌های رایانه‌ای، روش‌های متعددی تحت عنوان روشهای تشخیص نفوذ ایجاد گردیده است روش‌های تشخیص مورد استفاده در سامانه‌های تشخیص نفوذ به دو دسته، روش تشخیص رفتار غیر عادی وروش تشخیص سوء استفاده تقسیم می‌شوند در این مقاله سعی شده تا با معرفی این دو دسته از روشهای تشخیص نفوذ و همچنین با توجه به این که بیشتر تکنیک های داده کاوی در زمینه تشخیص نفوذ با موفقیت مورد استفاده قرار گرفته اند به بررسی آن تکنیک ها وروشها پرداخته می شود.

واژگان کلیدی: تشخیص نفوذ ، تشخیص رفتار غیر عادی ، تشخیص روش سوء استفاده ، داده کاوی

1- مقدمه

با گسترش روزافزون تبادل اطلاعات واستفاده از سیستم های برخط میزان حملات ونفوذ در سیستم های اطلاعاتی افزایش یافته است . به منظور جلوگیری از سوء استفاده از اطلاعات و رخنه در سیستم های حساس ومهم باید سیستم های مذکور امن شوند. مفهوم امنیت به معنای محافظت از محرمانگی ، یکپارچگی ودسترس پذیری است . اما محرمانگی ، یکپارچگی ودسترس پذیری چیست؟

محرمانه بودن داده ها به معنای این است که داده های در حال انتقال در شبکه تنها باید توسط افرادی که به شیوه مناسب احراز هویت شده اند مورد دستیابی قرار بگیرد. حفظ جا معیت به این داده ها از لحظه ای که دریافت می شوند ، نباید هیچ اختلال و فقدانی چه به دلیل رویداد های تصادفی وچه به دلیل فعالیت های خرابکارانه داشته باشند. دسترس پذیری در حالت مشهود به معنای آن است که سیستم درمقابل حملات انکار سرویس باید قدرتمند وضد ضربه باشد (حمیدی و ضیاعی ، 1388).

با توجه به اهمیت سیستم های تشخیص نفوذ در ارتقا امنیت سیستم های اطلاعاتی و موقعیت آن در مراحل ایمن سازی می توان به کمک این سیستم ها ، عملیات نفوذ در یک سیستم را نافرجام باقی گذاشت . اهمیت سیستم های تشخیص نفوذ در آنجا به اوج خود می رسد که قبل از اینکه حمله به اتمام برسد وقوع آن را گزارش کند ومانع از آسیب های احتمالی شود . بنابراین تشخیص نفوذ عبارت است از فرآیند شناسایی وپاسخ به فعالیت های مخرب که به صورت هدفمند ، منابع شبکه را مورد تهاجم قرار می دهد. بر اساس تعریف موسسه استاندارد وتکنولوژی "فرآیند نظارت بر رویدادهیی که در سیستم کامپیوتر و با شبکه رخ می دهد تجزیه وتحلیل آنها در جهت یافتن نشانه هایی از نفوذ ، که به معنای دور زدن محرمانگی وجامعیت و دسترس پذیری و یا عبور از مکانیزم های امنیتی یک کامپیوتر یا شبکه است ، می باشد" (kabiri et a1،2005) .
سامانه های تشخیص نفوذ به صورت سامانه های نرم افزاری و سخت افزاری ایجاد شده و هر کدام مزایا و معایب خاص خود را دارند. سرعت و دقت از مزایای سیستم های سخت افزاری است و عدم شکست امنیتی آنها توسط نفوذ گران ، قابلیت دیگر این گونه سیستم ها می باشد. اما استفاده آسان از نرم افزار ، قابلیت سازگاری در شرایط نرم افزاری و تفاوت سیستم های عامل مختلف ، عمومیت بیشتری را به سامانه های نرم افزاری می دهد وعموما این گونه سیستم ها انتخاب مناسب تری هستند.

