توسعه یک سیستم تشخیص نفوذ ترکیبی با استفاده از داده کاوی برای سیستم های قدرت

نوشته: Shengyi Pan، Thomas Morris، Uttam Adhikari

چکیده:

سیستم های Synchrophasor حجم بسیار زیادی از اطلاعات برای نظارت گسترده و کنترل سیستم های قدرت برای برآورده کردن تقاضای رو به افزایش انرژی قابل اعتماد، فراهم می کند. ساخت و ساز سیستم های تشخیص نفوذ سنتی (IDSS) که به صورت دستی بر اساس دانش تخصصی ایجاد شده اند، دانش فشرده است و و در زمینه مسائل با داده های بزرگ، مناسب نیستند. در این مقاله یک رویکرد سیستماتیک و خودکار برای ساخت یک IDS ترکیبی که مشخصات موقتی مبتنی بر حالت، برای حالات سیستم قدرت شامل اختلالات، عملیات عادی کنترل، و حملات سایبری را می آموزد، ارائه می کند. یک مسیر داده کاوی به نام کاوش مسیر مشترک استفاده می شود تا به صورت خودکار و صحیح به یادگیری الگوها برای حالات تلفیقی داده های اندازه گیری synchrophasor و لگاریتم حسابرسی سیستم قدرت، بپردازد. به عنوان یک اثبات مفهوم، یک طرح اولیه IDS اجرا شد و تایید شده است. نمونه اولیه IDS به دقت اختلالات، عملیات کنترل عادی، و حملات سایبری برای طرح حفاظت از راه دور برای سیستم انتقال قدرت دو خط سه اتوبوس، طبقه بندی کرده است. **اصطلاحات:** حملات سایبری، داده کاوی، حفاظت از راه دور، سیستم تشخیص نفوذ (IDS)، سیستم قدرت، سیستم synchrophasor.

www.nn4e.com – nn4e@aol.com – 09360703858 – گروه نامیرا

مقدمه

نسل بعدی سیستم قدرت ، که همچنین به عنوان شبکه هوشمند شناخته شده است، بر فن آوری های پیشرفته مانند سیستم های synchrophasor برای نظارت گسترده و کنترل سیستم های قدرت برای برآورده کردن تقاضای رو به افزایش انرژی قابل اعتماد تکیه می کنند. در حالی که در گذشته، اجزای سیستم های برق عایق شده بودند، آنها در حال حاضر از طریق زیرساخت های اطلاعات، به عنوان مثال، اترنت به هم پیوسته، و در نتیجه در معرض خطر حملات سایبری هستند. با توجه به نقش مهمی که سیستم های قدرت در جامعه ما بازی می کند ، یک توافق مشترک وجود دارد که شبکه برق بایستی امنیت بهتریداشته باشد تا اطمینان داشته باشیم که قدرت به طور مستمر در دسترس برای ملت وجود دارد [۱]. اسناد متعدد از سازمان های مختلف موجود است که توصیه ها و دستورالعمل هایی برای تامین امنیت بهتر امکانات صنایع را فراهم کرده است [۲]، [۳]. با این حال، مرکز پاسخگویی دولت ایالات متحده (GAO) نتیجه گرفته است که دستورالعمل های جاری برای به پیاده سازی امن شبکه های هوشمند کافی نیست و GAO خواستار تحقیق و توسعه به منظور بهبود مکانیزم های امنیتی فعلی، شده است [۴].

سیستم های تشخیص نفوذ (IDS) فعالیت هایی که سیاست امنیتی یک سیستم کامپیوتری یا شبکه را نقض می کند، تشخیص می دهند. IDS یک مکمل لازم برای مکانیزم های امنیتی پیشگیرانه مانند firewalls است. چراکه IDS حملاتی

www.nn4e.com – nn4e@aol.com – 09360703858 – گروه نامیرا

را تشخیص می دهد که کاستی ها و باگ های طراحی سیستم را کشف می کنند و IDS شواهد قانونی را فراهم می کند تا واکنش مدیر سیستم به حملات سایبری را اطلاع دهد {۵}. افزایش اتصال زیرساخت های سایبری و دستگاه های فیزیکی شبکه هوشمند باعث می شود یک IDS سنتی مبتنی بر میزبان ناکافی باشد چراکه IDS مبتنی بر میزبان، به صورت جداگانه میزبان را در سیستم مانیتور می کند در حالی که الگوریتم های کنترل سیستم قدرت مانند طرح حفاظت از راه دور معمولا شامل دستگاه های متعدد در مکان های مختلف هستند. بنابراین، IDS جدید باید توانایی در نظر گرفتن منابع اطلاعاتی متعدد و انجام نظارت stateful در سطح سیستم را دارا باشد. ساخت دستی یک IDS، stateful وظیفه ایست که دانش زیاد می طلبد که نیاز به تجزیه و تحلیل آسیب پذیری و ایجاد کتابچه راهنمای قوانین و الگوهایی که حملات و رفتارهای طبیعی را توصیف می کنند. فرایند تهیه کتابچه راهنما، مقیاس پذیری محدود و به روز رسانی کند و گران قیمت را نتیجه می دهد.

این مقاله یک رویکرد سیستماتیک و خودکار برای ساخت یک IDS ترکیبی ارائه می دهد که اهرم وار ویژگی های IDS مبتنی بر علامت و مشخصات را نیرو می بخشد. IDS رفتارهای سیستم را در طول زمان به عنوان اختلالات خاص، عملیات کنترل عادی، و یا حملات سایبری طبقه بندی می کند. سلسله ای از شرایط بحرانی، که مسیرهای مشترک نامیده می شوند، یک علامت یا مشخصه خاص را برای هر سناریو فراهم می کند. یک عنصر اساسی IDS که در این مقاله ارائه شده،

یک تکنیک داده کاوی است که داده های اندازه گیری synchrophasor را جمع کرده و لگاریتم های حاصل از دستگاه های سیستم چندتایی را حسابرسی می کند تا مسیرهای متداول را یاد بگیرد.

روش خودکار نیاز به تجزیه و تحلیل دستی و کدگذاری دستی الگوها را حذف می کند و قادر است مقادیر بسیار زیادی از داده ها را در خود داشته باشد.

مسیرهای متداول، علائم رویدادهای حاضر در یک پایگاه داده آموزشی هستند. همچنین این مسیرها خاص هستند چراکه رفتار مورد انتظار سیستم مربوط را که به رفتارهای نرمال مورد انتظار سیستم و حملات سایبری مربوط است، توصیف می کنند.

IDS با یک مجموعه موقت از وضعیت سیستم نظارت به مسیرهای متداول متصل می شود تا به یک طبقه بندی برسد. رفتارهایی که با یک مسیر متداول مطابقت ندارند، به عنوان وقایع نامعلوم در نظر گرفته می شوند و یا حملات روز صفر و یا رفتار های سیستم ناشناخته می باشند.

مطالعه موردی گنجانده شده است تا نشان دهد که IDS پیشنهادی، دقت تشخیص بالا برای هر دو سناریوی شناخته شده و ناشناخته فراهم می کند و در نتیجه برای ماموریت محیط های حساس مانند سیستم های قدرت مناسب است.

بقیه این مقاله به شرح زیر سازماندهی شده است. بخش دوم بررسی کارهای مرتبط است. یک نمای کلی از بستر آزمون و حالات شبیه سازی شده سیستم قدرت، در بخش سوم ارائه شده است.

www.nn4e.com - گروه نامیرا

بخش چهارم روشی برای ساخت IDS پیشنهاد می کند. آزمایشات و نتایج در بخش V آورده شده است. نتیجه گیری در بخش ششم ارائه شده است.

II. مطالعات مرتبط

A. IDS برای شبکه هوشمند

در سال های اخیر، ظهور شبکه هوشمند، پژوهش را به انواع مختلفی از تکنیک های IDS انگیزه داده است. مردم با پیشینه های مختلف IDS های مختلف را به وجود آورده اند که بر جنبه های مختلف شبکه هوشمند تمرکز کرده اند. یک نوع از این تحقیقات IDS بر امنیت هوشمند دستگاه های الکترونیکی (IED) در شبکه هوشمند تمرکز دارد [۶]، [۷]. این نوع از IDS معمولا مبتنی بر میزبان هستند و در نتیجه تنها حملات بر ضد عملکرد شبکه های مجرد IED در سیستم بر اساس رفتارهای در نظر گرفته شده است. در حالی که IDS مبتنی بر میزبان تامین امنیت دستگاه های منحصر به فرد در شبکه هوشمند، را بر عهده دارد، و نظارت stateful به سطح سیستم را ارائه نمی دهد. IDS پیشرفته تر از این نوع، رفتار دستگاه های متعدد در سیستم را در نظر می گیرد تا سطح تشخیص سیستم را به دست آورد. Mitchell and Chen [۸] یک IDS مبتنی بر قانون را برای شبکه برق با در نظر گرفتن رفتار های سه نوع از دستگاه های فیزیکی در شبکه برق پیشنهاد کردند: ۱) پایان های اصلی ۲) توزیع نقاط دسترسی/ نقاط تجمع داده؛ و ۳) اندازه انرژی مشترک. قرائت از ۲۲ سنسور از سه نوع دستگاه که به عنوان اجزای

www.nn4e.com – nn4e@aol.com – 09360703858 – گروه نامیرا

حالت استفاده شد. این روش هر یک از ۲۲ قطعه را در محدوده به تعداد محدودی با فرمول اندازه گیری کردند.

