امنیت و روش های یادگیری ماشین جهت تشخیص نفوذ شبکه ها

چکیده

امروزه توسعه روزافزون شبکه های رایانه ای و کاربرد وسیع آن در زندگی بشر، لزوم تأمین امنیت این شبکه ها را بیش از پیش نمایان ساخته است. جهت تأمین امنیت از ابزار و تجهیزات مختلفی استفاده می شود که سیستم تشخیص نفوذ از جمله آنها به شمار می رود. سیستم های تشخیص نفوذ ، اغلب از دو روش تشخیص سوء استفاده و تشخیص ناهنجاری به منظور تشخیص نفوذ استفاده می کنند . معماری های امروزی استفاده شده برای سیستم های تشخیص نفوذ، طراحان را در انتخاب نوع معماری کارایی که بتواند قابلیت اطمینان بیشتری در مورد تشخیص حملات داشته باشد با دشواری هایی مواجه کرده است و آن ها مجبور به استفاده از طرح های پیچیده ای برای بالا بردن توانایی این سیستم ها برای تشخیص تهاجم ها و مصون ماندن از حملات صورت گرفته بر علیه آن ها شده اند. همچنین در دنیای امنیت امروزی، بر خلاف گذشته، ابزار های دفاعی مبتنی بر پایگاه داده که در آنها قوانینی برای شناسایی حملات تعریف شده است، کارایی لازم را ندارند و در تامین امنیت شبکه ها به مشکل بر خورده اند. از این رو ابزار دفاعی مبتنی بر الگوریتم های یادگیری ماشین که توانایی مقابله با پیچیده ترین نوع حمله ها را دارند، مورد توجه قرار گرفته اند. از این رو در این مقاله به بررسی و ارزیابی روش های تشخیص نفوذ با استفاده از الگوریتم های یادگیری ماشین از جمله شبکه عصبی،مبتنی بر الگو، ماشین بردار پشتیبان و مدل مخفی مارکوف و .... خواهیم پرداخت تا معیارهای انتخاب یک سیستم تشخیص نفوذ کارا مورد ارزیابی قرار می گیرد.

کلمات کلیدی

امنیت، تهاجم، سیستم تشخیص نفوذ، شبکه های عصبی چند پخشی،مبتنی برالگو، ماشین بردار پشتیبان، مدل مخفی مارکوف

۱-مقدمه

با افزایش سرعت، اندازه و تعداد کامپیوتر ها بعد از سال ۱۹۷۰،اهمیت امنیت سیستم های کامپیوتری افزایش یافت و با توجه به افزایش حجم داده های ذخیره شده در دنباله های ممیزی مرور و تحلیل دستی آنها مشکل گردید، جیمز اندرسون اولین کسی بود که مسالة مرور خودکار دنباله های ممیزی را مطرح ساخت. تجربیات حاصله از ممیزی سیستم ها و ردیابی برخی از وقایع از روی دنباله های ممیزی به همراه ایده های مطرح شده توسط اندرسون، منجر به ظهور سیستم های تشخیص تهاجم در سال ۱۹۸۰گردید۱[۹,۱]. سیستم تشخیص نفوذ يك نرم افزار با قابلیت تشخیص، آشکارسازی و پاسخ (واکنش)به فعالیت های غیرمجاز یا ناهنجار در رابطه با سیستم می باشد. به عبارت دیگر سیستم تشخیص نفوذ يك سیستم محافظتی است که خرابکاری های در حال وقوع در شبکه را شناسایی می کند. در این سیستم ها با استفاده اطلاعاتی مانند پویش پورت ها و تشخیص ترافیك غیر متعارف، نفوذ خرابکاری ها را

می توان کشف و به مسئول شبکه گزارش داد[۲]. سیستم های تشخیص نفوذ ، اغلب از دو روش تشخیص سوء استفاده و تشخیص ناهنجاری به منظور تشخیص نفوذ استفاده می کنند . روشهای تشخیص سوء استفاده، نرخ تشخیص بالایی دارند اما توانایی شناسایی حملات جدید را ندارند . در مقابل ، روشهای تشخیص ناهنجاری، توانایی شناسایی حملات جدید را دارند اما نرخ هشدار غلط بالایی دارند .برای آنکه سیستم های تشخیص نفوذ قدرت کشف و تشخیص نفوذهای از قبل تعریف نشده را داشته باشند، به نوعی هوشمندی نیاز دارند. در این حالت، این سیستم ها قابلیت یادگیری دارند و می توانند بر روی بسته های وارد شده به شبکه تحلیل انجام داده و کاربران عادی و غیر عادی را تشخیص دهند.از جمله روش های هوشمند متداول که امروزه مورد استفاده قرار می گیرند: شبکه های عصبی؛ منطق فازی؛ تکنیك های داده کاوی و الگوریتم ژنتیك می باشند[۱].

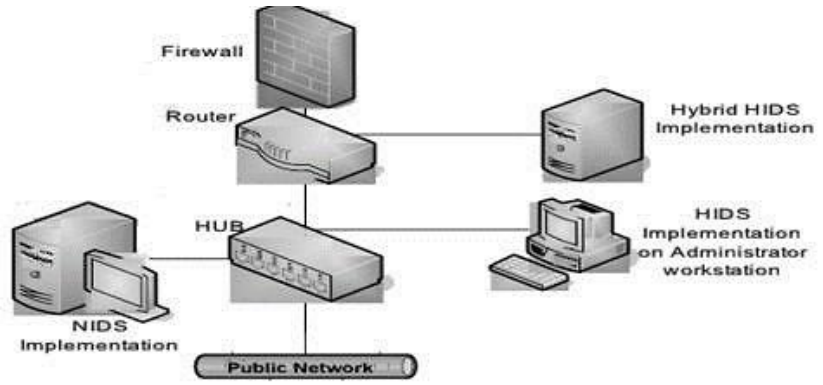
۲- پیشینه سیستم های تشخیص نفوذ

امنیت در شبکه های کامپیوتری یکی از چالش های اساسی دنیای امروز است در طی سال های مختلف لایه های مختلفی برای جلوگیری از ورود حملات ایجاد شده که در این بین سهم سیستم های تشخیص نفوذ چشمگیر بوده است، به طوری که تا قبل از سال ۱۹۸۰ روال کار سیستم های تشخیص نفوذ به صورت دستی انحام می شد بگونه ای که یك فرآیند به صورت دوره ای و زمانبندی شده وقایع رخ داده شده در سیستم را ثبت می کرد که به این فرآیند ثبت ممیزی گفته می شد، زمانی که یك حمله از جانب مهاجمان انجام می شد مدیر سیستم می توانست با مراجعه به ممیزی سیستم کارهایی از قبیل: ترمیم داده های آسیب دیده، کشف محل نفوذ ،بازسازی وقایع سیستم، زمان و تاریخ رخداد و همچنین نوع رویداد یا حادثه را انجام بدهد[۱]. ممیزی سیستم ها با اهداف مختلفی ممکن است صورت پذیرد، از جمله کنترل فعالیت های افراد دارای شناسه در سیستم، تشخیص خرابی ها،تعیین و تشخیص مشکلات سیستم و نواحی رخداد آنها و تشخیص استفاده های ناصحیح یا غیر مجاز از سیستم. یکی از معایب ممیزی این بود که عملیات تولید و ثبت رویدادهای سیستم زمان بر بود و در بیشتر مواقع نمی توانست اطلاعاتی دقیق از تشخیص وقوع حمله به مدیر سیستم بدهد[٤,٧]. اما سیر تکامل سیستم های تشخیص نفوذ با افزایش سرعت، اندازه و تعداد کامپیوتر ها بعد از سال ۱۹۷۰، اهمیت امنیت سیستم های کامپیوتری افزایش یافت و با توجه به افزایش حجم داده های ذخیره شده ، بحث خودکارسازی دنباله ممیزی از سال ۱۹۸۰ توسط اندرسون شروع شد، که با خودکار سازی ثبت رویدادها سرعت عملیات تشخیص را تسریع کرد ،که به دنبال آن چندین سیستم تشخیص نفوذ مثل IDES ,MIDAS, Haystack مطرح شدند[۳].

