دانشگاه صنعتی امیرکبیر

پلی تکنیک تهران

دانشکده مهندسی کامپیوتر و فناوری‌اطلاعات

پایان‌نامه کارشناسی‌ارشد مهندسی فناوری اطلاعات

گرایش امنیت اطلاعات

تشخیص مبتنی بر ناهنجاری نفوذ در میزبان از طریق تشابه نمایه کاربر

نگارش

امید مظفری

استاد راهنما

دکتر بابک صادقیان

پائیز ۱۳۸۶

بسمـه تعالـــی

دانشگاه صنعتی امیرکبیر
(پلی تکنیک تهران)
معاونت پژوهشی

فرم اطلاعات پایان نامه
کارشناسی ارشد و دکترا

تاریخ: .....

پیوست: .....

نام و نام خانوادگی:	امید مظفری	**دانشجوی آزاد** ☒	**بورسیه** ☐	**معادل** ☐
شماره دانشجوئی:	۸٤۱۳۱۰۰۳	**دانشکده:** مهندسی کامپیوتر	**رشته تحصیلی:**	فناوری اطلاعات
نام و نام خانوادگی استاد راهنما:	دکتربابک صادقیان			
عنوان پایان نامه به فارسی:	**تشخیص مبتنی بر ناهنجاری نفوذ در میزبان از طریق تشابه نمایه کاربر**			
عنوان پایان نامه به انگلیسی:	**Host - based anomaly detection of intrusions through user profile similarity**			
نوع پروژه:	**کارشناسی ارشد** ☒ / **دکترا** ☐	**کاربردی** ☒	**بنیادی** ☐	**توسعه ای** ☒ **نظری** ☐
تاریخ شروع:	۱۳۸۵/۶/۲۸	**تاریخ خاتمه:** ۱۳۸۶/۷/۳۰	**تعداد واحد:**	۶واحد
سازمان تأمین کننده اعتبار:				
واژه های کلیدی به فارسی:	*همبستگی، رگرسیون، گراف رابطه همبستگی، Log file، نشست*			
واژه های کلیدی به انگلیسی:	**Correlation,regression,graph,log file,session**			

نظرها و پیشنهادها به منظور بهبود فعالیت های پژوهشی دانشگاه:

استاد راهنما:

دانشجو:

امضاء استاد راهنما: **تاریخ:**

نسخه ۱: معاونت پژوهشی

نسخه ۲: کتابخانه و به انضمام دو جلد پایان نامه به منظور تسویه حساب با کتابخانه و مرکز اسناد و مدارک علمی

چکیده:
تشخیص نفوذبخش مهمی از حفاظت امنیت سیستم‌های کامپیوتری است. هر چند محصولات تجاری بسیاری وجود دارند، ولی تشخیص نفوذ با محدودیت‌هایی که روشهای جاری دارند، کاری مشکل است. بنابراین، روشهای پیشرفته موردنیاز می‌باشند. یکی از روش‌های پیشرفته، روش همبستگی داده‌ها است. با استفاده از روش همبستگی یک نمایه از کارهایی که کاربر انجام می‌دهد تشکیل می‌دهیم و آن را بعنوان رفتار هنجار تعریف می‌کنیم، رفتارهای غیر از آن را بعنوان رفتاری ناهنجار تشخیص می‌دهیم.

سیستم‌هایی که تاکنون در زمینه همبستگی داده‌ها ارائه شده‌اند اغلب همبستگی بین اعلام خطرهای تولید شده توسط چند سیستم تشخیص نفوذ را بررسی می کنند ولی در سیستم پیشنهادی ما با استفاده از رابطه ضریب همبستگی پیرسن، همبستگی بین پارامترهای log file کاربر خاص اندازه‌گیری شده و با استفاده از پارامترهای همبسته یک گراف تشکیل می‌شود. سیستم تشخیص نفوذ ما اطلاعات مورد نیاز خود را از log سیستم عامل لینوکس دریافت می‌کند. از روش همبستگی برای تشخیص ناهنجاری استفاده نمودیم که به نرخ تشخیص بالا و نرخ اعلام هشدار پایین برسیم.

هریال این گراف دارای یک وزن است که بیانگر ضریب همبستگی بین دو رأس آن (رأسها همان پارامترهای همبسته هستند) می‌باشد. این روش پیشنهادی همبستگی را بهبود بخشیدیم و برای هر یال این گراف یک رابطه رگرسیون اگر در نشستی رابطه رگرسیون بین دو پارامتر همبسته نقض شود همبستگی بین آنها تغییر نموده است و چنین اتفاقی می‌تواند به معنی تغییر رفتار در یک نشست باشد. که در نتیجه رفتار ناهنجار رخ داده است. با استفاده از ۴ روش آماری و همبستگی رفتار ناهنجار کاربری خاص را تشخیص می دهیم. با توجه به اینکه روش همبستگی، روش اصلی این پایان نامه است ، و بدلیل نرخ تشخیص ناهنجاری و اعلام هشدار بدست‌آمده مناسب نیست، روش پیشنهادی را بهبود بخشیدیم که نتایج آن نشان‌دهنده قابل قبول بودن عملکرد روش بهبود یافته همبستگی برای تشخیص ناهنجاری است.

فصل اول

مقدمه

حمله، نفوذ، از دست رفتن اطلاعات مهم، بدست آوردن رمز عبور ، دسترسی به کامپیوتر دیگران ، هر روز در بسیاری از مجله های کامپیوتر و سایت های اینترنتی با آنها مواجه می شویم. امروزه بیشتر مردم با دنیای کامپیوتر آشنایی نسبی بدست آورده و در منزل و یا محل کار با آن سروکار دارند. مشکلی که بسیاری از کاربران کامپیوتر با آن موجه هستند مشکل امنیتی آن است. حمله هایی که از راه دور به یک سیستم کامپیوتری می شود و یا حتی با ورود به یک سیستم کامپیوتری، از نزدیک آن را مورد نفوذ قرار می دهند. در اخبار ، مجلات و سایت های اینترنتی از ضعف های سیستم های کامپیوتری با اطلاع شده ایم ولی با وجود همه اینها از میزان استفاده و کاربری کامپیوتر کاسته نشده است.