2-لزوم استفاده از سیستم های تشخیص نفوذ

در دنیای امروز ،کامپیوتر وشبکه های کامپیوتری متصل به اینترنت نقش عمده ای در ارتباط و انتقال اطلاعات ایفا می کنند. در این بین افراد سودجو با دسترسی به اطلاعات مهم مراکز خاص با اطلاعات افراد دیگر وبا قصد اعمال نفوذ یا اعمال فشار ویا حتی به هم ریختن نظم سیستم ها ، عمل تجاوز به سیستم های کامپیوتری را در پیش گرفته واقدام به نفوذ به سیستم های دیگر کرده وامنیت آنها را به خطر می اندازند . بنابراین لزوم حفظ امنیت اطلاعاتی وحفظ کارایی در شبکه های کامپیوتری که با دنیای خارج ارتباط دارند کاملا محسوس است (حمیدی و ضیاعی ،1388).

سیستم های تشخیص نفوذ برای بسیاری از سازمانها ، ازدفاتر کوچک تا شرکت های چند ملیتی ، ضروری هستند. برخی از فواید این سیستم ها عبارتند از:

- کارایی بیشتر در تشخیص نفوذ ، در مقایسه با سیستم های دستی
- منبع دانش کاملی از حملات
- توانایی رسیدگی به حجم زیادی از اطلاعات
- توانایی هشدار نسبتا بلادرنگ که باعث کاهش خسارت می شود
- دادن پاسخ های خودکار ، مانند قطع ارتباط کاربر ، فعال سازی حساب کاربر ، اعمال مجموعه دستر های خودکار و غیره
- توانایی گزارش دهی

3- روش های تشخیص نفوذ

روشهای تشخیص مورداستفاده درسیستم های تشخیص نفوذبه دودسته تقسیم می شوند (Paul Innella and Oba McMillan).

1- روش تشخیص رفتار غیر عادی

2- روش تشخیص سوءاستفاده یا تشخیص مبتنی بر امضا

3-1 روش تشخیص رفتار غیر عادی

برای تشخیص رفتارغیر عادی ، باید رفتارهای عادی را شناسایی کرده و الگوها و قواعد خاصی برای آنها پیدا کرد . رفتار هایی که از این الگو ها پیروی می کنند ، عادی بوده و رویداد هایی که انحرافی بیش از حد معمول آماری از این الگو دارند، به عنوان رفتار غیر عادی تشخیص داده می شوند. نفوذهای غیر عادی برای تشخیص بسیار سخت هستند، چون هیچ گونه الگوی ثابتی برای نظارت وجود ندارد (Paul Innella and Oba McMillan).

یکی از حالات غیر عادی بودن استفاده از حد معمول از یک سیستم است. مثلا اگر شخصی که یک یا دوبار در روز وارد سیستم می شده امروز چندین برابر گذشته وارد سیستم شده است این فعالیت غیر عادی است . یا مثلا زمان استفاده از یک سیستم نیز می تواند عاملی برای تشخیص مشکوک بودن فعالیت باشد، مثلا اگر شخصی خارج از ساعت اداری وارد سیستم شود نیز یک فعالیت مشکوک انجام داده که می تواند برای تشخیص نفوذ آن وارد عمل شد.

تکنیک ها ومعیار هایی که در تشخیص رفتار غیر عادی به کار می روند عبارتند از(حمیدی و ضیاعی ، 1388):

3-1-1تشخیص سطح آستانه

تعداد ورود وخروج به سیستم ویا زمان استفاده از سیستم ، از مشخصه های رفتار سیستم ویا استفاده کننده است که می توان با شمارش آن به رفتار غیر عادی سیستم پی برد وآن را ناشی از یک نفوذ دانست، این سطح کاملا ایستا و اکتشافی است.

3-1-2 معیار های آماری

در نوع پارامتریک، مشخصات جمع شده بر اساس یک الگوی خاص در نظر گرفته می شود ودر حالت غیر پارامتریک بر اساس مقادیری که به تجربه حاصل شده است مقایسه صورت می گیرد ، از ids های معروف که از اندازه گیری آماری برای تشخیص نفوذ رفتار غیر عادی استفاده می کنند می توان NIDS را نام برد.