۲-۱ سیستم تشخیص نفوذ از لحاظ منبع داده

با توجه به گسترش شبکه های کامپیوتری منابع اطلاعاتی نیز گسترش یافتند و زمان تحلیل آن ها طولانی تر و پیچیده تر شده است، سیستم های تشخیص نفوذ برای اینکه قادر به تشخیص یك حمله باشند به یك مجموعه داده برای مطابقت حملات با آن نیاز دارند یا به طور کلی هر سیستم مبتنی بر پردازش داده برای انجام عملیات خود به منابع اطلاعاتی برای ردگیری اهداف خود نیاز دارد ، بنابراین یکی از نیازهای اولیه سیستم های تشخیص نفوذ بدست آوردن منابع اطلاعاتی است که متناسب با نوع سیستم تشخیص نفوذ

(تشخیص حمله یا جلوگیری از حمله)ممکن است متفاوت باشد[١٤]. در شکل زیر محل استقرار هر یك از روش ها قابل مشاهده است ، سیستم های تشخیص نفوذ از لحاظ تامین کننده منابع اطلاعاتی به سه دسته تقسیم می شوند:



تقسیم بندی روش های تشخیص نفوذ از لحاظ منبع اطلاعات

۲-۱-۱ سیستم های تشخیص نفوذ مبتنی برمیزبان

از زمان مطرح شدن سیستم های تشخیص نفوذ تا سال ۱۹۹۰ سیستم های تشخیص تهاجم به طور عمده، مبتنی بر میزبان بودند که تحلیل آنها منحصراً بر اساس داده های حاصل از دنباله های ممیزی سیستم عامل یا منابع اطلاعاتی میزبان گرای دیگر بود. سیستم های تشخیصی مبتنی بر میزبان وظیفه نظارت بر اتفاقات رخ داده شده بروی میزبان را برعهده دارند. که این نظارت با جستجو کردن در فایل های سیستمی انجام می شود و با این کار هر گونه تغییرات ایجاد شده در آن ها توسط نفوذگران را پیدا می کند[۱۲]. منظور از نظارت هر مکانیزم جمع آوری اطلاعات است که توسط سیستم تشخیص نفوذ بکار می رود .سیستم های تشخیص نفوذی که از منابع اطلاعاتی مبتنی بر میزبان بهره می گیرند، روی اطلاعات جمع آوری شده از داخل یك سیستم (میزبان) عمل می نمایند.

۲-۱-۲ سیستم های تشخیص نفوذ مبتنی برشبکه

روال کار سیستم های تشخیصی مبتنی بر شبکه با کارکردن مستقیم بروی ترافیك شبکه انجام می شود ، با توجه به اینکه یکی از منابع مهم اطلاعاتی این سیستم ها ترافیك شبکه است بنابراین قادر به استفاده از بسته های ترافیك موجود بروی یك بخش بعنوان داده های خود می باشد. که این کار با قرار دادن کارت واسط شبکه در مد تصادفی صورت می پذیرد ، مد تصادفی حالتی است که در آن کارت شبکه بگونه ای تنظیم گردد که برای کل ترافیك شبکه وقفه تولید نماید[۱]. یکی از مهم ترین مزیت های این روش نسبت به روش قبلی این است که هزینه نظارت بر منابع اطلاعاتی خیلی کم است چرا که نظارت می تواند به سادگی با خواندن بسته هایی که در بخشی از شبکه جابجا می شوند، صورت پذیرد، بدون آنکه این نظارت تأثیری بر کارایی دیگر سیستم های موجود در شبکه داشته باشد، به همین دلیل اضافه کردن یك سیستم تشخیص

تهاجم مبتنی بر شبکه، به یك شبکه موجود با کمترین تغییر در آن شبکه ممکن می باشد. و همچنین یکی از مهمترین معایب این روش این است که سیستم های تشخیص تهاجم مبتنی بر شبکه با وجود سادگی و قدرتشان، معمولاً نمی توانند در شبکه های مدرن سوئیچی کار کنند، چرا که سوئیچ ها، اتصالات بین میزبان ها را از هم مجزا می سازند به طوری که یك میزبان فقط بتواند ترافیکی که به مقصد آن آدرس دهی شده را ببیند.

۲-۱-۳ سیستم های تشخیص نفوذ ترکیبی

سیستم تشخیص نفوذ ترکیبی که گاهی توزیع شده گفته می شود از ترکیب دو روش قبلی بهره می برد بگونه ای که سطح امنیت و انعطاف پذیری بیشتری را بدنبال خواهد داشت، که این توزیع شدگی نه تنها در جمع آوری اطلاعات بلکه در تحلیل داده ها نیز استفاده می شود .از خصوصیات مهم این روش امکان ردیابی کاربر در شبکه تحت نظارت است، که با توجه به دشوار بودن ردیابی کاربران و فایل های رد و بدل شده بین آنها در یك شبکه بزرگ ،قابلیت مهمی به شمار می رود، بدین ترتیب تشخیص و جلوگیری از حمله های چند نفره و توزیع شده، می تواند به صورت کارا توسط این سیستم صورت پذیرد. برای بهبود کارایی این روش دو تکنیك وجود دارد: الف) تقسیم ترافیک: این روش بیشتر براساس جریان های داده ای و سیاسیت های امنیتی و ساختار IDS کار می کند. ب) متعادل کردن بار: در هر زمان مقدار بار مناسبی برای هر یك کدام از حسگرها در نظر می گیرد، به نحوی که از ظرفیت سیستم به بطور بهینه استفاده شود[۱]. این روش بدلیل ترکیب دو روش قبلی پیچیدگی زیادی را دارا می باشد. البته کارایی و دقت این روش در تشخیص حمله از دو روش قبلی بالاتر است[۲].

۲-۲ سیستم تشخیص نفوذ از لحاظ استراتژی تحلیل

در تحلیل اطلاعات نفوذگر ، مسألة اصلی آن است که چه واقعه ای در سیستم در حال انجام است و چه فعالیت و یا واقعه ای باید مورد توجه قرار گیرد. قاعدتاً در سیستم های تشخیص تهاجم، وقایعی که نشانی از حمله و یا تهاجم به سیستم مورد حفاظت را دارند، باید در تحلیل مورد توجه قرار گیرند[۱]. به این ترتیب در این قسمت سعی داریم دو روش تحلیل و تشخیص تهاجم، یعنی روش های تشخیص سوء استفاده و تشخیص ناهنجاری معرفی نماییم. باید توجه داشت که هر دو روش فوق ، سعی بر این دارند که به گونه ای وقایعی خاص و مشکوک را که نشانی از تهاجم و حمله را در خود دارا می باشند ،مشخص نمایند:

۲-۲-۱ تشخیص نفوذ مبتنی بر سوءاستفاده

در سیستم های تشخیصِ تهاجمی که از روش سوء استفاده بهره می برند، در همان ابتدای امر تعریف دقیقی از تهاجم های ممکن، به سیستم ارائه می گردد و پس از آن سیستم تشخیص تهاجم با نگریستن به وقایع رخ داده در سیستمِ موردِ حفاظت، در صورت مشاهدة یک امضاء و یا الگوی تهاجمی مطابق با الگوها یا امضاهای از قبل تعریف شده ،تهاجم را تشخیص و تولید اخطار می نماید. الگوهای تهاجم شامل ویژگی ها، شرایط، ترتیب و روابط بین فعالیت هایی هستند که منجر به نفوذ و یا سوء استفاده می گردند[۱]. پس بنابراین باید یك پایگاه دانشی از این الگوها یا امضاهای تهاجم در سیستم ایجاد گردد و بر اساس آن به تحلیل وقایع و تشخیص سوء استفاده پرداخته شود. یکی از مزایای عمده این روش این است که حملات

شناخته شده(حملاتی که در بانك اطلاعاتی ثبت شده است) به طور مطمئن و کارا تشخیص داده می شوند، البته یك عیب بزرگ دارند آن این است که قادر به تشخیص حملات ناشناخته و جدید نیستند[۱۳].