مشکل امنیتی کامپیوتر، قبل از آنکه شبکه های کامپیوتری بوجود بیایند، خیلی محدود بود ولی با ظهور شبکه های کامپیوتری این مشکل نیز بزرگ شد.

۱-۱ امنیت کامپیوتری

هدف امنیت کامپیوتری ایجاد اطمینان در یک سیستم کامپیوتری است که از طریق ۳ویژگی امکان پذیر است. محرمانگی[۱]، انسجام و یکپارچگی[۲]، و موجودیت[۳]، هدف محرمانگی محدود ساختن دسترسی به سیستم‌های کامپیوتری تنها به کاربران مجاز است (می‌تواند یک فرد، فرایند یا سیستم باشد). در این پایان نامه محرمانگی مدنظر است که کاربران مجاز حق وارد شدن به سیستم کامپیوتری و استفاده از آن را دارند، در غیر اینصورت فردی که مجاز به استفاده از سیستم کامپیوتری نیست، وارد سیستم شده و از اطلاعات و منابع سیستم بدون اجازه استفاده می کند، در نتیجه محرمانگی به خطر می افتد. هدف انسجام و یکپارچگی، محدود ساختن احتمال تغییر یا تخریب منابع کامپیوتری است (هم با سوء نیت و هم به طور تصادفی). هدف موجودیت، داشتن سیستم کامپیوتری به ارائه خدمات مناسب به کاربران مجاز، در زمانی است که انتظار می‌رود.

برای اجرای امنیت در یک سیستم کامپیوتری، خدمات امنیتی مختلفی مورد نیاز می‌باشند. در یک مجموعه امنیت پنج سرویس امنیتی عمومی مورد نیاز می‌باشند : تأیید اعتبار[۴]، کنترل دسترسی[۵]، محرمانگی، انسجام، و عدم تکذیب[۶]. هدف تأیید اعتبار، تأیید هویت ارائه شده توسط کاربر است. هدف

---

[1] Confidentiality
[2] Integrity
[3] Availability
[4] Authentication
[5] Access control
[6] Non-repudiation

کنترل دسترسی محدود ساختن دسترسی به اشیاء یا خدمات تنها به کاربران مجاز است. هدف عدم تکذیب، نفوذ ساختن اثبات انتقال و هویت فرستنده است.

اگر یک سیستم کامپیوتری یک یا چند آسیب‌پذیری داشته باشد، ایمن نیست. آسیب‌پذیری (یا نقص امنیتی) اغلب اوقات به صورت ضعفی در یک سیستم کامپیوتری تعریف می‌شود که نقض سیاست امنیتی را امکان‌پذیر می‌سازد.

سیستمی که بتوان به آن اتکا نموده و رفتار آن مطابق انتظار باشد، یک سیستم امن است. اینکه رفتار یک سیستم مطابق انتظار است یا خیر، وابستگی زیادی به تعریفی که ما از سرویسهای مورد انتظار از سیستم داریم، دارد.

۱-۲ آسیب پذیری[1]

آسیب‌پذیری عبارت است از ضعف در رویه‌های امنیتی، مدیریتی و کنترلی- سیستمی که باعث می-شود یکی از سرویسهای امنیتی توسط کاربران غیر مجاز نقض شود[99]. نقاط آسیب‌پذیر یک سیستم مهمترین عوامل ایجاد مشکلات امنیتی آن سیستم می‌باشند که می‌توانند مورد سوء استفاده قرارگیرند. جدول۱-۱ تاکسونومی نقایص امنیتی را با توجه به خاستگاه طبق کار [۷۱] ارائه می‌کند.

<table>
<tr><td rowspan="6">خاستگاه</td><td rowspan="6">تعمدی</td><td rowspan="5">نفوذی</td><td>اسب تروآ</td><td>غیر تکثیری</td></tr>
<tr><td></td><td>تکثیری</td></tr>
<tr><td colspan="2">درب تله[2]</td></tr>
<tr><td colspan="2">بمب ساعتی/ منطقی</td></tr>
<tr><td rowspan="3">غیرنفوذی</td><td rowspan="2">کانال پنهان[3]</td><td>ذخیره</td></tr>
<tr><td>زمانبندی</td></tr>
<tr><td colspan="2">موارد دیگر</td></tr>
<tr><td rowspan="2">ناخواسته</td><td colspan="3">خطای تأیید اعتبار (ناکامل / ناسازگار)</td></tr>
<tr><td colspan="3">خطای حوزه (شامل استفاده مجدد از جسم، باقیمانده‌ها، و</td></tr>
</table>

[1] Vulnerability
[2] Trapdoor
[3] Covert channel

خطاهای نمایش بی‌حفاظ)		
سری‌سازی/ ناهمواری لبه (شامل TOCT-TOU)		
تعیین هویت/ تأیید اعتبار ناکافی		
نقض شرایط مرزی (شامل استفاده بی‌حد از منابع و خطاهای قیود قابل نقض)		
خطاهای منطقی قابل بهره‌برداری دیگر		

جدول۱-۱ تاکسونومی نقایص امنیتی

۱-۳ نفوذ[1]

عملی که باعث شود یک کاربرغیرمجاز از طریق یک آسیب‌پذیری سیستمی سرویسهای امنیتی را نقض کند نفوذ نامیده می‌شود که درواقع منجر به تخطی از خط مشی‌های امنیتی می‌شود[99]. آقای اکسلسون در [۳] هر احتمالی مبنی بر وجود یک کوشش آگاهانه و بدون مجوز را برای دستیابی به اطلاعات، استفاده و تغییر آنها و نیز نامطمئن نمودن سیستم برای کاربران مجاز را یک نفوذ قلمداد می‌کند. برای یک کاربر غیرمجاز انجام فرآیند نفوذ می‌تواند دلایل زیر را داشته باشد:

- دستیابی به منابعی که اجازه استفاده از آنها را ندارد.
- صدمه زدن به یک سیستم و از کار انداختن آن.
- دستیابی به اطلاعات محرمانه و تغییر آن.