3-1-3 معیار های قانون گرا

شبیه به معیار های آماری غیر پارامتریک است . به طوری که داده ی مشاهده شده بر اساس الگوهای استفاده شده ی مشخصی به طور قابل قبول تعریف می شود. اما با الگوهایی که به عنوان قانون مشخص شده فرق دارد وبه صورت شمارشی نیست.

3-2 تشخیص مبتنی بر امضاء

این روش به این ترتیب است که روش های مختلف نفوذی که از قبل استفاده شده و مقابله با آن ها تجربه شده است و به صورت الگوهایی در سیستم قرار داده شده است . سیستم نیز فعالیت های انجام شده را با این الگوها مطا بقت می دهد و طبیعی است که درصورت مطابقت یک فعالیت بایکی از این الگوها باید هشدار لازم را بدهد . در این روش ها ، معمولا تشخیص دهنده دارای پایگاه داده ای از امضاها یا الگوهای حمله است وسعی می کند با بررسی ترافیک شبکه الگوهای مشابه با آن چه را که در پایگاه داده ی خود نگهداری می کند بیابد (Paul Innella and Oba McMillan).
طبیعی است که چنین سیستم هایی توان تشخیص نفوذهایی که باروش های جدید انجام شده والگوی آنها در سیستم موجود نیست را ندارند. در واقع این وظیفه ی مدیر سیستم است که با تحقیق والبته تجربه الگوهای جدید وبه روز نفوذدر سیستم تشخیص نفوذ قرار دهد. در مقابل این سیستم ها روش های نفوذی که شناخته شده بوده والگوی آنها در سیستم موجود است بسیار کارا هستند.

4- داده کاوی

داده کاوی (Data Mining) فرآیندی است که برای استخراج وتحلیل الگوها و اطلاعات مفید از انباره های داده با استفاده روش های نیمه خودکار(Semi-Automatic) وخودکار انجام می شود (بهروزیان نژاد ،1392).

5-داده کاوی در سیستم های تشخیص نفوذ

تکنیک های داده کاوی می توانند در موارد زیر در سیستم های تشخیص نفوذ استفاده شوند:

1. حذف رفتارهای عادی از داده های مشکوک به منظور تمرکز بر حملات واقعی
2. تشخیص مولد های داده هایی که به اشتباه مشکوک به خطر در نظر گرفته شده اند.
3. یافتن رفتارهای غیر عادی که حملات واقعی را آشکار می کنند.

برای به کار گیری تکنیک های داده کاوی فوق در سیستم های تشخیص نفوذ ، متخصصان داده کاوی روش های زیر را به کار می گیرند (jake ryan et a1,1997):

الف) خلاصه سازی(Summarization) داده ها و یافتن مقادیر خارج از محدوده به کمک روشهای آماری .

ب) بصری سازی (Visualization) و یا نمایش گرافی خلاصه داده ها.

ج) خوشه بندی داده ها.

چ)کشف قوانین وابستگی : تعریف رفتارهای عادی به منظور توانایی کشف رفتارهای غیر عادی.

ح)کلاس بندی : پیش بینی طبقه ای که یک رکورد بدان متعلق است.

6- مدل‌ها و الگوریتم‌های داده‌کاوی

بسیاری از محصولات تجاری داده‌کاوی از مجموعه این الگوریتم ها استفاده می‌کنند و معمولا هر کدام آنها در یک بخش خاص قدرت دارند و برای استفاده از یکی از آنها باید بررسی های لازم در جهت انتخاب متناسب‌ترین محصول توسط گروه متخصص در نظر گرفته شود. نکته مهم دیگر این است که در بین این الگوریتم ها و مدل ها ، بهترین وجود ندارد و با توجه به داده ها و کارایی مورد نظر باید مدل انتخاب گردد.