۲-۲-۲ تشخیص نفوذ مبتنی بر ناهنجاری

در یک سیستم، رفتارهای کاربران و یا برنامه هایی که از سیستم استفاده می نمایند را می توان به دو دسته رفتارهای هنجار و رفتارهای ناهنجار تقسیم نمود. آنچه که در سیستم تشخیص سوء استفاده مورد توجه قرار گرفته است، بخشی از مجموعه رفتارهای ناهنجار می باشد که همان حمله ها و تهاجم های شناخته شده است .بدین ترتیب به علت عدم شناخت حمله های جدید که بخش دیگر از این مجموعه رفتارهای ناهنجار را تشکیل می دهند، بنابراین تشخیص حمله های جدید و ناشناخته توسط سیستم های تشخیص سوء استفاده امکان پذیر نخواهد بود.

برای رفع نقص فوق در سیستم های تشخیص سوء استفاده، سیستم های تشخیص ناهنجاری با این ایده مطرح شدند که تمام فعالیت های تهاجمی لزوماً فعالیت هایی ناهنجار در سیستم هستند، پس اگر مجموعه رفتارهای هنجار را در یك سیستم بشناسیم، می توانیم هرگونه تخطی از آن را به عنوان رفتاری ناهنجار و احتمالاً رفتاری تهاجمی تشخیص دهیم[۱۱]. بنابراین در این سیستم ابتدا نمایه هایی از رفتارهای نرمال (سیستم، شبکه و یا کاربران آنها) تولید می شود و سپس رفتارهای اتفاق افتاده و واقعی با رفتارهای مورد انتظار و پیش بینی شده در نمایه های نرمال مقایسه می شود، که در صورت انحراف از این نمایه های نرمال، به عنوان یك رفتار ناهنجار و تهاجمی تشخیص داده می شود ،هرچند ممکن است این سیستم ها بعضی رفتارهای نرمال را به عنوان حمله شناسایی نمایند، اما قادر به تشخیص حملات ناشناخته هستند. یکی از مزیت های عمده این روش قابلیت تشخیص حملات جدید و ناشناخته و همچنین گونه های تغییر یافته حمله های شناخته شده می باشد ،بدون آنکه سیستم از قبل دانشی راجع به جزئیات آنها داشته باشد[۱۳,۱۰]. البته این سیستم معایبی نیز دارد که می توان به هزینه ثبت و بروز رسانی رفتارهای نرمال که دارای حجم قابل توجهی هستند و همچنین تعداد زیاد هشدار های مثبت و منفی نادرست که ناشی از رفتارهای غیر قابل پیش بینی کاربران یا شبکه می باشد اشاره کرد.

۲-۲-۳ تشخیص نفوذ مبتنی بر نمایه

این روش براساس نیازهای کاربر، نمایه ها را برای پروتکل های خاص گسترش می دهد. البته هنوز به صورت همه جانبه مورد استفاده قرار نگرفته است و کاربردش بیشتر در IPv6 است، معمولا مدل هایی که در این روش برای پروتکل های شبکه استفاده می شوند براساس استانداردهای بین المللی مانند استانداردهای موسسه IETF تعریف می شوند[۹]. با استفاده از این روش می توان ناهنجاری های بالقوه ای از جمله تغییرات در طول دستورات ،مقادیر حداقل و حداکثر برای صفات و توالی غیر منتظره ای از دستورات که توسط سیستم های تشخیصی مبتنی بر امضا و مبتنی بر ناهنجاری قابل تشخیص نیستند را به راحتی ردیابی کرد. اما یکی از چالش های عمده این روش ترافیك های مخربی است که ممکن است با استفاده از این روش به درستی تشخیص داده نشوند[۱۸]. در جدول زیر به طور خلاصه مزایا و معایب هر کدام از روش ها بیان شده است:

مکانیزم مورد استفاده	مزیت	عیب
مبتنی بر دانش(امضا)	۱- موثرترین روش برای شناسایی حملات از پیش شناخته شده است.	۱-ناتوانی در تشخیص حملات نا شناخته ۲-بروز نگه داشتن امضا ها و الگو ها مشکل است
مبتنی بر رفتار(ناهنجاری)	۱-توانایی تشخیص حملات ناشناخته ۲-وابستگی خیلی کم به سیستم عامل	۱-از دقت تشخیص پایینی برخوردار است چون وقایع همواره در حال تغییراند و ثابت نیستند. ۲-مقرون به صرفه نیست
مبتنی بر نمایه (پروتکل)	۱-شناسایی حالت پروتکل ها ۲-توانایی تشخیص توالی دستورات غیر منتظره	۱-ناسازگاری با برخی سیستم عامل های اختصاصی ۲-نیاز به منابع قابل توجه برای شناسایی حالت پروتکل ها ۳-مشکل تشخیص ترافیک های مخرب

جدول مزایا و معایب رو ش های تشخیص نفوذ از لحاظ استراتژِی تحلیل

۳- یادگیری ماشین

در حالت کلی و عمومی یادگیری ماشین یعنی اینکه چگونه می توان برنامه ای نوشت که از طریق تجربه یادگیری کرده و عملکرد خود را بهتر کند. یکی از تعاریف یادگیری ماشینی آنطور که از سوی تام میشل پروفسور دانشگاه کارنگی ملون ارائه گردید بدین شرح است: نوعی برنامه کامپیوتری که با توجه به برخی وظایف گروه T و عملکردP ، تجربه E را شکل می دهد، اگر عملکرد آن در گروه وظایف T آنطور که توسطP اندازه گیری شده با تجربه E بهبود پیدا کند. یادگیری ماشین به عنوان یکی از شاخه های وسیع و پرکاربرد هوش مصنوعی، به تنظیم و اکتشاف شیوه ها و الگوریتم های می پردازد که بر اساس آنها رایانه ها و سامانه ها توانایی تعلم و یادگیری پیدا می کنند. هدف یادگیری ماشین این است که کامپیوتر (در کلی ترین مفهوم آن) بتواند به تدریج و با افزایش داده ها کارایی بهتری در انجام وظیفه مورد نظر پیدا کند.گستره این وظیفه می تواند از تشخیص خودکار چهره با دیدن چند نمونه از چهره مورد نظر تا فراگیری شیوه گام برداری روبات های دوپا با دریافت سیگنال پاداش و تنبیه باشد[۱۰].