۱-۴ تهدید[2]

تهدید در واقع یک نفوذ بالقوه است که به مرحله عمل نرسیده است. یک تهدید قابلیت آسیب رسانی به سرویسهای امنیتی را دارد ولی هنوز بالفعل نشده است[99].

تلاش‌هایی نیز برای ایجاد تاکسونومی‌های تهدیدها، عوامل تهدید (نفوذها یا مقصرها) و تکنیک‌های نفوذت (نفوذ) صورت گرفته است. یک نمونه اولیه تاکسونومی عوامل تهدید است که در سال ۱۹۸۰ توسط [۷۱] ارائه شد و عوامل تهدید طبق جدول ۱-۲ به سه دسته مختلف تقسیم می‌کند:

[1] Intrusion
[2] Threat

نفوذکننده مجاز برای استفاده از منابع داده‌ها/ برنامه	نفوذ کننده غیرمجاز برای استفاده از منابع داده‌های برنامه	
N/A	مورد A: نفوذ خارجی	کاربرد غیرمجاز کامپیوتر توسط نفوذ کننده
مورد Misfeasance: c	مورد B: نفوذ داخلی	کاربرد مجاز کامپیوتر توسط نفوذ کننده

جدول ۱-۲- عوامل تهدید

۱-۵ تشخیص نفوذ

امروزه بحث رسیدگی امنیتی سیستم وتشخیص نفوذ به آن، بخش مهمی از مباحث مربوط به امنیت کامپیوتر و شبکه را تشکیل می‌دهد. فنآوری تشخیص نفوذ با ارائه امکان کشف و ردیابی نفوذ، بطور مستقیم و با نظارت کامل بر صحت سیستم و قابلیت مکانیسم‌های امنیتی دیگر، در خدمت تأمین اهداف امنیتی قرارگرفته‌است.

وقتی نوبت به تعریف تشخیص نفوذ می‌رسد، دو راهکار عمده وجود دارد. اولین راهکار تنها شامل عنصر تشخیص نفوذ است.

مربوط به تکنیک‌هایی که تلاش می‌کند با مشاهده اعمال، گزارشات امنیتی یا داده‌های حسابرسی، نفوذ به یک کامپیوتر یا شبکه را آشکار کنند. تشخیص نفوذها یا تلاش‌ها به صورت دستی یا از طریق سیستم‌های تخصصی نرم‌افزاری که بر گزارشات یا دیگر اطلاعات موجود در شبکه عملی می‌کنند.

راهکار دوم، عنصر پاسخ را به تشخیص نفوذ اضافه می‌کند.

«تشخیص نفوذ فرایند تعیین و پاسخ‌دهی به فعالیت‌های بدخواهانه‌ای است که منابع شبکه و محاسباتی را هدف گرفته‌اند.»

یک سیستم اتوماتیک کامپیوتری که به فرایند تشخیص نفوذ کمک می‌کند، یک سیستم تشخیص نفوذ[1] نامیده می‌شود. به بیان بسیار ساده می‌توان گفت سیستم تشخیص نفوذ یک سیستم اعلام خطر خطر برای کامپیوترها و شبکه‌هاست.

سیستم های تشخیص نفوذ با استفاده از فرآیند نظارت بر وقایع رخ داده در یک سیستم یا شبکه کامپیوتری و تحلیل آنها، بدنبال کشف موارد انحراف از سیاست های امنیتی می‌باشند. این سیستم‌ها شامل سه جزء اصلی می‌باشند:

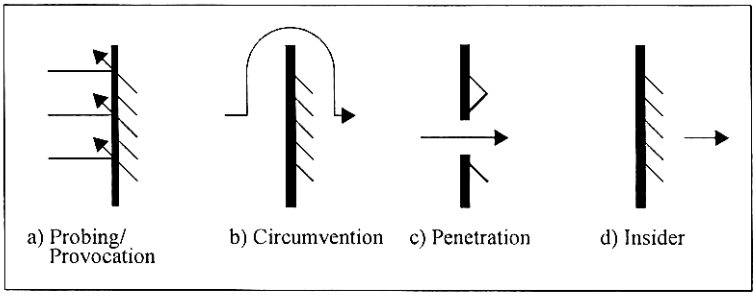
- یک منبع اطلاعاتی که جریانی از وقایع اتفاق افتاده در سیستم را فراهم می-کند.
- یک موتور تحلیل که عموماً "*موتور تشخیص*[2]" نامیده می‌شود و علائم نفوذ را پیدا می‌کند.
- یک پاسخ دهنده که براساس خروجی موتور تشخیص عکس‌العملهایی از خود نشان می‌دهد.

با افزایش سرعت، پیچیدگی و تعداد کامپیوترها نیاز به خودکار کردن سیستم ثبت‌ماوقع[3] به علائم و نشانه‌های نفوذ و همچنین پیوسته نمودن این عملیات بیش از پیش احساس می‌شود و این در حالی است که ثبت‌ماوقع در فاصله‌های زمانی کوتاه و مرور دستی آنها بسیار دشوار است. روال ثبت‌ماوقع شامل فرآیند تولید، ثبت و مرور یک سابقه تاریخی از وقایع سیستم است. ثبت‌ماوقع با اهداف مختلفی همچون نگهداری سابقه‌ای از فعالیتهای سیستم و اطلاعات شخص انجام دهنده آن، بازسازی وقایع سیستم، ارزیابی خسارت، ترمیم مناسب در موقع بروز خطا و کشف استفاده‌های نامناسب از سیستم انجام می‌گیرد.

در شکل ۱-۱ ، به رویدادهایی اشاره شده است، که بایستی سیستم تشخیص نفوذ به آنها پاسخ بدهند.