6-1شبکه‌های عصبی

هر شبکه عصبی شامل یک لایه ورودی می‌باشد که هر گره در این لایه معادل یکی از متغیرهای پیش‌بینی می‌باشد. گره‌های موجود در لایه میانی به تعدادی گره در لایه نهان وصل می‌شوند. هر گره ورودی به همه گره‌های لایه نهان وصل می‌شود. گره-های موجود در لایه نهان می‌توانند به گره‌های یک لایه نهان دیگر وصل شوند یا می‌توانند به لایه خروجی وصل شوند. لایه خروجی شامل یک یا چند متغیر خروجی می باشد. سیستم تشخیص نفوذ براساس شبکه های عصبی برای سیستم کامپیوتری ویژه شامل مراحل سه گانه زیر است(Jake Ryan et a1, 1997):

6-1-1 **مجموعه داده های آموزشی :** بدست آوردن دوره های زمانی چند روزه برای هرکاربر از طریق یک بردارنشان می-دهیم که یک کاربر چه دستوراتی را اجرا می کند.

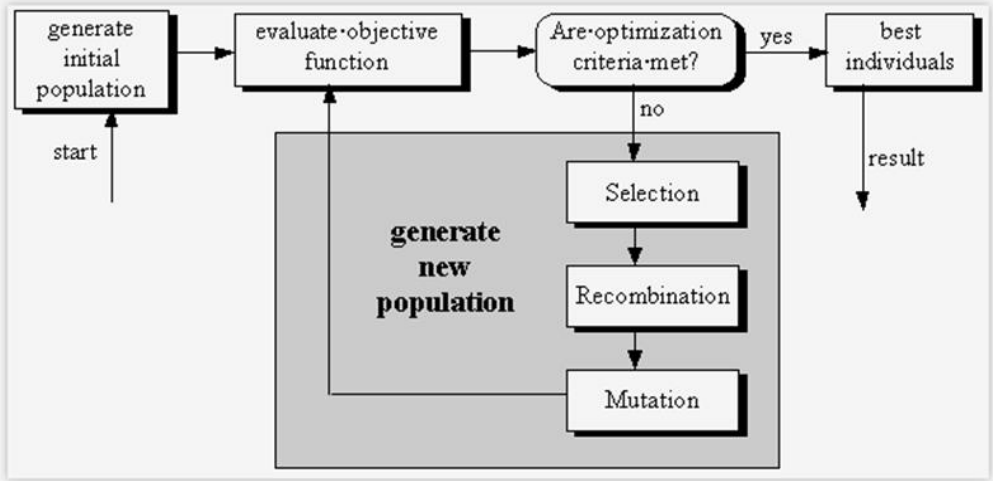
6-1-2 **آموزش:** شبکه عصبی را برای شناسایی کاربر بر اساس دستورراتی که در بردار می باشد.

6-1-3 **کارایی:** شبکه ، کاربررا برای هر دستور جدید شناسایی می کند ، یعنی اینکه اگر کاربری دستور جدیدی که در بردار مربوط به وی وجود ندارد را اجرا کند سیستم قادر به شناسایی آن کاربر خواد بود.

7 - الگوریتم ژنتیک

در این الگوریتم 2 فاز کلی وجود دارد. در فاز اول ما آموزش هایی را به سیستم می دهیم واطلاعاتی را در آن قرار می دهیم تا بتوان با الگو قرار دادن این آموزش ها و داده ها تشخیص نفوذ را انجام دهد. در فاز دوم به کمک این اطلاعات تشخیص نفوذ انجام می شود، در سیستم های تشخیص نفوذی که ازالگوریتم ژنتیک برای آموزش استفاده می نمایند، یک سری قوانین اولیه دسته بندی شده در پایگاه داده قرار می دهیم وبا بکارگیری الگوریتم ژنتیک قوانین جدیدی تولید شده و به قواعد قبل اضافه می شوند (شهبازی، لطیف شبگاهی،91).