۳-۱ انواع روش های یادگیری ماشین

روش های یادگیری ماشینی براساس قواعد و منابعی که برای بهبود عملکرد خود استفاده می کنند به سه دسته مختلف تقسیم می شوند[٦] :

۳-۱-۱ یادگیری نظارت شده

در این نوع یادگیری سیستم باید قبل از شروع به فعالیت تعلیم داده شود به عبارتی سیستم برای ادامه کار خود نیاز به یك منبع دانش دارد. در این روش هم ورودی و هم خروجی برای سیستم مورد نظر مشخص است یعنی به سیستم باید فهمانده شود که به ازای چه ورودی، کدام خروجی را فراخوانی کند. از مزایای

این روش سرعت بالا در جرای فرآیندها می باشد و از معایب این روش زمان یادگیری است چراکه سیستم باید برای اجرای هر کاری ابتدا آموزش های لازم را دریابد سپس فعالیت های خود ار آغاز کند.

۳-۱-۲ یادگیری نیمه نظارتی

در این نوع یادگیری تقریبا مشابه روش قبلی است اما به طور مستقیم ورودی و خروجی را در اختیار سیستم قرار نمی دهیم بلکه به صورت غیرمستقیم آن ها را در قالب سیگنال های صحیح و خطا یا اصطلاحا تنبیه و پاداش به سیستم می دهیم. به عبارتی یادگیری نیمه نظارتی که به یادگیری تقویتی نیز مشهور است مدلی برای مسائلی از این قبیل فراهم می آورد. در یادگیری تقویتی، سیستم تلاش می کند تا تقابلات خود با یک محیط پویا را از طریق آزمون و خطا بهینه نماید. یادگیری تقویتی مسئله ای است که یك عامل که می بایست رفتار خود را از طریق تعاملات آزمون و خطا با یك محیط پویا فرا گیرد، با آن مواجه است. در یادگیری تقویتی هیچ نوع زوج ورودی- خروجی ارائه نمی شود. به جای آن، پس از اتخاذ یک عمل، حالت بعدی و پاداش بلافصل به عامل ارائه می شود. هدف اولیه برنامه ریزی عامل ها با استفاده از تنبیه و تشویق است بدون آنکه ذکری از چگونگی انجام وظیفه آن ها شود. یك مثال خوب برای این نوع یادگیری انجام بازی است. اگر ماشین برنده بازی شود، سپس از نتیجه کار برای تقویت حرکات آتی خود در حین بازی بهره می گیرد.

۳-۱-۳ یادگیری بدون نظارت

روش های یادگیری ماشین از نوع نظارت نشده خیلی پرکاربرد هستند و نسبت به دو نوع قبلی رواج بیشتری دارند. روش کار این روش به این صورت است که هیچ ورودی و خروجی به سیستم داده نمی شود یعنی سیستم به اصطلاح به صورت خودمختار عمل می کند و هیچ ناظری برای آن ها وجود ندارد ،برخلاف دو نوع قبلی که تعیین خروجی به نوعی وابسته به وردی بود، اما در اینجا هیچ آموزش و تعلیمی در اختیار سیستم قرار داده نمی شود. بلکه سیستم خود به مرور زمان یاد می گیرد و خود را با شرایط وفق می دهد، به عبارتی این نوع یادگیری زمانی رخ می دهد که ماشین با استفاده از داده هایی آموزش می بینید که هیچگونه برچسب گذاری روی آنها انجام نشده. در این روش ،هرگز به الگوریتم یادگیری گفته نمی شود که داده ها نمایانگر چه هستند نکته کلیدی در مورد یادگیری نظارت نشده آن است که پس از پردازش اطلاعات بدون برچسب، تنها کافی است که یك نمونه از داده های برچسب گذاری شده در اختیار الگوریتم یادگیری قرار داده شود تا کارایی کامل پیدا کند. از مزایای این روش این است که هیچ نیازی به آموزش ابتدایی سیستم نیست که همین باعث محبوبیت این نوع سیستم یادگیری شده است، از معایب این روش این است که سیستم یادگیری باید فرایندهای بدون بچسب زیادی را اجرا کند تا بتواند از آن ها الگوبرداری کند و با اجرای تعداد کمی فرآیند نمی تواند نتیجه خاصی از آن بدست آورد.

٤- روش های یادگیری ماشین جهت تشخیص نفوذ

امروزه سیستم های تشخیص نفوذ به منظور حفاظت از امنیت سیستم های اطلاعاتی به کار گرفته می شود. فایروال هایی که برای تشخیص نفوذ استفاده می شود دارای اشکالات خاصی می باشند که توسط روش های مختلف داده کاوی می توان بر آن ها غلبه کرد. تعدادی از تکنیك های یادگیری ماشین مانند

قوانین انجمنی ،سری زمانی و ماشین های بردار پشتیبان و شبکه های عصبی و ...به طور گسترده به منظور ارتقاء تشخیص نفوذ استفاده می شوند. زیرا بکار گیری این روش ها عملیات تشخیص نفوذ را تا حدودی خودکار کرده است و برخلاف گذشته دیگر نیازی به عامل انسانی برای تشخیص فعالیت مشکوک نیست، علاوه براین با افزایش حجم اطلاعات و همچنین افزایش حملات شبکه ای امروزه وجود روشی هوشمند برای تشخیص فعالیت های مشکوک بیش ازپیش لازم و ضروری می باشد.

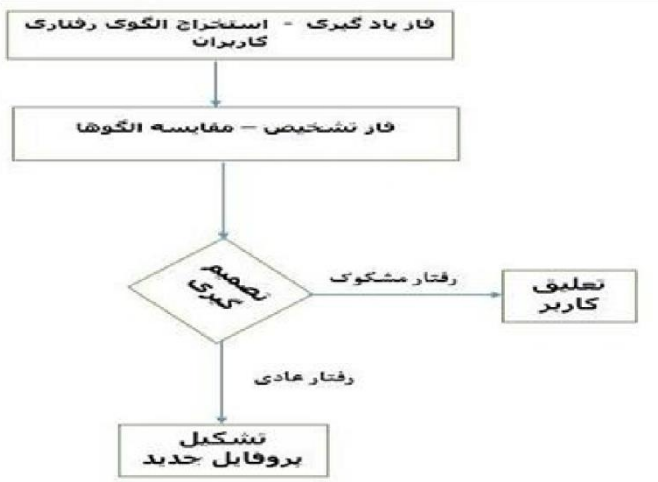
٤-١ تشخیص نفوذ با استفاده از شبکه های عصبی

ایده اصلی شبکه های عصبی (تا حدودی) الهام گرفته از شیوه کارکرد سیستم عصبی زیستی، برای پردازش داده ها، و اطلاعات به منظور یادگیری و ایجاد دانش قرار دارد علت الهام گرفتن از طبیعت می تواند این باشد که طبیعت به خاطر زمانی که در اختیار داشته است تقریبا راه حل بهینه را یافته است در این زمینه آقای اویلر ریاضی دان معروف می گویند " هر چیزی را که در طبیعت بررسی کنی بلاخره یا مینیمم یا ماکسیمم است". ایده استفاده از شبکه های عصبی در تشخیص تهاجم برای اولین بار در سال ۱۹۹۰ توسط Fox [۸] مطرح گردید و در سال های بعد از ۱۹۹۰ کارهای مختلفی در این زمینه انجام گرفت.

از سیستم های تشخیص نفوذ برای کشف حمله و رفتارهای مشکوك استفاده می شود. این سیستم ها در قبال هرگونه حمله یا رفتار مشکوك پیام هشداری را ایجاد می کنند. یکی از مشکلات مطرح سیستم های تشخیص نفوذ زیاد شدن پیام های هشدار و نیز تولید پیام های هشدار غلط است و به عنوان اساسی ترین مشکل سیستم های تشخیص نفوذ به شمار می رود، به گونه ای که در اکثر موارد هشدارها قابل استفاده نیستند. برای حل این مشکل باید با استفاده از مکانیسم های یادگیری ماشین، سیستم های تشخیصی را به نوعی هوشمند کرد یکی از بهترین مکانیسم ها برای برطرف کردن این مشکل استفاده از شبکه های عصبی مصنوعی است.