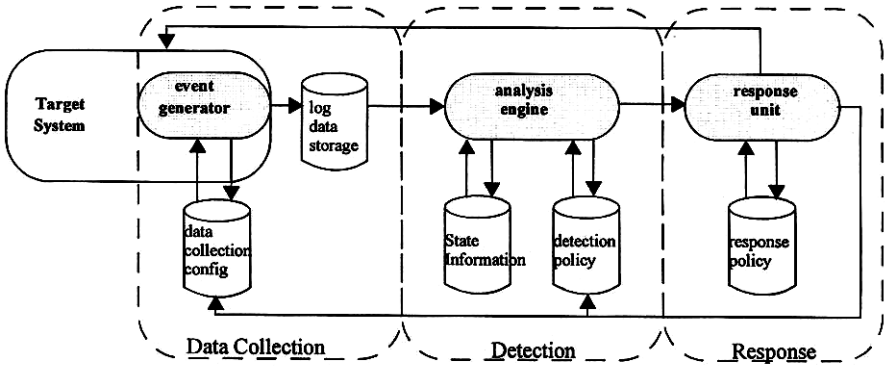
---

[1] Intrusion Detection System
[2] Detection Engine
[3] Auditing



شکل۱-۱ رویدادهایی که باید باعث پاسخ سیستم تشخیص نفوذ بشوند[۷۲]

برای ممکن ساختن عملکردی بودن بین سیستم‌های مختلف تشخیص نفوذ (و بخش‌های آن)، چارچوبی به نام چارچوب مشترک تشخیص نفوذ (CIDF[1]) پیشنهاد شد[۷۲]. در CIDF، یک سیستم تشخیص نفوذ متشکل از چهار بخش اصلی است که تولید اتفاقها[۲]، موتورهای تجزیه و تحلیل[۳]، مکانیزم‌های ذخیره[۴] و واحدهای پاسخ[۵] نامیده می‌شوند. تولیدکننده رویداد، رویدادها را جمع‌آوری کرده و آنها را به موتور تجزیه و تحلیل ارسال می‌کند. موتور تجزیه و تحلیل، رویدادها را براساس شیوه تشخیصش تجزیه و تحلیل می‌کند. مکانیزم‌های ذخیره، اطلاعات مورد نیاز سیستم تشخیص نفوذ را ذخیره می‌کنند. واحد پاسخ برای شروع کردن نوعی پاسخ اتوماتیک دستی یا اتوماتیک مورد استفاده قرار می‌گیرد.



شکل۱-۲ مدل کلی اجزای سیستم تشخیص نفوذ براساس CIDF

[1] Common Intrusion Detection Framework
[2] Event generator
[3] Anlysis engines
[4] Storage Mechanism
[5] Response units

شیوه‌های اصلی تشخیص به طور کلی به تشخیص سوء استفاده[1] و ناهنجاری[2] تقسیم می‌شوند.در [۳] از عبارت تشخیص ناهنجاری استفاده می‌کند، اما تشخیص سوء استفاده را به تشخیص امضاء تغییر داده است. در [۷۳] در عوض عبارات تشخیص مبتنی بر رفتار و آگاهی را اتخاذ کرده‌اند. در هر صورت، تمام این موارد معنای تقریباً یکسانی دارند، یعنی:

- تشخیص ناهنجاری به یک شیوه تشخیص اشاره دارد که رفتار کنونی را با نمایه‌های رفتار گذشته، عمدتاً با استفاده از شیوه‌های آماری، مقایسه می‌کند.

- تشخیص سوء استفاده به یک شیوه تشخیص اشاره دارد که رویدادهای کنونی، مثلاً log entries، را با رویدادهای مشخص بد مقایسه می‌کند که اغلب اوقات با امضاهای نفوذ ذخیره شده در یک پایگاه دانش و آگاهی نشان داده می‌شوند.

علاوه بر این، [۳] شیوه سومی را تعریف می‌کند که الهام یافته از امضاء نامیده می‌شود و هم رفتار نرمال سیستم و هم رفتار مرتکب را به حساب می‌آورد.

امنیت سیستم تشخیص نفوذ بسیار مهم است. این موضوع شامل حفاظت از تمام مراحل در فرایند تشخیص و پاسخ به نفوذ، و نیز حفاظت از سیاست‌های امنیتی توزیع شده است.

کارآیی تشخیص نفوذ به توانایی دسته‌بندی صحیح رویدادهای حسابرسی تحت عنوان رویداد نفوذ و تحمیلی یا رویداد غیر نفوذ اشاره دارد. این موضوع اغلب اوقات شامل یک تصمیم باینری است، اما تصمیم می‌تواند شامل درصدی قطعیت نیز باشد.

از سوی دیگر، راندمان و کارایی به توانایی پردازش مؤثر و کارآمد رویدادهای ورودی اشاره دارد. راندمان کم ممکن است به سرازیر شدن داده‌های حسابرسی به سیستم تشخیص نفوذ بینجامد، که به از دست رفتن رویدادها یا رویدادهای دیرهنگام منجر می‌شود.

اگر به توانایی دسته‌بندی صحیح رویدادی حسابرسی تحت عنوان نفوذ یا غیرنفوذ برگردیم، ممکن است چهار وضعیت زیر روی بدهند.

---

[1] Misuse detection
[2] Anomaly

رویداد غیرنفوذ[۱]	رویداد نفوذ[۲]	
مثبت نادرست[۳]	مثبت درست[۴]	تصمیم مبنی بر نفوذ[۵]
منفی درست[۶]	منفی نادرست[۷]	تصمیم مبنی برغیرنفوذ[۸]

جدول۱-۳ چهاروضعیت رویداد بازرسی[۹]

مثبت نادرست، سیستم به اشتباه کاربر مجاز را غیرمجاز تشخیص بدهد.

منفی نادرست، سیستم به اشتباه کاربر غیرمجاز را مجاز شناخته و اعلام هشدار ندهد.