سه دستگاه حالت با ۳۴۵۶، ۱۷۲۸ و ۳۴۵۶ تعداد حالت برای سه دستگاه به صورت دستی ساخته شد و ماشین های حالت به عنوان خصوصیات خاص برای سه نوع از دستگاه ها عمل می کنند. ساخت و ساز دستی چنین IDS هزینه سنگینی دارد و مقیاسش را برای سیستم های قدرت بزرگتر نمی توان تغییر داد. علاوه بر این، تغییرات رفتار سیستم نیاز به روز رسانی دستگاه های حالت خصوصیات خاص از طریق فرایند کتابچه راهنمای کاربر، دارد.

IDS مبتنی بر شبکه ترافیک ارتباطات در زیرساخت های اطلاعات شبکه هوشمند برای تشخیص حملات سایبری را نیرو می بخشد. IDS می تواند سیستم اعتماد که ارتباطات را کنترل می کند، نیرو بخشیده و از یک دستگاه [۲۴] ارتباطات را اعتبار بخشیده و فرمان را محدود و اعمال را کنترل کند به آن چیزی که مورد تایید سیستم اعتماد است. Yang et al. [9] یک IDS برای سیستم های synchrophasor پیشنهاد داده است که تشخیص حملات سایبری با استفاده از لیست سفید از بسته های با IP آدرس های قانونی، فرمت بسته درست، و مقادیر مجاز برای زمینه ها را انجام می دهد. IDS پیشنهادی یانگ برای انسان در وسط (MITM) و حملات انکار سرویس در برابر دستگاه های synchrophasor با استفاده از پروتکل IEEE C37.118 بررسی شد. ژانگ و همکاران [۱۰] یک IDS توزیع شده که تجزیه و تحلیل ترافیک ارتباطات در سطوح مختلف شبکه از شبکه

www.nn4e.com – nn4e@aol.com – 09360703858 – گروه نامیرا

های هوشمند از جمله شبکه های خانگی، شبکه های منطقه محله، و شبکه های گسترده را انجام می دهد. یک نمونه هوشمند در هر سطح گسترده شد تا طبقه بندی داده های مخرب و حملات سایبری احتمالی را با استفاده از الگوریتم های داده کاوی طبقه بندی کند. این نمونه ها سپس ارتباط برقرار کرده تا یک دیدگاه از سطح سیستم شبکه های ارتباطی را به منظور بهبود دقت تشخیص، فراهم کند. Hadeli و همکاران [۱۱] یک روش تشخیص ناهنجاری برای سیستم های کنترل صنعتی پیشنهاد داده که لیست سفید الگوهای ارتباطی قاونونی استخراج شده از پروتکل های سیستم کنترل صنعتی مختلف موجود در سیستم را ارائه می دهد.

Hadeli IDS از یک فایل توضیحات سیستم برای ارائه یک توصیف از الگوهای ارتباطی کلی که در سیستم کنترل صنعتی انتظار می رود، استفاده کرد. IDS پیشنهاد شده توسط [۹] - [۱۱] می تواند تغییرات مخرب در ترافیک شبکه را تشخیص دهد، اما هر سه IDS قادر به تشخیص ظرفیت مخرب که منجر به تغییرات نامعتبر در سیستم فیزیکی می شود، نیستند. به عنوان مثال، روش Hadeli et al.'s [11] نمی تواند یک دستور تزریقی اما معتبر تشخیص دهد تا حفاظ ایمنی را از یک IP آدرس معتبر انتقال دهد که می تواند یک خط انتقال را از سرویس خارج کرده و باعث خاموشی شود. یک IDS مبتنی بر خصوصیت خاص برای پیگیری حوادث پی در پی در یک زیرساخت پیشرفته انجام محاسبات مربوط (AMI) توسعه داده شده است. یک ماشین حالت دستی برای استخراج حالت های پی در پی قانونی سیستم از دو پروتکل AMI ها و دستگاه های حالت استفاده شد.

برای اثبات صحت دستگاه حالت، یک تکنیک بررسی مدل به منظور بررسی خصوصیات خاص مورد استفاده قرار گرفت. این IDS قابل استفاده با سیستم های انتقال نیست، چرا که سیستم های انتقال اعمال کنترل و اختلالات به مراتب بیشتر از AMI دارند. به این ترتیب، ایجاد یک چنین ماشین حالتی به صورت دستی می تواند بسیار گران قیمت باشد.

IDS های پیشنهادی دیگران برای شبکه های هوشمند سبب نیرومندسازی تئوری سیستم قدرت شده است. به عنوان مثال، والنزولا و همکاران [۱۳] از برنامه های جریان برق بهینه برای تشخیص حملات سایبری که داده های اندازه گیری سیستم را تغییر داده و منجر به این می شود که جریان قدرت به اشتباه اعزام شود. طالبی و همکاران [۱۴] یک مکانیزم برای شناسایی حملات داده های بد در یک سیستم قدرت با استفاده از تخمین وزن حالت ارائه داده است. اگر چه همه ی این کارها قادر به تشخیص داده تغییر یافته هستند، اما این IDS ها به یک نوع حمله محدود و نمی توانند برای تشخیص حملات دیگر علیه سیستم های قدرت، به کار روند.

B. دقت IDS خصوصیت محور

دقت تشخیص IDS خصوصیت محور بستگی به این دارد که با چه دقتی خصوصیات رفتار سیستم را توضیح می دهند. یک راه امیدبخش برای بهبود دقت خصوصیات استفاده از داده کاوی است. مسیر داده کاوی بر روی یک چارچوب IDS پیشنهاد شده توسط لی و همکاران [۱۵] اعمال شد، که IDS مبتنی بر علامت و IDSمبتنی بر ناهنجاری ترکیب شده است. برنامه های داده کاوی به حجم زیادی

www.nn4e.com – nn4e@aol.com – 09360703858 – گروه نامیرا

از اطلاعات ورودی اعمال می شود تا علائم حمله ها و الگوهای رفتار طبیعی را آموخته و به طور خودکار قواعد تشخیص را به وجود آورد. لی و همکاران [۱۵] نشان دادند که علائم برای حملات و الگوها برای رفتارهای طبیعی سیستم ایجاد شده با استفاده از مسیر داده کاوی، دقیق هستند و این کار را با مقایسه نتایج تشخیص خود با همه شرکت کنندگان در این برنامه ارزیابی تشخیص نفوذ آژانس پروژههای تحقیقات پیشرفته دفاع، آماده شده توسط آزمایشگاه دانشگاه MIT لینکلن، انجام دادند. IDS لی و همکاران [۱۵] در اصل برای IDS بدون تابعیت طراحی شده بود به همین دلیل آن را نمی توان به طور مستقیم بر IDS مبتنی بر خصوصیات اعمال کرد. یک الگوریتم داده کاوی جدید بایستی برای کشف رویدادی پی در پی برای خصوصیات طراحی شود.

C. تکنیک های داده کاوی برای یادگیری خصوصیات

یک خصوصیت برای یک سناریو شامل یک توالی از رویدادهای اجرایی یا حالت های سیستم است. ماهیت خصوصیات نیاز به تکنیک داده کاوی اعمال شده به IDS پیشنهادی دارد تا قادر باشد الگوهای ترتیبی را واکاوی کرده و رابطه وابسته بین حوادث را شناسایی کند. تکنیک داده کاوی مورد استفاده در این مقاله، از تکنیک الگوهای پی در پی واکاوی استفاده می کند که الگوهای فعالیت داده های زمان دستور را کشف می کند. الگوریتم استخراج الگوهای ترتیبی در ابتدا در [۱۶] ذکر شده بود. لین و همکاران. [۱۷] آن را به کار بردند تا داده های الگوهای موجود در فرآیند مدیریت مراقبت بالینی از مشتری که شامل سوابق بیمار و داده ها در

www.nn4e.com – nn4e@aol.com – 09360703858 – گروه نامیرا

طول یک دوره از زمان درمان است، را کشف کنند. این تکنیک در [۱۸] با به کارگیری شبکه های Bayesian تمدید شد تا به صورت گرافیکی الگوهای فرآیندهای همودیالیز مختلف که متشکل از یک توالی از حالات فیزیولوژیکی بیماران که شامل عکس از داده های بالینی و سوابق بیمار است، به عنوان مثال، دمای بدن، نبض و غیره هستند را نشان دهد. در کار لین و همکارانش [۱۸] حالات به همراه احتمال به منظور پیش بینی قرار گرفتند.