٤-٢ تشخیص نفوذ با استفاده از داده کاوی الگوی رفتاری

روش ارائه شده در [٥] بر مبنای یافتن ناهنجاری ها در رفتار کاربران محلی می باشد و از تکنیك های استخراج الگوی رفتار ترتیبی و زمانی استفاده می نماید. همانطور که معماری کلی این روش در شکل زیر قابل مشاهده است، روش ارائه شده شامل دو فاز یادگیری و تشخیص می باشد. هدف فاز یادگیری ساختن پروفایلی دقیق از رفتار کاربران بوسیله کاویدن الگوی رفتار آنان در سه سطح تراکنش،دستور و عمل از مستندات ثبت وقایع سیستم می باشد. در فاز تشخیص هر فعالیت جدیدکاربر با پروفایل وی مقایسه شده تا ناهنجاری احتمالی شناسایی و اخطار مناسب تولید شود. در صورت وجود رفتار مشکوك سیستم با توجه به نوع رفتار و میزان سطح دسترسی کاربر می تواند آن را تعلیق کند و اگر هم فعالیت های کاربر عادی تشخیص داده شوند برای ایجاد پروفایل جدید ارسال می شود تا بعد از آن به عنوان یك الگوی شناخته شده راحت تر تشخیص داده شود.

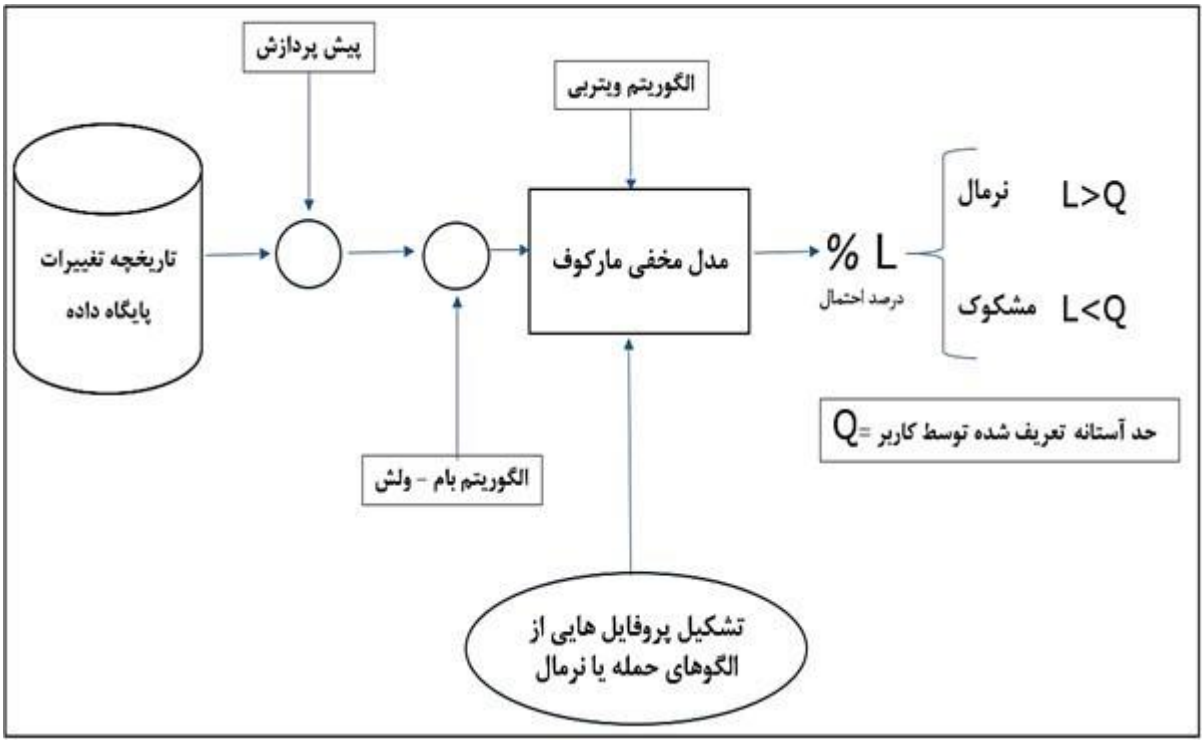


معماری روش تشخیص نفوذ مبتنی بر الگوی رفتاری

۳-٤ تشخیص نفوذ با استفاده از مدل مخفی مارکوف

یك مدل آماری است که در آن سیستم مدل شده به صورت یك فرآیند مارکوف با حالت های مشاهده نشده فرض می شود. به عبارتی یك مدل مخفی مارکوف را می توان یك شبکه بیزی پویا فرض کرد. ایده استفاده از مدل مخفی مارکوف در اواخر ۱۹٦۰ میلادی معرفی گردید، مدل مخفی مارکوف یا HMM هم در همان سال ها معرفی شد، و در حال حاضر گسترش کاربرد آن روبه افزایش است.

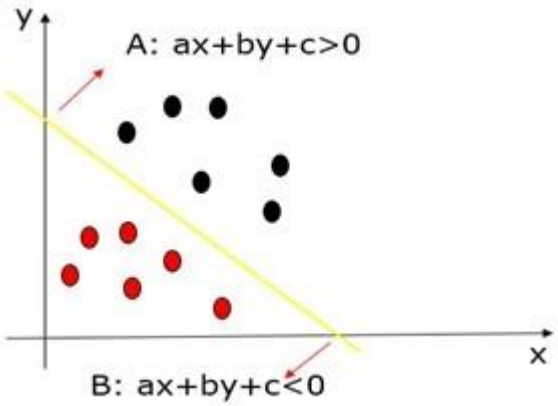
دو دلیل مهم برای این مسئله وجود دارد. اول اینکه این مدل ازلحاظ ساختارریاضی بسیار قدرتمند است و به همین دلیل مبانی نظری بسیاری از کاربردها را شکل داده است. دوم اینکه مدل مخفی مارکوف اگر به صورت مناسبی ایجاد شود می تواند برای کاربردهای بسیاری از جمله تشخیص نفوذ، تشخیص صدا ،پیش بینی وضع هوا مورداستفاده قرار گیرد[۸].



معماری روش تشخیص نفوذ مبتنی مدل مخفی مارکوف

٤-٤ تشخیص نفوذ مبتنی بر ماشین بردار پشتیبان

با توجه به گسترش شبکه های کامپیوتری در سازمانها و مراکز مهم تجاری ، اطلاعاتی، نظامی و... تامین امنیت این شبکه ها ، یکی از مباحث مهم در این سازمانها است. یك سیستم تشخیص نفوذ باید قادر باشد که حمله های جدید را شناسایی کند و همچنین دارای دقت بالایی باشد .منظور از دقت مفهوم عمومی آن یعنی نسبت تعداد داده هایی که درست تشخیص داده شده اند به کل داده های موجود می باشد. یکی دیگر از ویژگیهای یك سیستم تشخیص نفوذ مناسب این است که بتواند به صورت بلادرنگ نفوذ را تشخیص و به مدیر اطلاع دهد. در این بخش از ماشین بردار پشتیبان که مبتنی بر یادگیری ماشین است ،برای تشخیص نفوذ استفاده شده است. ماشین بردار پشتیبان نیز به عنوان نمونه ای از تکنیك های داده کاوی در سیستم های تشخیص نفوذ به شمار می رود. بدین وسیله الگوریتم تشخیص الگو حملات را شناسایی و دسته بندی می کند. هدف استفاده از ماشین بردار پشتیبان، انتخاب بهترین تفکیك کننده خطی است که خطای تعمیم را به حداقل می رساند. اما در تکنیك ماشین بردار پشتیبان به زمان آموزشی نسبتا بالایی نیاز خواهد بود. هر ماشین بردار پشتیبان وظیفه طبقه بندی دو یا چند کلاس از داده های ورودی را بر عهده می گیرد.این کار با استفاده از مقایسه داده های ورودی بر اساس توزیع آنها و با ایجاد یك ابر صفحه برای جداسازی داده ها صورت می گیرد. این تکنیك به عنوان یك جداساز بهینه کارایی خوبی را در طبقه بندی داده ها از خود نشان داده است.