آنچه در این پایان نامه مورد بررسی قرار می‌گیرد، استفاده از روشهای آماری و همبستگی داده‌ها[۱۰] که از مباحث کلاسیک آماری است در موتور تشخیص می‌باشد. امروزه سیستم‌های تشخیص نفوذ برای افزایش قدرت تشخیص نفوذ خود از منابع اطلاعاتی کامل‌تری نسبت به سیستمهای قدیمی استفاده می‌کنند. این منابع اطلاعاتی که براساس معماری سیستم تشخیص نفوذ و موقعیت حسگرهای آن که وظیفه جمع‌آوری اطلاعات اولیه را برعهده دارند متفاوت خواهد بود دارای یکسری اطلاعات غیر ضروری است که می‌توانند از سوی موتور تشخیص نادیده گرفته شوند. برای انتخاب سودمندترین اطلاعات موجود در منابع اطلاعاتی و حذف داده‌های کم اهمیت راههای مختلفی وجود دارد که یکی از این راهها استفاده از روشهای همبستگی داده‌ها و قرار دادن داده‌های همبسته در گروههای مجزا و درواقع تولید اطلاعات چگال‌تر است تا علاوه بر کاهش حجم داده‌ها، هیچ یک از داده‌های موجود در منابع اطلاعاتی کنار گذاشته نشوند.

ما دراین پایان نامه روشهایی جهت دسته‌بندی این داده‌ها و پیش‌پردازش آنها ارائه می‌نماییم که در نهایت تنها اطلاعات سودمند موجود در منابع اطلاعاتی را در اختیار موتور تشخیص قرار می‌دهد و البته موتور تشخیص نیز براساس همین روابط همبستگی موجود بین داده‌های جمع‌آوری شده فرآیند تشخیص

---

[1] Non-intrusive event
[2] Intrusive event
[3] False Positive
[4] True Positive
[5] Intrusive Deciosion
[6] True Negative
[7] False Negative
[8] Non-intrusive decision
[9] Audit events
[10] Data Correlation

را اجرا می‌کند. موتور تشخیص پیشنهادی ما براساس تشخیص ناهنجاری عمل می‌کند و رویکردی آماری (در ادامه این پایان نامه رویکرد آماری را معرفی خواهیم نمود) دارد.

ساختار فصل‌های پایان نامه

پس از اینکه در این بخش آشنایی مختصری با انواع مفاهیم امنیتی،و سیستم تشخیص نفوذ بدست آوردیم در فصلهای بعدی به ترتیب بصورت زیر عمل خواهیم نمود:

در فصل دوم روشهای ترکیب داده ها را توضیح خواهیم داد. ایده های نو در سیستم های تشخیص نفوذ هستند که درنتیجه کمتر مورد استفاده قرار گرفته اند. پس از مطالعه و بررسی این روشها، روش همبستگی را برای تشخیص نفوذ مبتنی بر میزبان در نظر گرفتیم که روی این موضوع کاری صورت نگرفته است.

در فصل سوم روشهای همبستگی داده را که یکی از روش‌های ترکیب داده‌ها است، به همراه مباحث آماری استفاده شده در سیستم تشخیص نفوذ خود را معرفی می‌کنیم. سپس کارهایی را که در این زمنیه ارائه شده را بیان می‌کنیم. فصل چهارم این پایان نامه به معرفی روشهای بکارگرفته شده در موتور تشخیص سیستم تشخیص نفوذ ارائه شده اختصاص دارد. روشهای استفاده شده در سیستم تشخیص نفوذ پیشنهادی را در فصل چهارم معرفی و توضیح می دهیم با استفاده از روش همبستگی پروفیل کاربر خاص را تشکیل داده و در نتیجه با مقایسه با پروفیل کاربری می توان رفتار ناهنجار را تشخیص داد ، سه روش آماری در نظرگرفته شده و آزمایشات و تستهای تشخیص نفوذ را بروی log های کاربران انجام می دهیم و در آخر نیز روش اصلی این پایان نامه که روش همسبتگی است، پیشنهاد می شود ، نرخ تشیخص و اعلام خطرهای نادرست را از روی نتایج روش پیشنهادی بدست می آوریم. که در فصل پنجم نشان داده می شود. همچنین تستها و آزمایشات ۳ روش آماری نیز در این فصل می آوریم. با توجه به پایین بودن نرخ تشخیص و بالا بودن نرخ اعلام خطر نادرست روش پیشنهادی را بهبود بخشیده و سپس نرخهای گفته شده را محاسبه می کنیم. روش بهبود یافته همراه با آزمایشات و نتایج آن در فصل ششم آورده ایم. در نهایت نیز با استفاده از نتایج بدست آمده از مرحله ارزیابی و تست موتور تشخیص یک جمع بندی نهایی از کل پایان نامه در فصل هفتم ارائه خواهیم نمود.

فصل دوم

ترکیب داده ها برای سیستم تشخیص نفوذ

۲-۱ طبقه بندی سیستم‌های تشخیص نفوذ

تقسیم بندی دیگری که برای سیستم‌های تشخیص نفوذ می‌توان ارائه نمود براساس منبع اطلاعاتی مورد استفاده در این سیستم‌ها است. بطور کلی منابع اطلاعاتی مورد نیاز برای یک سیستم تشخیص نفوذ جهت تصمیم‌گیری در مورد فعالیتهای مشاهده شده می‌توانند در دو دسته زیر قرار گیرند:

- بر اساس شبکه
- بر اساس میزبان

در واقع این دو معماری نحوه جمع‌آوری اطلاعات مورد نیاز برای منابع اطلاعاتی یک سیستم تشخیص نفوذ را مشخص می‌کنند. هریک از این دو معماری دارای خصوصیات مختص به خود می‌باشند و البته هریک مکمل دیگری نیز می‌باشد. این روشها دارای نقاط ضعف و قدرت خاص خود هستندکه در ادامه آنها را بررسی خواهیم نمود[۱۱].