برای کاری که در این مقاله آورده شده است، الگوریتم رشد FP [۱۹] در فرایند آموزش برای واکاوی الگوهای متوالی دائمی از داده های سیستم قدرت استفاده شد. رشد FP یک پیاده سازی از موارد متوالی است که واکاوی را انجام می دهند. یک مثال متداول از موارد متوالی انجام واکاوی، آنالیز جعبه بازار است که در آن فروشگاه ها تلاش می کنند تا ارتباط مشترک بین کالاهای خریداری شده توسط چندین مشتری را بیابند، برای مثال یافتن محصولاتی که اغلب با یکدیگر خریداری می شوند. مسیر متداول کاوش مشابه با آنالیز جعبه بازار است به جز مسیر متداول کاوش که حالات سیستم را می یابد که معمولا در یک مجموعه با هم پیدا می شوند. مسیر متداول کاوش همچنین دستورهای موقت حالات سیستم را حفظ می کند.

III. مسیر متداول کاوش

این مقاله از مفهوم مسیر متداول برای نشان دادن الگوهای رمزی شده در یک هم آمیزی زمانی که سنسورهای داده را نشانه گذاری کرده، استفاده کرده است. یک

www.nn4e.com – nn4e@aol.com – 09360703858 – گروه نامیرا

مسیر متداول شامل یک توالی از حالات سیستمی بحرانی در یک ترتیب موقت است. توصیف الگوریتم مسیر متداول کاوش نیاز به تعریف مفاهیم حالت، ویژگی، توالی، و مسیر دارد.

حالت برای توصیف وضعیت خود به خودی یک سیستم به کار می رود. یک حالت شامل یک مجموعه از اندازه گیری های مشاهده شده سیستم است یا ویژگی های f همچنین یک نشانه نرمال سازی شده زمانی (TS)، است. برای نمونه، S = { TS, f1, . . . , fn}. ارزش یک ویژگی از روی یک سنسور خوانده می شود. ارزش های ممکن برای یک ویژگی در یک محدوده هستند که دامنه آن نامیده می شود. یک ویژگی که ارزش های پیوسته در دامنه خود دارد، بایستی به محدوده های پایان پذیر جداسازی شوند تا از فضای حالت نامحدود جلوگیری شود.

یک مسیر که P نامیده می شود، یک لیست از حالات مشاهده شده سیستم است که در یک توالی موقت بر اساس TS هایشان مرتب شده اند، که Pi = { S1, S2, . . . , Sn} نامیده شده و با افزایش زمان مرتب شده اند. یک توالی، s، یک زیرمجموعه از مسیر است، برای نمونه s ⊆ P. ما یک توالی s را با { Si + 1, Si + 2, . . . , Si + m } مشخص می کنیم. یک مسیر P شامل توالی s است، اگرکه تمامی اجزای s در P در یک ترتیب ظاهر شوند. در یک مجموعه از توالی ها، یک توالی حداکثری است اگر که توالی در هیچ توالی دیگری قرار نداشته باشد.

در نظر بگیرید که G مجموعه ای از مسیرهای مشاهده شده برای یک سناریو Q باشد، بنابراین { G = { P1, P2, . . . , Pn در جاییکه n تعداد مسیرهای مشاهده

شده برای Q باشد. یک مسیر توالی s را حمایت می کند اگر که توالی در مسیر قرار داشته باشد. حمایت می تواند به صورت متریک تعریف شود در حالیکه حمایت توالی s درصدی از مسیرها در G است که توالی s را نیز شامل می شود. یک مسیر متداول برای سناریو Q هر توالی است که حمایتش بزرگتر از حمایت حداقلی آستانه بوده و بزرگتر از ماکزیمم باشد. اینجا ممکن است چندین مسیر متداول برای یک سناریو وجود داشته باشد. مسیرهای متداول، حالاتی را که مرتبا برای یک حالت اتفاق می افتد، بازتاب می دهد.

الگوریتم مسیر متداول کاوش شامل شش مرحله است. پنج مرحله اول مسیر،P، را برای هر سناریو، به وجود می آورد. اول، داده های خام از سناریوهای مختلف در سیستم، جمع آوری می شود. سپس، داده های خام با یک پایگاه داده منفرد ترکیب می شود. سنسورها ممکن است در زمان ها و فرکانس های مختلف اندازه گیری را انجام دهند. داده سنسور فرکانس پایین تر، با مقدار بالاتر اندازه گیری شده، بنابراین همه ی اندازه گیری های فرکانس بالا نگه داشته می شوند. سوم، اندازه هایی که پیوسته هستند، با فرمول اندازه گیری می شوند تا تعداد کل حالت های ممکن در پایگاه داده را کاهش بدهیم. دانش متخصصان برای طراحی محدوده برای هر سنسور، استفاده می شود.پایگاه داده، یک جدول است که هر سنسور را در یک ستون جای داده و ردیف ها حالت سیستم را در زمان افزایش TS نمایش می دهند. در مرحله چهارم، پایگاه داده تجزیه می شود تا همه ی حالات منحصر به فرد، به دست آید. پنجم، پایگاه داده با ترکیب همه ی ردیف هایی که حالت مشابه

دارند، فشرده سازی می شود.در مرحله ششم، همه ی مسیرهای شناخته شده برای یک سناریو، مجموعه G، با الگوریتم الگوهای متداول حداقلی رشد FP {۱۹}،بررسی می شوند، تا متداول ترین توالی حالت ها واکاوی شود. آستانه حمایتی توسط سعی و خطا و یا استفاده از دانش متخصصان، تعیین می شود. توالی های متداول حداکثری، مسیرهای متداول برای سناریو ها هستند.

مثال: مجموعه ای از مسیرها را در نظر بگیرید که در جدول ۱ آمده است. برای مثال، G = {P1, P2, P3, P4, P5}. اگر آستانه حمایتی حداقلی بر روی ۶۰٪ تنظیم شده باشد، مجموعه توالی های پی در پی در G که آستانه حداقلی را به دست می آورند، شامل {S1, S2, S3, S4, S5}, {S1, S3, S4, S5}, و { S1, S4, S5} است. برای این مثال {S1, S2, S3, S4, S5} حداکثر است و بنابراین مسیر متداول است. توالی {S1, S3, S4, S5} و {S1, S4, S5} حداکثر نیستند چراکه در { S1, S2, S3, S4, S5} قرار گرفته اند. در مقابل، اگر حداقل آستانه حمایتی تا ۷۰٪ تغییر کند، مجموعه ای از توالی ها در G که آستانه حداقلی حمایتی را به دست می آورند فقط شامل {S1، S4، S5} است. از آنجا که {S1، S4، S5} در این مورد آستانه را به دست آورده اند، این حداکثر است و یک مسیر متداول است.

جدول ۱ نیز مثال هایی از انواع ممکن مسیرها است که می تواند در مجموعه داده یافت شود. P1 نشان دهنده حالت ایده آل برای یک مسیر است که یک سناریو را نشان می دهد. P2 با P1 مطابقت می کند به جز در مواردی که یک زیر مجموعه

از حالت ها به تعویق می افتد. این ممکن است به علت یک خطای اندازه گیری و یا به دلیل پویایی سیستم قدرت رخ دهد. P3 دارای یک حالت اضافی است. حالت های اضافی هنگامی ممکن است رخ دهد که یکی از ویژگی ها در طول انتقال حالت نوسان کند. P4 نشان دهنده موردی است که مسیر مشابه است، اما یک حالت متفاوت از حالت ایده آل است. این می تواند هنگامی اتفاق افتد که یک رویداد که باید در T2 رخ دهد به جای آن در T3 رخ دهد، که حالت های S2 و S3 را عوض کند (آنها به S11، S12 تغییر می کنند). P5 نشان دهنده یک مسیر خطا است. در مسیر خطا هیچ توالی مطابقت با مسیر متداول نهایی ندارد.

جدول I. نمونه مسیر برای یک سناریوی

	T6	T5	T4	T3	T2	T1	
حالت ایده آل		S5	S4	S3	S2	S1	P1
حالت های تاخیر	S5	S4	S3	S2		S1	P2
حالت های اضافی	S5	S4	S3	S2	S10	S1	P3
حالت های تعریف شده		S5	S4	S12	S11	S1	P4
مسیر خطا		S25	S24	S23	S22	S1	P5

مسیر متداول به عنوان یک خصوصیت در طول طبقه بندی استفاده می شود. تغییر آستانه حداقل حمایتی، تعداد حالات در یک مسیر متداول را تغییر می دهد و می تواند دقت طبقه بندی را تحت تاثیر قرار دهد. ضروری نیست که یک مسیر متداول بیابید که با مسیر ایده آل مطابقت داشته باشد، و نه هدف این است که یک مسیر متداول بیابیم که منحصر به فرد برای یک سناریو باشد و منجر به حداکثر دقت

طبقه بندی شود. برای یک سیستم پر سر و صدا یک مسیر متداول کوتاه تر ممکن است نتایج طبقه بندی بهتری به دست دهد.