نحوه تفکیک داده های نرمال و مشکوک

۵- انواع سیستم های تشخیص نفوذ

۵-۱- سیستم های تشخیص نفوذ تحت شبکه(NID)

نام NIDS از منظر محلی به موقعیت قرار گیری IDS در شبکه است. این سیستم می تواند بر تمام شبکه نظارت داشته باشد همیشه بهترین محل قرارگیری یک سیستم تشخیص نفوذ تحت شبکه در کنار فایروال شبکه است تا مکمل یکدیگر باشند.

یک سیستم تشخیص نفوذ تحت شبکه در حقیقت یکی از انواع IDSها میباشد که خود را به شبکه متصل کرده و از این طریق ترافیک شبکه را پایش و گزارش های خود را ارائه میکند .یکی از مشکلات NIDS این است که به خاطر جمع آوری اطلاعات و تجزیه و تحلیل با سرعت بالا ممکن است بسیاری از بسته ها را از دست بدهد. بعضی از NIDS ها در سرعت های بالاتر از Mbps511دچار مشکل میشوند.[۱۷]

۵-۲- سیستم های تشخیص نفوذ میزبان(HIDS)

یک سیستم تشخیص نفوذ تحت میزبان به گونه ای طراحی شده است که بتواند بر روی سیستم عامل های میزبان (شخصی) افراد در قالب نرم افزار نصب شده و فعالیت کند.این سیستم ها معمولا دارای یک سرویس می باشند که در پس زمینه فعالیت عادی سیستم عامل انجام میشود در SDIهای مبتنی بر میزبان احتمال هشدارهای نادرست بسیار کم است،چرا که اطلاعات مستقیما به کاربران برنامه های کاربردی بر می گردد.

۵-۳- سیستم های تشخیص نفوذی توزیع شده(DIDS)

این سیستم ها از چندین NIDS یا SSDI یا ترکیبی از این دو نوع همراه با یک ایستگاه مدیریت مرکزی تشکیل شده است. بدین صورت که هر IDS که در شبکه موجود است گزارش های خود را برای ایستگاه مدیریت مرکزی ارسال میکند .ایستگاه مرکزی وظیفه بررسی گزارش های رسیده و آگاه سازی مسئول امنیتی سیستم را بر عهده دارد [۱۸]

٦- انواع روش های تشخیص نفوذ

٦-١- سیستم های تشخیص نفوذ-تشخیص سوء استفاده

در تشخیص سوء استفاده(Misuse Detection)شناسایی نفوذ مطابق داده های مشاهده شده با توصیف های از پیش تعریف شده از رفتار سرزده است.این نوع SDI بیشتر با استفاده از ارزیابی حملات بر اساس شناسه (امضا)های حمله(Attack_Signatures) و اثرات آن(Audit-Trails) در شبکه فعالیت میکند. بنابراین نفوذ شناخته شده می تواند به طور موثر با نرخ خطا False Positive (FP)کم تشخیص داده شوند.به همین دلیل،این روش به طور گسترده ای در اکثر سیستم های تجاری در اختیار گرفته شده است،نفوذها معمولا چند شکلی هستند و به طور مداوم تکامل می یابند.تشخیص مبتنی بر امضا(سوءاستفاده) زمانی که با نفوذ ناشناخته مواجه شود،می تواند با شکست مواجه شود .یک راه برای مقابله با این مشکل این است که به طور منظم،یا به صورت دستی ،یا به صورت خودکار با کمک الگوریتم های یادگیری، در شبکه عصبی به روز رسانی شود[١٩].

٦-٢- سیستم های تشخیص نفوذ-تشخیص رفتارهای غیرمتعارف

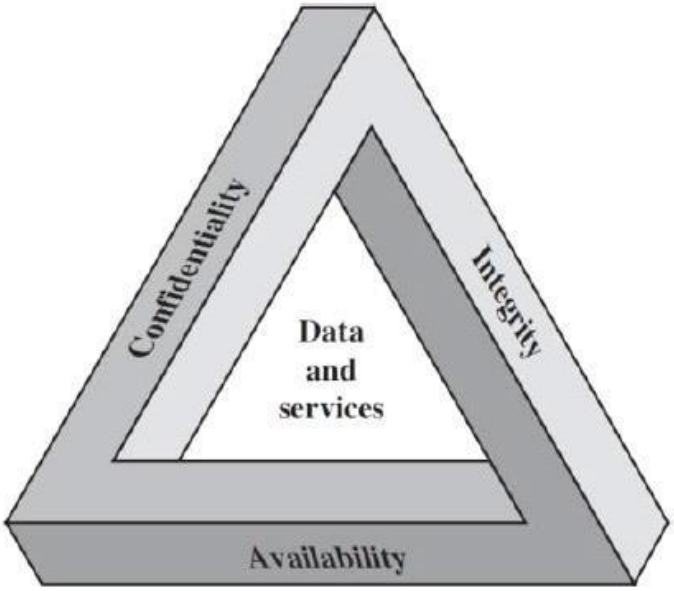
برای تشخیص رفتار غیر متعارف(Anomaly Detection)باید رفتارهای عادی را شناسایی کرده و الگوها و قواعد خاصی برای آن ها پیدا کرد.رفتارهایی که از این الگوها پیروی می کنند،عادی بوده و رویدادهایی که انحرافی بیش از حد معمول آماری از این الگوها دارند،به عنوان رفتار غیرعادی تشخیص داده می شود.نفوذهای غیر عادی برای تشخیص بسیار سخت هستند،چون هیچگونه الگوی ثابتی برای نظارت وجود ندارند.معمولا رویدادی که بیشتر از دو مورد انحراف از رفتار استاندارد داشته باشد،غیرعادی فرض میشود[٢٠].متاسفانه تشخیص دهندگان نفوذ های غیر عادیَ(Anomaly_Based) و IDSهایی از این نوع ،باعث ایجاد تعداد زیادی هشدار نادرست می شوند و آن هم به خاطر این است که الگوهای رفتاری از جانب استفاده کنندگان و سیستم بسیار متفاوت است.برخلاف روش های تشخیص مبتنی بر امضاء روشهای تشخیص رفتار غیرعادی،قادر به کشف انواع حملات جدید هستند.

٧- امنیت شبکه کامپیوتری

حفاظت، پشتیبانی و نگهداری از داده های رایانه ای، اطلاعات مهم، برنامه های حساس و نرم افزارهای مورد نیاز و یا هر چیزی که مورد اهمیت باشد امنیت رایانه ای گفته می شود. و وقتی که در بستر شبکه بزرگی مانند اینترنت، کامپیوترها با هم متصل شده و تبادل داده و اطلاعات انجام می دهند به امنیت این اطلاعات امنیت شبکه کامپیوتری گفته می شود. تمامی کامپیوترها از کامپیوترهای موجود در منازل تا کامپیوترهای موجود در سازمان ها و موسسات بزرگ،در معرض آسیب وتهدیدات امنیتی می باشند.