۲-۱-۱ سیستم تشخیص نفوذ بر اساس شبکه

ترافیک شبکه یکی از رایج‌ترین منابع اطلاعات برای سیستم‌های تشخیص‌نفوذ می‌باشد. در این حالت داده‌ها از ترافیک شبکه جمع‌آوری و مورد تحلیل قرار می‌گیرند. اطلاعات بدست آمده از ترافیک شبکه از جنبه‌های مختلف دارای اهمیت می‌باشد. یکی از علت‌ها، نرخ ورود بسته‌ها می‌باشد. در اکثر موارد، نرخ ورود بسته‌ها به اندازه‌ای نیست که دریافت آنها در کارآئی سیستم مشکلی ایجاد کند. یکی دیگر از مزایای استفاده از اطلاعات شبکه این است که دریافت اطلاعات از دید کاربر مخفی می‌باشد. علاوه بر این موارد، با بررسی اطلاعات شبکه می‌توان حملاتی را تشخیص داد که با بررسی اطلاعات سیستم‌عامل و یا برنامه کاربردی قابل تشخیص نبوده است.

خصوصیات چنین روشی را می‌توان در موارد ذیل خلاصه نمود:

- هزینه کم
- دشواری در ازبین بردن مدارک نفوذ توسط نفوذ گران
- تشخیص و عکس العمل مناسب و بلادرنگ به حملات
- توانایی در تشخیص تلاشهایی برای حمله که با موفقیت به پایان نرسیده‌اند
- مستقل از سیستم عامل

۲-۱-۲ سیستم تشخیص نفوذ بر اساس میزبان

در این روش تمامی رخدادهایی که در سیستم ثبت شده‌اند مورد بررسی قرار می‌گیرند. نظارت بر سیستم، رخدادهای سیستمی و رخدادهای امنیتی بهترین منابع اطلاعاتی در این روش می‌باشند. در چنین سیستمهایی، فایلهای سیستمی و اجرایی برای مشاهده تغییرات غیرمنتظره توسط Checksum بررسی می‌شوند. در این روش تنها فعالیتهای نفوذی کشف می‌شوند که اثر آنها در سیستم باقی‌مانده باشد و فعالیتهای مخربی که به حمله منتهی نشده‌اند قابل شناسایی نخواهند بود. در مقایسه با مدل شبکه‌ای می‌توان گفت سرعت پاسخ‌دهی کمتری دارند اگرچه تمام عملیاتهای کاربران را درفاصله اتصال آنها به شبکه ثبت خواهند کرد. این سیستم درصد زیادی از حملاتی را که مدل قبلی تشخیص نداده است، تشخیص خواهد داد.

۲-۲ سیستم‌های ترکیبی

رویکرد ترکیبی از طریق استفاده از انواع مختلف سیستم‌های تشخیص نفوذ و قرار دادن آنها در نقاط بحرانی ورودی شبکه و همچنین در میزبان هایی که توابع اصلی (عملیاتهای اصلی) بر عهده آنهاست امکانپذیر می باشد. با ترکیب تکنیکهای مختلف در یک سیستم ترکیبی در واقع کلیه مزایای ممکن جهت غلبه بر بسیاری از مشکلات را در اختیار خواهیم داشت. سیستم‌های ترکیبی خیلی سریع از نیازمندیهای اساسی محیطهای گسترده سازمانی بشمار آمدند. سیستمهای ترکیبی بشکل وسیعی و از طریق افزایش قدرت کاربران، یک سیستم تشخیص نفوذ امن و مناسب برای سازمانها ارائه نمودند. هنگامیکه یک طراح سیستم امنیتی از یک سیستم تشخیص نفوذ ترکیبی استفاده می‌کند مسلماً به ضعف‌های سیستم‌های تشخیص نفوذ ساده واقف است و می‌داند که آنها نمی‌توانند تمام نیازهایش را برطرف کنند.

مشخصه های یک سیستم ترکیبی خوب در زیر نشان داده شده است:

- این سیستم نباید به آسانی فریب بخورد
- این سیستم باید به اندازه کافی در برابر خطا تحمل پذیر باشد. بگونه ای که یک ازکارافتادگی سیستمی را بدون نیاز به راه اندازی مجدد پشت سر بگذارد
- از آنجا که هر سیستم از الگوهای خاصی استفاده می‌کند با هم تفاوت دارند
- براحتی بتواند با ابزارهای امنیتی و چهارچوب کاری آنها مجتمع شود و قابلیت تطبیق بالایی داشته باشد

این سیستم می‌بایست عملکرد موفق اثبات شده ای را در یک محیط پیچیده داشته باشد

- علاوه بر اثبات تشخیص رفتارهای غیرعادی، حسگرها می‌بایست به تناسب هر شبکه خاص قابل پیکربندی باشند
- صحت و درستی امضاها
- قابلیت تفسیر و نگهداری داده هایی که بتوان آنها را در دادگاه ارائه نمود
- بروزرسانی بموقع و بجای امضاها
- برخورداری از حمایت پرسنل با تجربه
- صحت نصب آن در محیطهای پیچیده اثبات شده باشد
- توانایی نظارت بر عملکرد خودش جهت جلوگیری از تخریب خود را دشته باشد

۲-۳ متمرکزشده

اگر یک موتور تشخیص برروی یک سیستم مرکزی که وظیفه اجرای فرآیندهای تشخیص نفوذ را براساس اطلاعات جمع آوری شده از عاملهای توزیع شده در سطح شبکه دارد قرار دهیم و اطلاعات مورد نیاز آن را از طریق آن عاملها فراهم نمائیم در واقع از یک سیستم تشخیص نفوذ مرکزی استفاده کرده‌ایم. هنگامیکه از لفظ عامل استفاده می‌کنیم ممکن است این سوالها در ذهن تداعی شود که آیا هر عامل خود به تنهایی یک سیستم تشخیص نفوذ است و یا تنها به جمع آوری داده‌های خام مورد نیاز برای یک سیستم تشخیص نفوذ مرکزی کمک میکند؟ و یا اینکه این عاملها بطور ثابت برروی هر ماشین قرار دارند و یا اینکه متحرک هستند؟ و ...