یک مسیر متداول برای یک خط به زمین (SLG) منفرد خطا بایستی دنباله ای از حالت های بحرانی نشان دهنده "در حال حاضر بالا رفتن "، "حرکت امدادی" و "در حال حاضر سقوط به صفر" است. توانایی یافتن یک مسیر متداول تا حد زیادی وابسته به کیفیت مسیرهای G دارد. برای مثال، اگر مسیرهای خطای بسیاری در G وجود دارد بسیار دشوار خواهد شد که توالی هایی را بیابیم که آستانه حداقل حمایتی را به دست آورند.

طبقه بندی با مقایسه حالت های سیستم مشاهده شده با حالت های مسیرهای متداول شناخته شده انجام می شود. مسیر تحت آزمون (PUT) با تمام مسیرهای متداول مقایسه می شود. اگر $cpi \subseteq PUT$ باشد، cpi یکی از داوطلبین مسیر متداول است. PUT به عنوان سناریوی منطبق حداکثر داوطلب مسیر متداول از میان مجموعه داوطلبین مسیر متداول، طبقه بندی می شود. اگر بیش از یک داوطلب مسیر متداول، حداکثر باشند، PUT به عنوان ناشناخته طبقه بندی می شود.

بقیه این مقاله یک مطالعه موردی است که الگوریتم واکاوی مسیر متداول را برای سیستم انتقال سه اتوبوس دو خط، برای طبقه بندی ۲۵ سناریو سیستم قدرت اعمال می کند.

IV. آزمون بستر معماری

A. حفاظت مسافت برای خطوط انتقال

طرح حفاظت مسافت یکی از محبوب ترین طرح ها برای حفاظت از خطوط انتقال است. اصل عمل تشخیص می دهد که مقاومت ظاهری خط انتقال ولتاژ بالا تقریبا متناسب با طول آن است. این به این معنی است که مقاومت ظاهری "دیده شده" توسط ایستگاه تقویت در طول یک گسل متناسب با فاصله بین نقطه از گسل و ایستگاه تقویت است. فاصله ایستگاه تقویت با مناطق حفاظت چندگانه کد گذاری شده است. به هر منطقه یک آستانه مقاومت ظاهری و زمان حرکت اختصاص داده شده است. ایستگاه های تقویت مناطق حفاظت همپوشان دارند تا فراوانی سیستم محافظت را فراهم کنند. یک منطقه ایستگاه تقویت۱، بخشی از یک منطقه دیگر از ایستگاه تقویت ۲ و غیره است. در این مطالعه، طرح حفاظت مسافت با غیرفعال کردن زمان معکوس تاخیر پشتیبان گیری و محدود کردن تعدادی از مناطق حفاظت برای هر ایستگاه تقویت به ۲ عدد، ساده سازی شد. شکل۱ یک سیستم انتقال دو خط سه اتوبوس از سیستم IEEE چهار اتوبوس سه ژنراتور است. مناطق ایستگاه تقویت R1، شماره های ۱ و ۲ به صورت جعبه های با خطوط منقطع نشان داده شده است. هر ایستگاه تقویت حفاظت اولیه تا ۸۰٪ از خط (منطقه ۱ حفاظت) و حفاظت پشتیبانی (منطقه ۲ حفاظت) تا ۱۵۰٪ خط در صورتی که حفاظت اولیه موفق نشود، را فراهم می کند. زمان حرکت برای حفاظت منطقه ۱ در نظر گرفته شده که فوری باشد در حالی که زمان حرکت برای حفاظت منطقه

www.nn4e.com – nn4e@aol.com – 09360703858 – گروه نامیرا

۲، زمان تاخیر برای جلوگیری از قطع نادرست است مگر اینکه ایستگاه تقویت اولیه با شکست مواجه شده باشد.

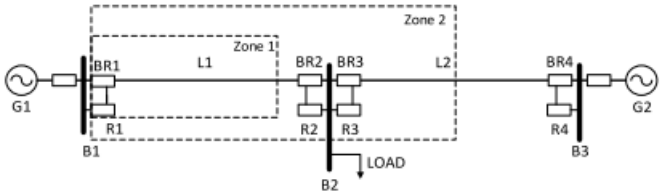
B. آزمون بستر معماری

یک سخت افزار در حلقه بستر آزمون، که در شکل ۲ نشان داده شده است، برای پیاده سازی سناریوی سیستم قدرت و تولید داده ها استفاده شد. یک شبیه ساز دیجیتال زمان واقعی (RTDS) برای شبیه سازی خطوط انتقال، قطع کننده ها، ژنراتورها و بار استفاده شد. چهار ایستگاه تقویت فیزیکی در یک پیکربندی سخت افزار به شکل حلقوی، با سیم به RTDS وصل شدند. ایستگاه های تقویت، طرح حفاظت مسافت دو منطقه را اجرا کردند. ایستگاه تقویت می لغزد و هنگامی که خطائی در خط انتقال رخ می دهد، قطع کننده ها را باز می کند. تمام ایستگاه های تقویت شامل عملکرد واحد اندازه گیری یکپارچه فازور (PMU) برای اندازه گیری حالت خط انتقال سیستم قدرت دولتی می باشد، با این حال، PMUs به طور جداگانه در نمودار کشیده شده بودند، زیرا ایستگاه های تقویت توسط اندازه گیری synchrophasor جریان های MODBUS/پروتکل کنترل انتقال (TCP) و PMUs با استفاده از IEEE پروتکل C37.118. PMUs کنترل می شوند. PMUs داده های اندازه گیری synchrophasor در نرخ ۱۲۰ نمونه / S، به متمرکز کننده داده فازور (PDC) که قاب های شبکه مشترک را از PMU چندگانه جمع کرده و به به برنامه OpenPDC می فرستد. مجموعه ای از اسکریپت ها شبیه سازی را با استفاده از ایجاد تغییرات تصادفی، گرفتن اندازه گیری، برچسب

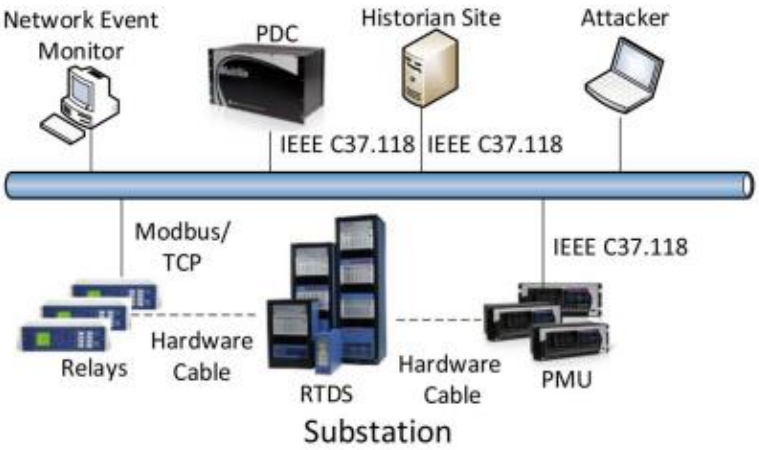
www.nn4e.com – nn4e@aol.com – 09360703858 – گروه نامیرا

داده ها بر اسا نوع سناریو، و ادغام داده ها از منابع مختلف را به یک فایل کنترل می کنند. داده های اندازه گیری synchrophasor شامل فرکانس، فازور فعلی، فازور ولتاژ و قطعات توالی، است. چهار ایستگاه تقویت، منابع تغییر حالت ایستگاه تقویت نشانه گذاری شده با زمان بودند. یک IDS مبتنی بر علامت، Snort بر روی کامپوتر اجرا می شود تا فعالیت شبکه را تشخیص دهد. Snort هنگامی که فرمان فعالیت های حرکت راه دور در شبکه را تشخیص دهد، اخطار می دهد. . Snort به خودی خود، نمی تواند بین دستورات مشروع و نامشروع حرکت راه دور تشخیصی انجام دهد، چراکه آنها همه در شبکه به یک شکل ظاهر می شوند.

شکل۱. طرح حفاظت از راه دور در یک سیستم انتقال سه اتوبوس، دو خط.



شکل۲. سخت افزار در بستر آزمون حلقه.