٧-١- مثلث امنیت شبکه

امنیت صرفا جلو گیری از افشای اطلاعات نیست. بلکه در دسترس نبودن اطلاعات در زمان مورد نیاز و تغییر اطلاعات در مسیر تبادل نیزباعث از دست دادن امنیت می گردد سه عامل رازداری، یکپارچگی و در دسترس بودن مثلث امنیت شبکه را تشکیل می دهند. (روغنی ،١٣٩٢) [٢٠]



مثلث امنیت شبکه

۸- تعریف حمله

راه هایی که خواسته یا ناخواسته از طریق سیستم یا اشخاص سبب از بین رفتن اطلاعات، دسترسی به اطلاعات و یا خسارت زدن به منابع اطلاعاتی باشد را حمله می گوند. (مرتضوی ،۱۳۹٤) [۲۱]

۸-۱- انواع حمله

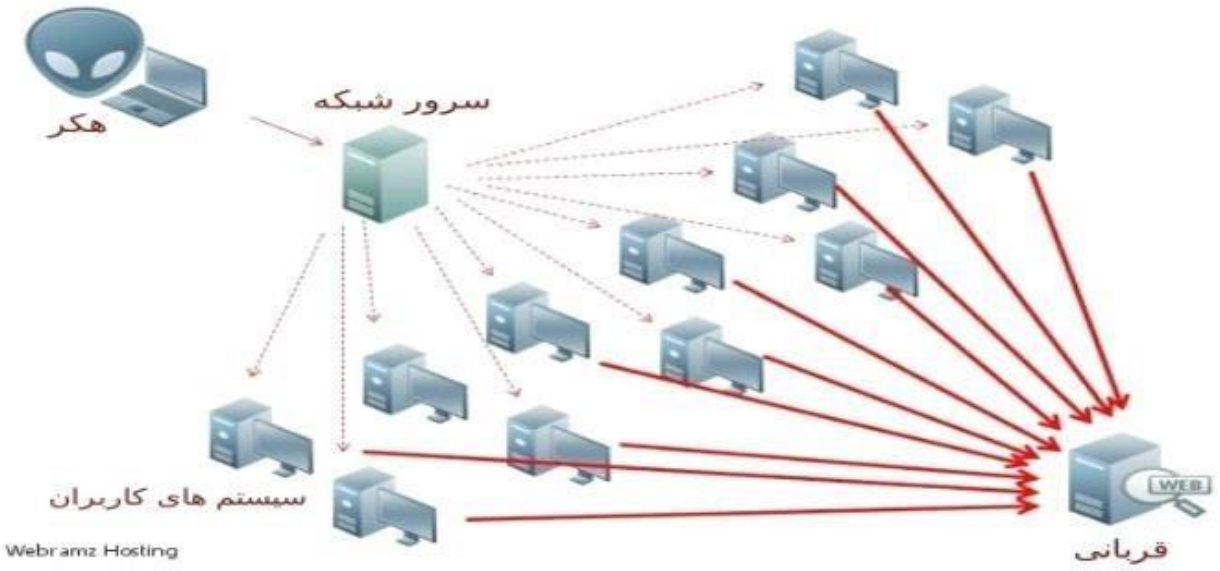
۸-۱-۲- حمله جلوگیری از سرویس

در این نوع حمله کاربرمجاز نمی تواند از منابع در اختیار استفاده نماید. در این نوع حمله به مهاجم اجازه دسترسی و تغییر اطلاعات اصلی معمولا داده نمی شود. ولی از سرویس دهی به کاربران مجاز جلوگیری می کند. هدف مهاجم در این حمله معمولا خرابکاری است که به این حملات Denial-of-Service نیز گفته می شود.

احتمالاً حملات DoS خطرناکترین تهدیدها است که برای توضیح دادن هم مشکل هستند. آنها بدین دلیل خطرناکترین هستند که به آسانی می توانند اجرا شوند، به سختی رهگیری می شوند (برخی مواقع غیرممکن است)، و سرپیچی از درخواست حمله کننده آسان نیست حتی اگر این درخواست غیر قانونی باشد.

منطق یک حمله DoS ساده است. درخواستهای زیادی به ماشین ارسال می شود که از اداره ماشین خارج است. ابزارهای در دسترسی در محافل زیر زمینی وجود دارد که این کار را به صورت یک برنامه در می آورند و به آن می گویند در چه میزبانی درخواستها را منتشر کنند. برنامه حمله کننده به راحتی با برخی از پورتهای

خدماتی ارتباط برقرار می کند، شاید اطلاعات عنوان پاکت را که می گوید بسته از کجا آمده را جعل می کند و آنگاه ارتباط را قطع می کند.



حمله هکر

۸-۱-۳- دسترسی غیرمجاز

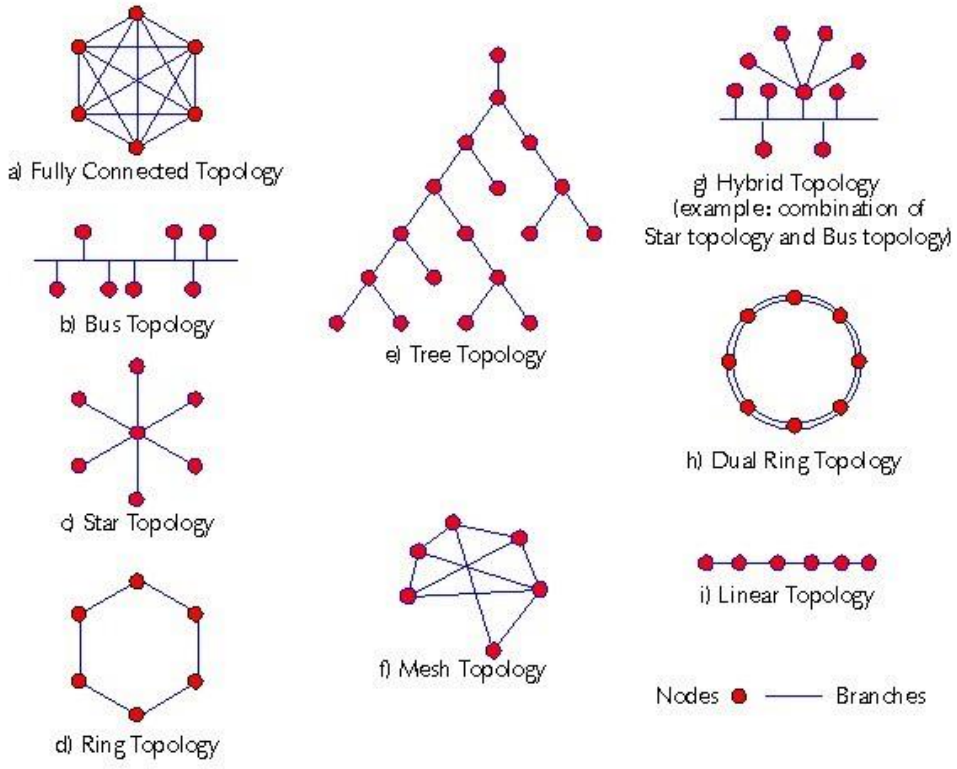
دسترسی غیر مجاز یک واژه سطح بالا است که می تواند به انواع مختلف حملات مرتبط باشد. هدف از این نوع حملات دسترسی به برخی منابع است که ماشین شما نبایستی آنرا در اختیار حمله کنندگان قرار دهد. برای مثال، یک هاست می تواند یک وب سرور باشد و بایستی صفحات وب را برای هر کسی که در خواست میکند در اختیار قرار دهد. با اینحال این هاست نباید دسترسی به لایه فرمان را بدون اینکه مطمئن شود که فرد درخواست کننده مجاز به این کار است، مثل یک مدیر محلی فراهم آورد.