هر عامل تنها وظیفه جمع‌آوری اطلاعات خام حاصل از حسگرهای موجود برروی هر شبکه را برعهده دارد و هیچ پردازشی روی آنها انجام نمی‌دهد. سیستم تشخیص نفوذ مرکزی که برروی یکی از ماشینها قرار دارد، عاملهای خود را به سایر ماشینهای شبکه ارسال می‌کند و این عاملها داده‌های خام را پس از دسته‌بندی در فیلدهای مناسب و در بعضی مواقع دورریختن اطلاعات غیرضروری جمع‌آوری شده توسط حسگرها و بالاخره قراردادن آنها در یک قالب استاندارد از پیش تعیین شده به سمت سیستم مرکزی ارسال می‌نمایند. چنین طرحی نیاز به تعریف قواعدی کلی جهت جمع‌آوری داده‌های مطلوب، فیلتر نمودن اطلاعات جمع‌آوری شده غیر مفید، نحوه دسته‌بندی اطلاعات مرتبط با هم وبالاخره نرمال‌سازی و به شکل یک قالب استاندارد درآوردن داده‌ها دارد که در نهایت ارتباط و وابستگی بین این اطلاعات به آسانی قابل تمیز دادن باشد تا در نهایت نرخ هشدار خطر اشتباه در سیستم کاهش یابد و از سوی دیگر نیز یک هشدار نفوذ درصد قطعیت بالایی پیداکند.

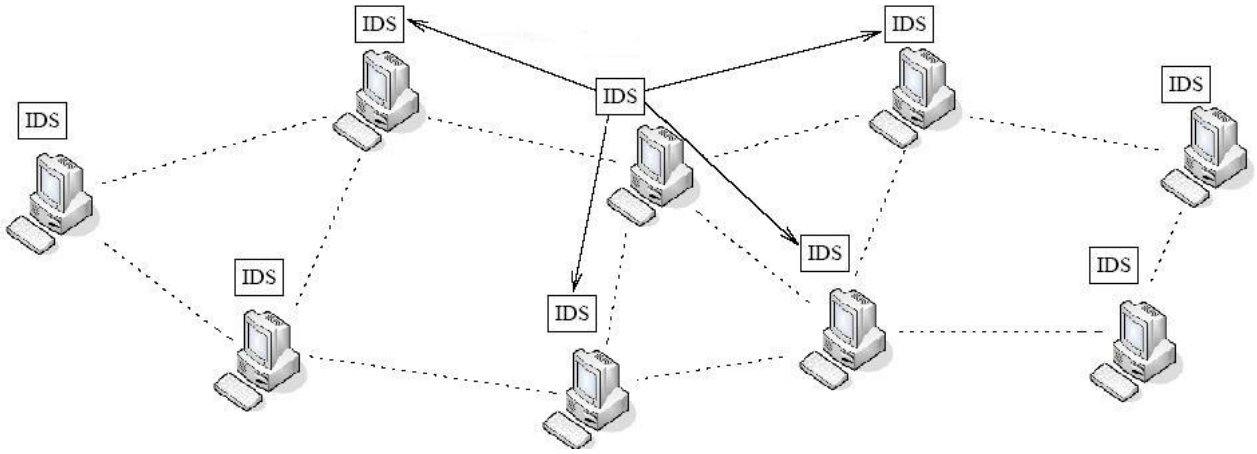
۲-۴ توزیع شده

طرح دیگری که با استفاده از این عاملها می‌توان پیاده‌سازی نمود خود مختار بودن هرعامل برای هشدار نفوذ به سیستمی است که عامل مذکور روی آن قرار دارد. در چنین سیستمی که معمولاًیک طرح مناسب برای سیستمهای تشخیص نفوذ مورد استفاده در شبکه‌های بیسیم است، برروی هر ماشین میزبان روی شبکه یک سیستم تشخیص نفوذ مجزا نصب می‌شود و با استفاده از اطلاعاتی که حسگرهای تعبیه شده روی هر ماشین در اختیار سیستم تشخیص نفوذ خودش قرار می‌دهد آن سیستم را در اجرای فرآیند تشخیص یاری می‌دهد. هر سیستم تشخیص نفوذ علاوه بر رسیدگی به وضعیت امنیتی ماشین مربوطه، با توجه به اطلاعاتی که در اختیار دارد می‌تواند به روند آگاه‌سازی سایر ماشینهای شبکه از تهدیدهای امنیتی که احیاناً با آنها مواجه خواهند شد کمک نماید. این عملکرد دو جانبه توسط سیستمهای تشخیص نفوذ توزیع شده در سطح شبکه را می‌توان به شکل زیر توجیه نمود:

اگر یکی از ماشینها یک فعالیت مخرب محلی را برروی سیستم خودش تشخیص دهد بدون تأثیرپذیری از سایر اطلاعات دریافتی از دیگر ماشینهای شبکه با حمله مورد نظر برخورد خواهد نمود. در چنین حالتی ارسال یک گزارش از آنچه در این ماشین بوقوع پیوسته است برای سایر ماشینهای شبکه می‌تواند به روند تشخیص نفوذ در سایر ماشینها در زمانهای بعدی کمک مؤثری بنماید.