در شکل زیر ساختار یک الگوریتم ژنتیک ساده نشان داده شده است.



توضیح تصویر

الگوریتم های ژنتیک ، از تکامل ژنتیکی به عنوان یک الگوی حاصل مسئله استفاده می کند . راه حل ها طبق یک الگو کد گذاری می شوندکه تابع هرراه حل کاندیدرا ارزیابی می کندکه اکثرآن نام داردوبرازندگی می شوند(شهبازی، لطیف شبگاهی،91).

تکامل از یک مجموعه کاملا تصادفی از جامعه اولیه شروع می شود ودر نسل های بعدی تکرار می شود، تا این که به آخرین مرحلهبرسیم(شهبازی، لطیف شبگاهی،91).

شرایط خاتمه الگوریتم ژنتیک می تواند به صورت زیر باشد:

- به تعداد ثابتی از نسل ها برسیم.
- بودجه اختصاص داده شده تمام شود.
- بیشترین درجه بر ارزش فرزندان حاصل شود یا دیگر نتایج.
- بازرسی دستی
- ترکیب های موارد ذکر شده در بالا

8 -ماشین های بردار پشتیبان

ماشین های بردار پشتیبان خصیصه های ورودی با مقادیر حقیقی را با نگاشت غیر خطی به فضایی با ابعاد بالاتر می برد وبا قرار دادن یک مرز خطی ، داده ها را جدا می کند.پیدا کردن یک مرز تفکیک برای جداسازی داده ها به مسئله بهینه سازی درجه دوم تبدیل می شود واز مرز خطی برای تقسیم بندی استفاده می شود . اما همه مسائل از ویژگی به نام توابع پایه SVMبه صورت خطی قابل تفکیک نیستند و برای حل این مشکل استفاده می کند. این توابع الـگوریتـم های خطی را به غیر تبدیل می کند وبا بردن داده ها به فضایی با ابعاد بالاتر،تفکیک خطی رادرآن فضا ممکن می سازند(شهبازی، لطیف شبگاهی،91).

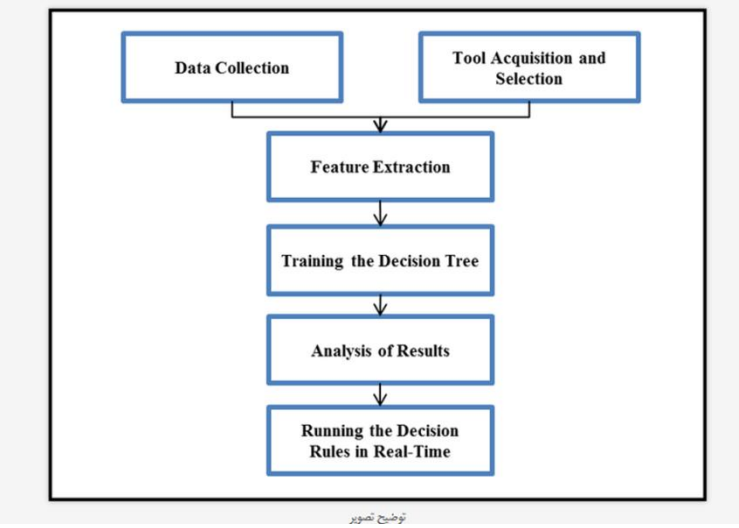
9- درخت تصمیم

درخت های تصمیم روشی برای نمایش یک سری از قوانین هستند که منتهی به یک رده یا مقدار می شوند.یکی از تفاوت ها بین متدهای ساخت درخت تصمیم این است که این فاصله چگونه اندازه گیری می شود. درخت های تصمیمی که برای پیش بینی متغیرهای دسته ای استفاده می شوند درخت های classification نامیده می شوند، زیرا نمونه ها رادر دسته ها و یا رده - ها قرار می دهند. درخت های تصمیمی که برای پیش بینی متغیرهای پیوسته استفاده می شوند درخت های regression نامیده می شود(J.Han and M.Kamber,2001).