جدول II . سناریو های شبیه سازی

نام سناریو	شرح
Q1-Q2	گسل منفرد خط به زمین به ترتیب بر روی خط L1 و L2.
Q3-Q4	گسل بازپخش حمله که یک گسل معتبر روی L1 و L2 را به ترتیب، تقلید می کند.
Q5-Q6	از راه دور، هر دو ایستگاه تقویت را در هر دو انتهای خط انتقال (به ترتیب L1 و L2) برای حفظ و نگه داری، باز می کند.
Q7-Q10	فرمان تزریق حمله برای بازگشایی از راه دور یک ایستگاه تقویت (به ترتیب R4, R3, R2, R1).
Q11-Q12	فرمان تزریق حمله برای بازگشایی از راه دور دو ایستگاه تقویت (به ترتیب R1 و R2 یا R3 و R4). این حمله، سناریوی نگه داری(Q5-Q6) را تقلید می کند.
Q13-Q16	یک ایستگاه تقویت (به ترتیبR4, R3, R2, R1) در طول یک گسل بر روی خطی که به ایستگاه تقویت متصل است، از کار می افتد.
Q17-Q20	یک ایستگاه تقویت (به ترتیبR4, R3, R2, R1) در طول یک رویداد نگه داری بر روی خطی که به ایستگاه تقویت متصل است، از کار می افتد.
Q21-Q22	دو ایستگاه تقویت (به ترتیبR4, R3, R2, R1) در طول یک گسل بر روی خطی که به آن ایستگاه های تقویت متصل است، از کار می افتند.
Q23-Q24	دو ایستگاه تقویت (به ترتیبR4, R3, R2, R1) در طول یک رویداد نگه داری بر روی خطی که به آن ایستگاه های تقویت متصل است، از کار می افتند.
Q25	عملکرد معمول سیستم. هیچ رویدادی اتفاق نمی افتد.

یک کنترل پنل کامپیوتر عملکرد سیستم مدیریت انرژی (EMS) را شبیه سازی می کند. شبیه سازی EMS برای قطع یک خط انتقال جهت تعمیر و نگهداری از راه دور توسط ایستگاه تقویت حرکت راه دور و از طریق بسته شبکه MODBUS / TCP، استفاده شد. یک لگاریتم EMS، TS یک چنین رویداد تعمیر و نگهداری

خط را فراهم می کند. در این مقاله، فرض بر این است که یک کامپیوتر مهاجم با موفقیت به شبکه عملیاتی ابزار نفوذ کرده و می تواند حملات سایبری از یک گره در شبکه عملیاتی را راه اندازی کند. سناریوهای اختلالات سیستم قدرت، عملیات عادی، و حملات سایبری سیستم های قدرت علیه سیستم قدرت شبیه سازی شده و اجزای آن اعمال می شود. لگاریتم داده ها از سیستم synchrophasor، ایستگاه تقویت، Snort، و EMS شبیه سازی شده، به دست آمده اند. تمام لگاریتم داده های زمان مهر شده و با نام سناریو، شبیه سازی شده اند.

D. سناریوهای بستر آزمون

سناریوهای سیستم قدرت مورد استفاده برای آموزش و اعتبار IDS ارائه شده در این مقاله به سه گروه طبقه بندی شده است: ۱) سیستم قدرت تک خط به زمین گسل؛ ۲) عملیات عادی؛ و ۳) حملات سایبری. هر دسته در این بخش با جزئیات آورده شده است. در مجموع ۲۵ سناریو وجود دارد که هر یک با حرف "Q" و همراه با یک شماره نامگذاری شده است. بار سیستم در آغاز هر سناریو به صورت تصادفی شده است. گسلهای SLG سیستم قدرت متعلق به خانواده گسل انتقال هستند و تا ۷۰٪ از گسلهای یک سیستم قدرت [۲۰] را تشکیل می دهند. در این مقاله، تنها فاز یک به زمین گسل شبیه سازی شده است چراکه هر فاز به زمین گسل دارای ویژگی های مشابه است. گسل فاز یک به زمین به عنوان "گسل" به صورت مختصر در ادامه این مقاله، آورده می شود. جدول II خلاصه ای از سناریوهای شبیه سازی شده که برای اعتبار بخشی به IDS استفاده می شود،فراهم

www.nn4e.com – nn4e@aol.com – 09360703858 – گروه نامیرا

می کند. برای سناریوهای گسل SLG (Q1, Q2) ایستگاه تقویت بلافاصله برای منطقه ۱ عمل می کند و پس از یک تاخیر زمانی برای گسل ها در منطقه ۲ عمل خواهند کرد. مدل های طرح خودکار بستن یک طرح بازبست سه فاز با سرعت بالا [۲۱] که پس از یک ثانیه قطع کننده ها را می بندد.

SLG حملات مجدد گسل (Q3 و Q4) تلاش می کند تا با یک گسل معتبر با تغییر و اندازه گیری سیستم پس از ارسال دستور حرکت غیر قانونی به ایستگاه تقویت در انتهای خط انتقال، برابری کند. این حمله ممکن است به سردرگمی منجر شود و به طور بالقوه باعث شود که یک اپراتور یک فعالیت کنترلی نامعتبر را انجام دهد. یک اسکریپت غیب گو برای وارد کردن یک حمله MITM بین سخت افزار PDC و نرم افزار OpenPDC استفاده شد. حملات اندازه گیری synchrophasor از یک گسل SLG معتبر را دوباره انجام می دهد و پس از آن پخش مجدد دستورات حرکت ایستگاه تقویت را در خط آسیب دیده انجام می دهد. سناریوهای تعمیر و نگهداری خط انتقال (Q5 و Q6) وضعیت زمانی که یک اپراتور از راه دور حرکتهای ایستگاه تقویت برای باز کردن قطع کننده در هر دو انتهای یک خط انتقال را به خط خارج از سرویس برای تعمیر و نگهداری خط، انجام می دهد را شبیه سازی می کند. اپراتوری که آغاز دستورات حرکت از راه دور را ثبت می کند و زمان در لگاریتم کنترل پنل ثبت می شود.

سیستم قدرت حملات سایبری ممکن است از داخل، هکرهای آماتور، فعالان سیاسی، سازمان های جنایی، دولت ها و تروریست ها سرچشمه بگیرد. حملات

سایبری ممکن است به عنوان یک مزاحم به نظر رسد و یا ممکن است برای سیستم سقوط را به ارمغان بیاورد. حملات می تواند از درون پستهای سیستم قدرت، یک مرکز کنترل، و یا در انتقال و توزیع زیرساخت با بهره برداری از نقاط ضعف در سیاست های امنیتی فیزیکی انجام شده باشد. در روش دیگر، حملات ممکن است از نقص امنیت و آسیب پذیری در نرم افزار، دستگاه، زیرساخت های ارتباطی، و یا پروتکل های ارتباطی برای نفوذ الکترونیکی در شبکه های عملیاتی سیستم قدرت، استفاده کند. سه نوع از حملات شبیه سازی شده است: ۱) تزریق فرمان حرکت ایستگاه تقویت ۲) غیر فعال کردن عملکرد ایستگاه تقویت. و ۳) عمل مجدد گسل SLG. حملات تزریق فرمان حرکت ایستگاه تقویت (Q7-Q12) ایجاد احتمالات با ارسال دستورات حرکت ایستگاه تقویت غیر منتظره از راه دور از کامپیوتر یک مهاجم به ایستگاه تقویت در انتهای دو خط انتقال است. حمله تزریقی فرمان حرکت که برای این مقاله استفاده می شود، از نزدیک سناریوی تعمیر و نگهداری خط را تقلید می کند. دستور حرکت مخرب از گره دیگر در شبکه ارتباطات با یک IP آدرس قانونی جعلی سرچشمه می گیرد. از آنجا که این حمله از کنترل پانل کامپیوتر صورت نگرفته، در لگاریتم کنترل پنل ثبت نخواهد شد، با این حال، Snort مانیتور شبکه ترافیک این دستور حرکت از راه دور را تشخیص می دهد. حملات غیر فعال ایستگاه تقویت (Q13-Q24) اثرات داخلی ها که اقدامات کنترل غیر قانونی را انجام می دهند و یا نرم افزار مخرب کنترل سیستم های نرم افزاری برای دستکاری دستگاه های کنترل را به دست می گیرند،

تقلید می کند. یک اسکریپت پایتون به ثبت داخلی یک ایستگاه تقویت از طریق دستورات MODBUS / TCP فرستاده شده از کامپیوتر مهاجم، دسترسی دارد، که تغییر در تنظیمات ایستگاه تقویت مربوطه را تعریف می کند. حملات غیر فعال ایستگاه تقویت در حوادث خطا و تعمیر و نگهداری با هم همپوشانی دارند. سناریوی نهایی، Q25، نشان دهنده یک حالت سیستم با ثبات است. برای این سناریو، بار ممکن است تغییر کند، اما هیچ یک از حملات، اختلالات، و یا اقدامات کنترل شبیه سازی نمی شود. سناریوها با سیستم در حالت پایدار، شروع و پایان می یابد. به این ترتیب، تمام گسل ها پاک، خطوط انتقال خارج شده از سرویس برای تعمیر و نگهداری، به خدمت بازگشته، و تمام حملات قبل از شبیه سازی سناریو بعدی پایان می یابند.