اگر یکی از ماشینها یک فعالیت مخرب را برروی شبکه بگونه‌ای تشخیص دهد که احتمال تأثیر گذاری آن برروی کل شبکه زیاد باشد و امنیت سایر ماشینهای شبکه را تهدید نماید، سایر ماشینها را در جریان این حمله قرارخواهدداد. در پاسخ به چنین هشدار فراگیری سایر ماشینها می‌توانند بعنوان مثال هویت خود را با اجرای یک فرآیند هشدار هویت مجدد سراسری برروی کل شبکه به سایرین ارسال نمایند تا اگر یکی از آنها در دسترس یک نفوذگر قرار گرفته است تشخیص داده‌شود. از مزایای این طرح می‌توان به عدم تمایل یک نفوذگر برای اجرای دروغین فرآیند هشدار مجدد کلی در سطح شبکه برای دستیابی به هویت سایر کاربران اشاره نمود زیرا در اینصورت هویت اصلی خود نفوذگر افشا میشود. پس اجرای چنین فرآیندی نمی‌تواند جعلی باشد و از جهتی دیگر بهترین راه برای تشخیص نفوذ در شبکه خواهد بود.

این طرح با همه مزایایش یک ایراد عمده نیز دارد و آن نیز هزینه بالای حاصل از قرار دادن سیستمهای تشخیص نفوذ مجزا برروی تمام ماشینهای شبکه میباشد[۱۲]. شکل۲-۱ یک معماری توزیع شده از سیستمهای تشخیص نفوذ را به تصویر میکشد.



شکل۲-۱ یک معماری توزیع شده از سیستمهای تشخیص نفوذ

در این بخش چند ایده نو در این سیستمها که امروزه زمینه‌های پژوهشی فراوانی را در مقوله امنیت ایجاد نموده‌اند معرفی می‌کنیم.

در [11] چهار ایده بعنوان ایده‌های نو در زمینه سیستمهای تشخیص نفوذ معرفی شده است:

۲-۵ ایده های نو در سیستم های تشخیص نفوذ

۲-۵-۱ سیستم تشخیص نفوذ بر اساس ایمنی شناسی

یک مدل ایمنی شناسی تشخیص بصورت توزیع شده و نیز توسعه یافته معمولاً تشخیص منفی نامیده می‌شود که کارآیی آنها هم اکنون مورد مطالعه طراحان این سیستم ها قرار گرفته است. چنین رویکردی معمولاً بر اساس تناسب رفتاری / زیستی و مقایسه آن با رفتارهای عادی صورت می‌گیرد. اینکه سیستم دفاعی بدن چگونه در مقابل بیماریها مقابله میکند می‌تواند یک ایده خوب برای مقابله با نفوذگران به دنیا اطلاعات باشد. سیستم تشخیص نفوذ، نفوذگران را از طریق تغییرات مشکوک در رفتارشان مشخص می‌کند. هدف سیستم تشخیص، تمیز دادن بین رفتار قانونی و متعارف با رفتار غیرقانونی است. این سیستم احتیاج به تشکیل یک پایگاه داده از فراخوانیهای سیستمی مشاهده شده برای یک برنامه دارد تا هنگامیکه ترتیب اجرای برنامه یک ترتیب خارج از این پایگاه داده را نشان بدهد. در این حالت می‌گوئیم سیستم یک رفتار نفوذی جدید را تشخیص داده است (تشخیص غیرنرمال) و این اطلاعات جدید را در جهت بهبود الگوهای قبلی (تشخیص قانون بنیاد) بکار می‌برد.

۲-۵-۲ سیستم تشخیص نفوذ مبتنی بر داده‌کاوی

یک رویکرد نمونه از تشخیص نفوذ است، که با تحلیل وقایع ثبت شده در سیستم بدنبال کشف الگوهای غیرنرمال می باشد. مساله ای که در فایل‌های ثبت ماوقع وجود دارد اینست که امکان افزایش هزینه تحلیل سیستم بدلیل نرخ تغییرات داده ها و اثرنامطلوب آن زیاد است. اخیراً بعضی از توسعه دهندگان این سیستم ها علاقهٔ خود را به استفاده از رویکرد جمع آوری اطلاعات از روش داده کاوی نشان‌داده‌اند. این مدلهای تشخیص نفوذ از حملات شناخته شده و رفتارهای طبیعی جهت تشخیص حملات ناشناخته استفاده می‌کنند.

در حال حاضر چندین مورد مبهم و مشکل ساز در پیاده سازی این روش و نیز Application هایی که از این روش استفاده می‌کنند وجود دارد که عمدهٔ آنها عبارتند از:

- صحت تشخیص.
- قابلیت استفاده و تأثیرگذاری.

بطور نمونه سیستمهای تشخیص نفوذی که براین اساس عمل می‌کنند (بویژه سیتمهای تشخیص رفتارهای غیرعادی) نسبت به سیستمهای قدیمی که براساس امضاء عمل می‌کردند نرخ هشدار اشتباهی خطر بیشتری داشتند تشخیص نفوذ با استفاده از داده کاوی هنوز از تکنیکهای خام و ابتدایی این مقوله به شمار می‌رود که موارد زیادی جهت اصلاح و گسترش این سیستم وجود دارد.

سیستمهای تشخیص نفوذ در شبکه به یک جز استاندارد در سازمان امنیتی در آمده اند بطوریکه administrator های شبکه می توانند تخلف های امنیتی را تشخیص دهند. تخلف های امنیتی می تواند در بازه ای از حمله های خارجی که می خواهد دسترسی غیر مجازی از سیستم داشته باشد تا کارمندان داخلی که از دسترسی های خود سوءاستفاده می کنند قرار می گیرد.

متناوباً، administrator می تواند جاهایی از سیستم را که نیاز به بهبود دارند را شناسایی کند، آسیب پذیرهایی شناخته نشده، سیستمی که درست patch نشده است.

متاسفانه، سیستمهای امروزه در تشخیص حمله های جدید ضعیف هستند. برای حل این مشکل، پیشنهاد می شود تکنیکهای داده کاوی را که در اتصلات شبکه در یک محیط offline بکار می رود را استفاده کرد، سنسور های موجود real-time را افزایش داد.

در این قسمت به dataset های مختلف در دسترس برای ارزیابی سیستم ها نگاهی می اندازیم سپس feature های مهمی که برای روشهای داده کاوی انتخاب می شود را مرور می کنیم. درانتها روشهای داده کاوی را توضیح می دهیم.