9-1 پیاده سازی درخت تصمیم برای تشخیص نفوذ

نخستین کاری که باید انجام شود این است که داده ها وابزارهای موجود که دریافت می کنیم باید پیش پردازش شوند. این پیش پردازش باید داده ها را به فرمی در آورد که درخت تصمیم بتواند از آنها استفاده کند. در واقع داده های خام و پردازش نشده نمی تواند به عنوان ورودی درخت تصمیم مورد استفاده قرار گیرد. نتیجه پردازش داده ها برای مرحله بعد که تعیین مجموعه قوانین است بسیار مهم است.در مرحله بعد آنالیز روی این داده ها انجام شده واز نتیجه ای آن برای تعیین قوانین تصمیم گیری استفاده می شود.

شکل زیر مراحل کار درخت تصمیم را نشان می دهد(Jeff Markey,2011).



توضیح تصویر

نتیجه گیری:

امروزه با توجه به رشد فناوری اطلاعات و شبکه های اینترنتی مسئله امنیت و نفوذ روز به روز گسترش می یابد. سیستم‌های تشخیص نفوذ سخت‌افزار یا نرم‌افزاری است که کار نظارت بر شبکه کامپیوتری را در مورد فعالیت‌های مخرب و یا نقص سیاست-های مدیریتی و امنیتی را انجام می‌دهد و گزارش‌های حاصله را به بخش مدیریت شبکه ارائه می‌دهد. سیستم‌های تشخیص نفوذ وظیفه شناسایی و تشخیص هر گونه استفاده غیر مجاز به سیستم، سوء استفاده و یا آسیب رسانی توسط هر دودسته کاربران داخلی و خارجی را بر عهده دارند. هدف این سیستم‌ها جلوگیری از حمله نیست و تنها کشف و احتمالا شناسایی حملات و تشخیص اشکالات امنیتی در سیستم و اعلام آن به مدیر سیستم است. عموما سیستم‌های تشخیص نفوذ درکنار دیوارهای آتش و بصورت مکمل امنیتی برای آن‌ها مورد استفاده قرار می‌گیرد. سیستم های تشخیص نفوذ سنتی نمی توانند خود را با حملات جدید تطبیق دهند از این رو امروزه سیستم‌های تشخیص نفوذ مبتنی بر داده‌کاوی مطرح گردیده‌اند.که با با استفاده از تکنیک های داده کاوی میتوان به طرز چشم گیری از حمله ونفوذ پیشگیری کرد. در این مقاله سعی شد که با معرفی برخی از الگوریتم های مهم داده کاوی در تشخیص نفوذ پرداخته شود تا به توان با استفاده از تکنیک های داده کاوی از حمله و نفوذ پیشگیری کرد.

منابع:

لاله حمیدی ، سیده مارال ضیایی ، معرفی سیستم های تشخیص نفوذ، آزمایشگاه تخصصی آپا در حوزه امنیت سرویس های شبکه وتجهیزات بی سیم ، تیر 88

ابراهیم بهروزیان نژاد ، " بررسی چند نمونه از سیستم های تشخیص نفوذ مبتنی برشبکه عصبی "، اسفند 1395

وحید شهبازی ، غلامرضا لطیف شبگاهی ، " مرور ودسته بندی روش های تشخیص نفوذ در شبکه های کامپوتری" آذر 91.

kabiri, Peyman, and Ali A. Ghorbani ."Research on Intrusion Detection Response : A Survey ." IJ Network Security 1.2(2005):84-102.

Paul Innella and Oba McMillan, "An Introduction Detection Systems"2001.

Jake Ryan, Meng-Jang Lin, RistoMiikkulainen, " Intrusion Detection with Networks ", Texas University, 1997.

J.Han, and M.Kamber, "Data Mining: Concepts and Techniques", San Diego Academic Press, 2001.

Jeff Markey, "Using Decision Tree Analysis for Intrusion Detection: How- To Guide ", June 5 2011.