D. داده های آزمون

داده های آزمون مورد استفاده برای این مقاله شامل لگاریتم داده ها در ارتباط با ۱۰۰۰۰ نمونه شبیه سازی شده از ۲۵ سناریو فوق الذکر است. لگاریتم داده ها یک فایل کاملا جدا با برچسب چندتایی است که شامل ۵۶ اندازه گیری سنسور و یک TS است. منابع ۵۶ داده شامل۵۲ اندازه گیری synchrophasor هستند؛ (۱۳) از هر ایستگاه تقویت در شکل ۱ تعیین مکان شده است. اطلاعات synchrophasor از یک ایستگاه تقویت منفرد شامل ولتاژ فاز و اندازه فازور فعلی، ولتاژ توالی صفر، مثبت و منفی و فازور فعلی و مقاومت ظاهری خط آشکار می باشد. داده synchrophasor در ۱۲۰ بار / S نمونه گیری شد. اطلاعات

وضعیت ایستگاه تقویت ، حوادث قطع کننده ها، هشدار Snort، و هشدار کنترل پانل نیز وارد شد. همه داده های وارد شده به یک مجموعه داده واحد ادغام شدند. یک نمونه از یک سناریو واحد توسط حدود ۲۰۰۰ تاپل در مجموعه داده ها از آزمون ارائه شده است. این مربوط به حدود ۱۷ ثانیه از زمان سیستم شبیه سازی شده در هر سناریو است. در مجموع، داده های آزمون بیش از دو میلیون تاپل دارند. هر تاپل در داده های آزمون برچسب شده است. حدود نیمی از داده ها ی آزمون برای آموزش طبقه بندی کننده و نیم دیگر برای تست دقت طبقه بندی استفاده شد. در این مقاله، ۱۵ ویژگی مورد استفاده قرار گرفت. اندازه فاز فعلی که در هر ایستگاه تقویت اندازه گیری شده، وضعیت ایستگاه تقویت برای هر ایستگاه تقویت، وضعیت هشدار Snort برای هر ایستگاه تقویت، و کنترل پنل وضعیت حرکت از راه دور.

V. آموزش IDS

این بخش، روند ساخت و ساز IDS را سندسازی می کند. در مرحله اول، مرحله قالب بندی اطلاعات، لگاریتم داده های ورودی را به یک پایگاه داده وقایع اندازه گیری شده (MED) تبدیل می کند. در مرحله بعد، مراحل یادگیری خصوصیات، MED را پردازش کرده تا روش های متداول، مجموعه ای منحصر به فرد از حالات سیستم در یک ترتیب موقت، برای هر سناریو ی برچسب شده، را بیاموزد. در نهایت، یک نمودار ساخته می شود که شامل مسیر متداول برای همه سناریو ها است.

A. قالب بندی داده ها

گام اول از روند قالب بندی داده ها تدریج ویژگی است. تدریج ویژگی نیاز به تخصص دامنه دارد. ویژگی های با ارزش که می تواند ارزش پیوسته بگیرند، در یک محدوده متناهی برای محدود کردن اندازه فضای حالت، ترسیم شده اند. ویژگی هایی که مقادیر گسسته می گیرند، به طور کلی بدون تغییر گذاشته می شوند، مگر اینکه تعداد ارزشهای گسسته بزرگ باشد. اندازه گیری فاز جریان یک عدد حقیقی است و به همین دلیل باید در یک محدوده گسسته گروه بندی شود. اندازه فاز فعلی، به محدوده های طبیعی و بالا جدا سازی شد. محدوده نرمال۰ تا A۱۱۹۹ بود. محدوده بالا تمام مقادیر بزرگتر یا مساوی با ۱۲۰۰ A بود. حالت ایستگاه تقویت، هشدار Snort، و کنترل پنل ویژگی های وضعیت حرکت از راه دور همه باینری بود. ارزش ممکن ایستگاه تقویت حرکت و یا بی حرکت است. ارزش ممکن وضعیت هشدار Snort، هشدار و بدون خطر است. ارزش ممکن وضعیت کنترل پنل حرکت از راه دور، حرکت و بی حرکت است. MED یک مجموعه داده فشرده ادغام شده با ویژگی های کوانتیده است. اطلاعات از سنسورهای با نرخ نمونه پایین تر، از بالا نمونه برداری شده تا با نرخ نمونه برداری از سنسور با بالاترین نرخ نمونه برداری، مطابقت کند. روند نمونه برداری از بالا، بستگی به نوع سنسور دارد. نمونه گیری مداوم سنسور، ارزش آنها را در هر نمونه دوره بر اساس حالت رایج اندازه گیری، به روز رسانی می کند. مقدار جریان و ایستگاه تقویت به طور مداوم نمونه گیری می شوند. سنسور مبتنی بر رویداد هنگامی که تغییر حالت رخ می دهد، یک

پیام می فرستد. ویژگی های هشدار Snort و حالت کنترل پنل حرکت از راه دور، مبتنی بر رویداد هستند. برای هر یک، هنگامی که سنسور حضور یک رویداد را تشخیص می دهد، سنسور یک پیام فراهم می کند که نشان دهد این رویداد رخ داده است. در یک لگاریتم داده، اندازه گیری سنسوری که به طور مداوم نمونه گیری شده، یک ارزش می گیرد و این ارزش را در سراسر نمونه های متعدد حفظ می کند تا هنگامی که تغییرات حالت اتفاق افتد.

در مقابل، در لگاریتم داده ها، ویژگی های مبتنی بر رویداد برای تنها یک نمونه برای هر رویداد اندازه گیری حمایت می شوند. هنگامی که نمونه برداری رو به بالا انجام می شود، اندازه گیری های مداوم سنسور نمونه برداری شده، به نزدیکترین دوره نمونه پس از اندازه گیری ترسیم می شوند. تمام نمونه های بدون ارزش، ارزش نزدیکترین نمونه غیر قطعی را می گیرند. در این مقاله، اندازه گیری های مقدار جریان در ۱۲۰ نمونه / S اندازه گیری شد که بالاترین نرخ نمونه برداری از تمام ویژگی ها است. وضعیت ایستگاه تقویت، هشدار Snort، و کنترل پنل ویژگی های لگاریتم رو به بالا با توجه به روش فوق نمونه گیری شد.

یک MED نشان دهنده یک نمونه از یک سناریو است. TS های ردیف ها در MED با حذف زمان ردیف اول از سایر ردیف ها، نرمال سازی می شود. این باعث می شود که تمام MED ها از زمان ۰ شروع شوند.

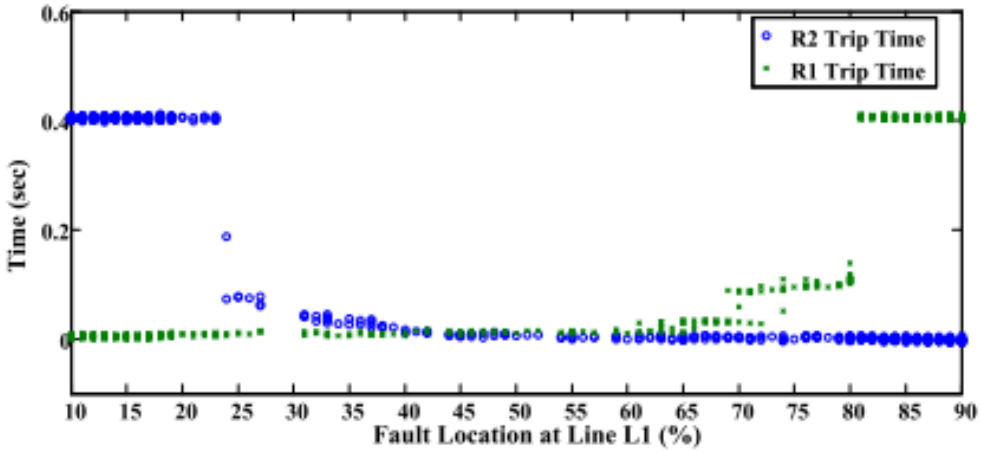
B. ایجاد و گروه بندی مسیرها

مسیر، یک لیست از حالات مشاهده شده سیستم به ترتیب زمانی است. مسیر ها توسط نمونه برداری پایین از MED در حالی که تمام حالت انتقال حفظ شده است، به دست می آیند. تغییر حالت، یک تغییر در ارزش هر سنسور بین دو نمونه MED است. این MED برای شناسایی تمام دوره های حالت سازگار، تجزیه می شود. دوره های حالت سازگار، رو به پایین نمونه گیری شده، از یک کاربر که دوره نمونه را تعریف کرده، استفاده می کند. در این مقاله، دوره نمونه ۰٫۵ s بود. هر حالت منحصر به فرد، یک حالت شناسایی (SID) معین می شود، و تمام حالات شناخته شده در یک پایگاه داده حالت ذخیره می شوند. یک مسیر برای هر MED استخراج شده است. یک سناریوی منفرد، ممکن است بسیاری از مسیرهای منحصر به فرد با توجه به ماهیت پویای سیستم های قدرت، تغییرات در دستور حالتها در یک مسیر، و به علت تغییرات در زمان رویداد، داشته باشد. استفاده از مسیرهای خام به دست آمده از فرایند استخراج برای طبقه بندی، دقت ضعیف طبقه بندی را نتیجه می دهد. الگوریتم استخراج مسیر متداول استفاده می شود تا گروه های بزرگتر مسیر را به یک مجموعه مسیر متداول که نشان دهنده تنوع طبیعی است و به عنوان مجموعه ای از علائم برای هر سناریو هستند، کاهش دهد. گروه بندی یک مرحله اختیاری است که داده های ورودی را پیش پردازش می کند تا کلاس های بزرگ را به زیر کلاس های کوچکتر جداسازی کند. گروه بندی می تواند به طبقه بندی دقیق تر بیانجامد، بویژه وقتیکه زیر کلاس ها به اندازه کافی از یکدیگر متفاوت هستند.