وانواع آن نیز عبارتند از:

جاسوسی, استراق سمع یا شنود, حائل شدن یا واسطه شدن

۹- توپولوژی شبکه

طراحی توپولوژیکی شبکه، یکی از عوامل اصلی است که در زمان رخداد حملات فیزیکی میتواند از خطای کلی شبکه جلوگیری کند.



توپولوژی شبکه

۱۰-نتیجه گیری

با توجه به گسترده شدن شبکههای کامپیوتری و استفاده از اینترنت هر سیستم یا شبکه کامپیوتری آسیب پذیری خود را دارد. با گسترش کاربرد کامپیوتر در جوامع امروزی، مکانیزه کردن اطلاعات و تبادل آن امری اجتناب ناپذیر است .

بنابراین حفظ امنیت سیستم و اطلاعات ضروری میباشد. استفاده از تکنولوژی های برقراری امنیت شبکه تا حدی میتواند این ریسك را کاهش دهد. سیستم های تشخیص نفوذ وظیفه نظارت بر استفاده از سیستم های اطلاعاتی را بر عهده دارند تا هر وضعیت ناامنی را تشخیص دهند و با یك پیغام هشدار آن را به مسئول سیستم تشخیص نفوذ ارسال کند به همین دلیل احتیاج به کنترل مداوم دارد و همچنین تنها افراد متخصص می توانند دادهای ارسالی را تفسیر کنند، تقسیم بندی روش های تشخیص نفوذ این مزیت را دارد که تحلیلگر بتواند با توجه به نوع محیطی در آن قرار دارد سیستم تشخیصی خود را انتخاب کند. حجم فایل های ایجاد شده ناشی از تحلیل بسته های ورودی توسط سیستم سنتی تشخیص نفوذ مثل Snort خیلی بالا است بنابراین امروزه استفاده از روش های داده کاوی برای فرآیند تشخیص فعالیت های مشکوك به شدت مورد توجه قرار گرفته است از طرفی خودکار سازی فرایند تشخیص نفوذ با بکارگیری مکانیسم های یادگیری ماشین می تواند پیشرفت قابل توجهی باشد که با توجه این که روش های سنتی تشخیص نفوذ نیاز به منابع انسان زیادی برای جمع آوری داده های ممیزی دارند. در این گزارش پس از بیان مفاهیم و اصطلاحات مطرح در امنیت کامپیوتر، نحوه ظهور سیستم های تشخیص تهاجم و روند تکاملی این سیستم ها در قالب طبقه بندی های مختلفی مورد بررسی قرار گرفتند. استفاده از روش های

یادگیری ماشینی جهت تشخیص نفوذ می تواند نرخ تشخیص درست قابل توجهی به نسبت روش های مشابه قبلی دارد دلیل این امر بهره گیری روش های یادگیری ماشینی از دانش تجربه و یادگیری از مشاهدات جاری و استفاده آن در مشاهدات آینده که ما در این گزارش روش های مطرح در این زمینه را بررسیکردیم

اینترنت وشبکه های دیگر با وجود تمام مزایا و فناوری هایی را که برای ما به ارمغان می آورند تهدیدات و خطرهایی را نیز برای ما به دنبال دارند که با شناخت این خطرها ومنابعی که باید در مقابل این خطرات محافظت شوند و همچنین با انجام برخی از اقدامات امنیتی می توان چنین خطرهایی را به حداقل رساند.

۱۱-منابع

[1] R.G. Bace, ‘‘Survey Intrusion Detection,’’, Macmillan Technical Publishing, USA, 2000.

[2] H.Debar, M.Dacier,A.Wespi, ”Towards a taxonomy of intrusion-detection systems ,” , IBM Research DiÕision, Elsevier Science 1999.

[3] H.Liao, Ch.Richard Lin, Y. Lin, K.Tung " Intrusion detection system : A comprehensive review,", Department of Computer Science and Information Engineering, Cheng Shiu University & Elsevier Ltd. All rights reserved 2012.

[4] John Vacca, John Mallery, Scott R. Ellis, Bill Mansoor , “Computer and Information Security Handbook”, copyright 2013 by Elsevier inc. all right reserved, 2013.

[5] Ayres, J., Flannick, J., Gehrke, J., and Yiu, T. “Sequential Pattern Mining using A Bitmap Representation” , In ACM SIGKDD Conference, pp.429435, 2002.

[6] A. Smola and S. Vishwanathan ,“INTRODUCTION TO MACHINE LEARNING “,Departments of Statistics and Computer Science Cambridge University Press 2008.

[7] Debar, H., Becker, M., and Siboni, D., "A Neural

Network Component for An Intrusion Detection System",

Proceedings of the IEEE Computer Society Symposium, pp. 240-250, 1992.

[8] S. Kumar and M. Manoria , “ Intrusion Detection using Hidden Markov Model” Ph.D. Scholar, Mewar University, Rajasthan, Director and Professor, International Journal of Computer Applications (0975 – 8887) Volume 115 – No. 4, April .5102

[9] D.E. Denning, “An Intrusion-Detection Model”, IEEE Transaction on software engineering, 1987.

[10] J. Göbel,“Advanced Honeynet Based Intrusion Detection”, Department of Computer Science, Diploma Thesis, July 2006.

[11] B. Zaugg , “An Overview of Intrusion Detection Systems Technology and Research”, Published on 19 June 2010.

[12] H. Debar , M.Dacier , A. Wespi , “Towards a taxonomy of intrusion-detection systems”, IBM Research DiÕision, Zurich Research Laboratory, Saumerstrasse 4,

CH-8803 Ruschlikon, 1999 Elsevier Science B.V. All rights reserved.

[13] Ch. Prajkta , A.Raut, “Hybrid Model For Intrusion Detection System, International Journal Of Engineering And Computer Science ISSN:2319-7242 Volume1 Issue 3 Dec 2012 .

[14] E. Guillen, J. Sánchez and R. Paez,” Inefficiency of
IDS Static Anomaly Detectors in Real-World Networks”,
Engineering Systems Department, Xaverian University, Colombia, Published: 6 May 2015.

[15] Barbara, D., Goel, R., and Jajodia, “S. Mining Malicious Data Corruption with Hidden Markov Models”. In Proceedings of the 16th Annual IFIP WG 11.3 Working Conference on Data and Application Security, Cambridge, England, July 2002.

[16] [15] Raju, E. and Sravanthi K'Network intrusion detection using Support Vector Machines', International Journal of Computer Science and Management Research, 2(1), pp 1314-1319,2013.

[17] W. Wang, X. Guan, X. Zhang, L. Yang, Profiling behavior for anomaly intrusion detection based on the and frequency property of computer audit data Security, vol. 25, no. 7, pp. 539550,

[18] T. Sproull, J. Lockwood, Distributed Intrusion Prevention in Active and Extensible Networks, In: Active Lecture Notes Networks Computer Science, Vol. 3912, pp 54-65, 2007.

[19] D.E.Denning,Anintrusiondetectionmodel, IEEE Transactions On Software Engineering, Special issue on Computer And Security Privacy, vol. 13, no. 2, pp. 222-232, 1987.

[۲۰]. روغنی، محمدتقی ،۱۳۹۲، شبکه صفرتاصد، جلد اول، چاپ اول، تهران

[۲۱]. مرتضوی، سیدعلی، ۲۹۳۱، امنیت شبکه های کامپیوتری، کنفرانس بین المللی علوم و مهندسی ،۷