شکل. ۳ به وضوح مناطق ۱ و ۲ مرزهای حرکت برای هر دو ایستگاه تقویت را نشان می دهد. علاوه بر این، شکل ۳ نشان می دهد که بار حرکت ایستگاه تقویت با محل خطا به خصوص در منطقه محل خطا از ۲۴ درصد به ۷۹ درصد از خط انتقال متفاوت است. زمان حرکت ایستگاه تقویت برای شکل ۳ از MED محاسبه شد و به عنوان حالت ایستگاه تقویت، انتقالات از بسته به باز منهای زمان اولیه جریان خط که مقدار بالایی است. رفتار سیستم همچنین با تغییرات بار سیستم تغییر می کند. در حالت ایده آل، موارد گسل SLG از دو منطقه طرح حفاظت از راه دور، می تواند به سه گروه با توجه به منطقه خط که در آن گسل رخ می دهد، جدا شود. گروه ۱ شامل گسل از طول خط است که توسط منطقه۱، ایستگاه تقویت R1 و منطقه ۲ ایستگاه تقویت R2 حفاظت می شود. از شکل ۲ می توان دریافت که، گروه ۱ شامل گسل هایی است که بین ۱۰٪ و ۲۳٪ از خط رخ می دهد. برای گروه ۱ گسل ها، ایستگاه تقویت R1 باید فورا حرکت و R2 باید پس از ۰,۴ ثانیه حرکت کند. گروه ۲ شامل گسل حفاظت شده توسط منطقه ۱ ایستگاه تقویت R1 و R2 است. هر دو ایستگاه تقویت باید فورا برای گسل های گروه ۲ حرکت کنند. از شکل ۳ می توان دریافت که، گسل گروه ۲ بین ۲۴٪ و ۷۹٪ از خط رخ می دهد. گروه ۳ شامل گسل هایی است که توسط منطقه ۲ ایستگاه تقویت R1 و منطقه ۱ ایستگاه تقویت R2 حفاظت می شود. ایستگاه تقویت R1 باید پس از ۲۰ چرخه حرکت و R2 باید فورا برای گروه ۳ گسل ها حرکت کند. از شکل ۳ می توان دریافت که گروه ۳ گسل ها بین ۸۰٪ و ۹۰٪ از خط رخ می دهد.

زمان حرکت مشاهده شده در گروه ۲ تمایل دارد که افزایش یابد همینکه به نقاط مرزی مناطق ۱ و ۲ نزدیک می شود. برای جبران این رفتار مشاهده شده، مسیر گسل SLG توسط محل گسل در گروه های زیر گروه بندی شده است: ۱۰٪ -۲۳٪ و ۲۴٪ -۲۹٪، ۳۰٪ -۳۵٪، ۳۶٪ -۴۰٪، ۴۱٪ -۶۰٪، ۶۱٪ -۶۵٪، ۶۶٪ -۷۰٪، ۷۱٪ -۸۰٪ و ۸۱٪ -۹۰٪. علاوه بر این، مشاهده شد که زمان های حرکت تا حدی با بار سیستم همبسته است. به عنوان یک نتیجه، مسیر گسل SLG توسط محل گسل و بار گروه بندی شد. چهار محدوده بار مورد استفاده قرار گرفت: ۲۰۰-۲۴۹، ۲۵۰-۳۹۹، ۳۰۰-۳۴۹، ۳۵۰-۳۹۹ مگاوات. این گروه بندی مسیرهای SLG گسل را به ۹*۴= ۳۶ زیر گروه تقسیم می کند.

شکل۳. زمان حرکت ایستگاه تقویت در مقابل مکان گسل برای ایستگاه های تقویت R1 و R2.



C. واکاوی مسیر متداول

برای این آزمایش مجموعه G شامل ۵۰۰۰ مسیر خام از ۵۰۰۰ نمونه ۲۵ سناریو است. الگوریتم واکاوی مسیر متداول، ۴۷۷ مسیر متداول در تمام سناریو ها تولید کرده است. تعداد حداقل و حداکثر مسیرهای متداول برای یک سناریوی واحد به ترتیب ۴ و ۵۳ به دست آمد. ۱۵ سناریو گسل SLG دارای ۴۲۱ مسیر متداول گسترش یافته در میان آنها است. ده سناریو باقی مانده ۵۶ مسیر متداول داشتند. تعداد زیاد مسیر متداول برای گسل SLG به دلیل تنوع زیاد در زمان های حرکت ایستگاه تقویت، و همچنین محل گسل و بارگذاری متفاوت سیستم است.

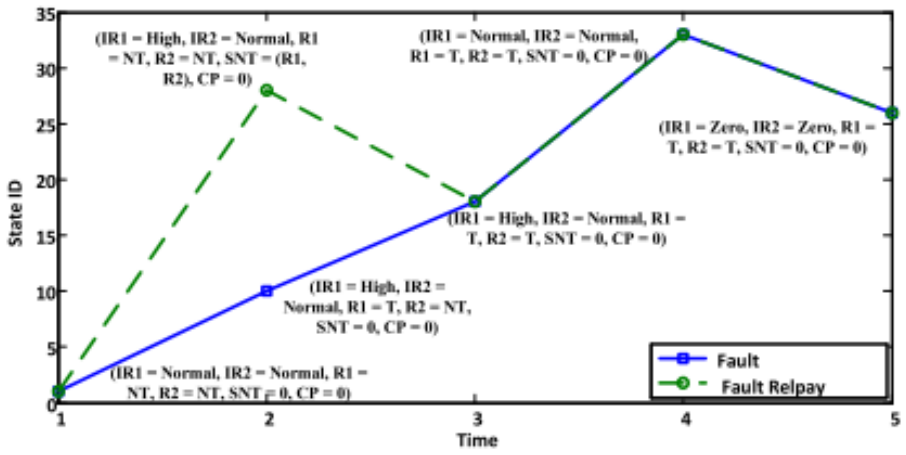
مسیر متداول می تواند دو بعدی ترسیم شود و مختصات با محور y نشان دهنده کد شناسایی حالت (ID حالت) و محور x نشان دهنده TSS نرمال سازی شده است. لبه بین دو راس نشان دهنده انتقال زمانی بین دو حالت است. هر راس با اطلاعات حالت مشخص شده است. توجه داشته باشید که برای صرفه جویی در فضا، تنها ویژگی های لازم نمایش داده می شود. شکل ۴ مسیر مشترک برای دو سناریو، یک گسل در ۳۶٪ -۴۰٪ محل گسل از خط L1 و یک حمله باز پخش گسل در خط L1 را نشان می دهد. گسل و مسیر بازپخش گسل هر دو در حالت عادی سیستم شروع می شود. برای گسل های واقعی، PMU جریان بالا را در زمانی که یک گسل موجود است، اندازه گیری می کند در حالی که برای حمله بازپخش گسل، مهاجم اندازه گیری های جریان بالا به PDC را تزریق می کند.

این باعث می شود حالت دوم هر دو مسیر متداول بالا تشخیص داده شود در ایستگاه تقویت R1، به عنوان مثال، IR1 = high. با این حال، این مسیرها

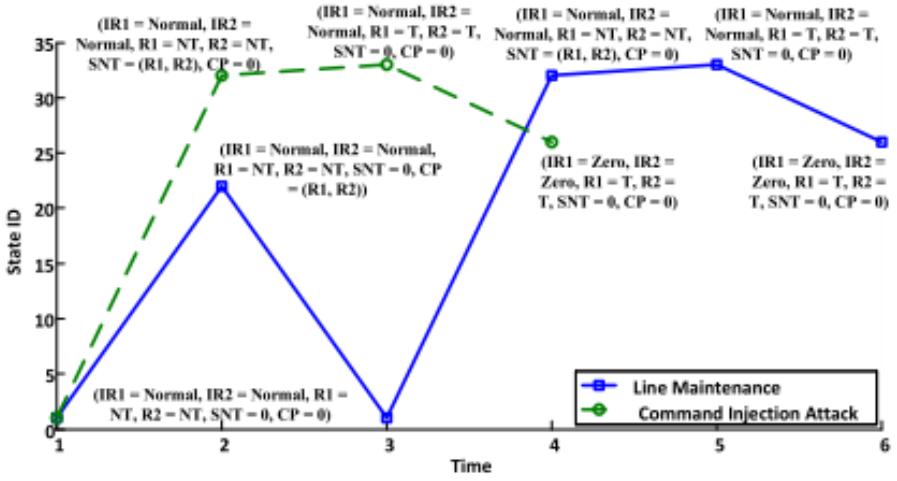
بلافاصله فرق می کنند به خاطر اینکه برای بازپخش گسل، مهاجم بایستی فرمان حرکت ایستگاه تقویت را به ایستگاه تقویت R1 و R2 به طور هم زمان، تزریق کند. به این ترتیب، حالت دوم برای حمله بازپخش گسل است که دستورات سفر به R1 و R2 تشخیص داده شده توسط Snort را دارد، به عنوان مثال، SNT = (R1, R2) در شکل ۴.

شکل ۵ مسیرهای متداول برای تعمیر و نگهداری خط و تزریق فرمان سناریوهای حمله را نشان می دهد. تفاوت اصلی بین دو سناریو، فرمان برای باز کردن ایستگاه تقویت R1 و R2 است که از کنترل پنل کامپیوتر برای سناریوی تعمیر و نگهداری خط، سرچشمه می گیرد. این باعث می شود که لگاریتم کنترل پنل شامل یک پیام دستور حرکت باشد. راه معمول برای سناریوی تعمیر و نگهداری خط شامل یک حالت است که تشخیص لگاریتم کنترل وقایع (برای مثال CP = (R1, R2))) و حالت هایی که دستور حرکت از راه دور تشخیص داده شده Snort بسته شبکه (برای مثال SNT = (R1, R2)) را ثبت می کند. مسیر متداول برای تزریق فرمان شامل هشدار Snort است اما حالت لگاریتم کنترل پانل را در بر نمی گیرد. شکل ۴ و ۵ نشان می دهد که مسیر متداول شامل حالت های بحرانی برای سناریو های مختلف است. سهم اصلی الگوریتم واکاوی مسیر، توانایی به وجود آوردن خودکار مسیرهای منحصر به فرد برای هر نوع سناریو از مجموعه داده است که رفتار مرتبط با سناریو را اندازه گیری می کند.

شکل ۴. مستندسازی مختصات دو بعدی گسل (Fault) در مقابل گسل بازپخش (Fault replay) حمله مسیرهای متداول



شکل ۵. مستندسازی مختصات دوبعدی تعمیر و نگهداری خط (Line maintenance) در مقابل فرمان تزریق حمله (command injection attack) مسیرهای متداول



VI. ارزیابی

سه روش برای ارزیابی IDS استفاده شد. اول، IDS برای طبقه بندی ۵۰۰۰ نمونه از سناریوهای مجموعه داده های آزمون شرح داده شده در بخش چهارم این مقاله

استفاده شد. ماتریس های سردرگمی تهیه شده اند تا دقت IDS را نشان دهد. بررسی دقیق توانایی الگوریتم برای طبقه بندی گسل SLG توسط محل گسل نیز ارائه شده است. دوم، آموزش و آزمایش با مجموعه ای از چهار سناریوی از دست رفته از مجموعه داده ها تکرار شد. این تست برای نشان دادن توانایی IDS ها در تشخیص حملات روز صفر و سناریو های ناشناخته استفاده شد. در نهایت، هزینه و عملکرد IDS با اندازه گیری مقدار زمان پردازش و حافظه مورد نیاز در طول آموزش و ارزیابی اندازه گیری شد.

جداول III و IV ماتریس های سردرگمی را ارائه می کنند که برای ۲۵ سناریوی تست شده است. ماتریس های سردرگمی در دو جدول جدا آورده شدند تا به آنها اجازه دهد که در عرض ستون این مقاله جای شوند. ردیفی که با برچسب "Oth" نشان داده شده، حالات Q14-Q25 در جدول III و Q1-Q13 در جدول IV را نشان می دهد. ردیفی که برچسب "Unk" دارد، تعداد نمونه هایی که با مسیر مشترک منطبق نبودند را فراهم می کند. در نهایت، ردیف با برچسب "UNC" تعداد موارد با طبقه بندی نمشخص را نشان می دهد که به علت تطبیق با بیش از یک مسیر متداول، از بیش از یک سناریو فراهم شده است.

www.nn4e.com – nn4e@aol.com – 09360703858 – گروه نامیرا

جدول III. ماتریس سردرگمی برای سناریوی Q1-Q13

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13
Q1	505	0	31	0	0	0	0	0	0	0	0	0	0
Q2	0	502	0	34	0	0	0	0	0	0	0	0	0
Q3	10	0	301	0	0	0	0	0	0	0	0	0	0
Q4	0	6	0	321	0	1	0	0	0	0	0	31	0
Q5	0	0	0	0	130	0	0	0	0	0	0	0	0
Q6	0	0	0	0	0	108	0	0	0	0	0	0	0
Q7	0	0	0	0	0	0	67	0	0	0	0	0	0
Q8	0	0	0	0	0	0	0	54	0	0	0	0	0
Q9	0	0	0	0	0	0	0	0	99	0	0	0	0
Q10	0	0	0	0	0	0	0	0	0	57	0	0	0
Q11	0	0	6	0	1	0	0	0	0	0	127	0	0
Q12	0	0	0	0	0	0	0	0	0	0	0	104	0
Q13	0	0	0	0	0	0	0	0	0	0	0	0	179
Oth.	1	2	3	1	0	0	0	0	0	0	0	0	0
Unk	3	1	4	0	0	1	35	32	0	26	0	0	0
Unc	0	2	8	4	0	0	0	0	0	0	0	0	0

جدول IV. ماتریس سردرگمی برای سناریوی Q14-Q25

	Q14	Q15	Q16	Q17	Q18	Q19	Q20	Q21	Q22	Q23	Q24	Q25
Q14	220	0	0	0	0	0	0	0	0	0	0	0
Q15	0	208	0	0	0	0	0	0	0	0	0	0
Q16	0	0	162	0	0	0	0	0	0	0	0	0
Q17	0	0	0	40	0	0	0	0	0	0	0	0
Q18	0	0	0	0	73	0	0	0	0	0	0	0
Q19	0	0	0	0	0	33	0	0	0	0	0	0
Q20	0	0	0	0	0	0	45	0	0	0	0	0
Q21	0	0	0	0	0	0	0	424	0	0	0	0
Q22	0	0	1	0	0	0	0	0	413	0	0	0
Q23	0	0	0	0	0	0	0	0	0	122	0	0
Q24	0	0	0	0	0	0	0	0	0	0	112	0
Q25	0	0	0	0	0	0	0	0	0	0	0	114
Oth	0	0	0	2	1	5	0	0	0	0	0	0
Unk	10	12	2	23	0	12	0	6	33	16	19	0
Unc	37	58	0	0	0	0	0	0	0	0	0	0

در مجموع، ۹۰,۴ درصد موارد تست شده به درستی طبقه بندی شد و ۲,۷ درصد موارد غلط طبقه بندی شد. ۴,۷٪ از موارد به عنوان ناشناخته و ۲,۲٪ با عنوان نا مشخص طبقه بندی شده است. تمامی موارد نامشخص مربوط به نمونه های گسل SLG بود که یک مسیر متداول را برای بیش از یک سناریو گسل همسان کرد.

IDS می تواند مثبت های کاذب به خصوص در مورد سناریوهای طراحی شده برای تقلید از یک سناریو یا رویداد غیرتهاجمی، تولید کند. در این مقاله، نرخ مثبت کاذب برای تمام سناریو های غیر تهاجمی محاسبه شده و به غلط مهاجمی طبقه بندی شده اند. سناریوهای Q1 و Q2، هر دو گسل SLG، 2.1٪ و ۱,۶٪ به ترتیب نرخ مثبت کاذب داشتند. در هر دو مورد، اکثریت مثبت کاذب به عنوان حملات بازپخش گسل طبقه بندی شده بودند. حملات بازپخش طراحی شده اند تا حملات SLG را تقلید کنند. یکی از یازده مثبت کاذب به عنوان یک ایستگاه تقویت غیر فعال کردن حمله طبقه بندی شد. سناریوهای Q5 و Q6، هر دو رویداد تعمیر و نگهداری خط، به ترتیب ۰,۸٪ و ۰,۹٪ نرخ مثبت کاذب را داشتند، که یک مثبت کاذب برای Q5 و Q6، به ترتیب بود. برای سناریوی Q5، مثبت کاذب یک تزریق فرمان حمله برای باز کردن هر دو ایستگاه تقویت در پایان خط انتقال بود. برای سناریوی Q6، مثبت کاذب، یک حمله بازپخش گسل بود. در هر دو مورد، توالی حالت ها در مسیر متداول برای سناریوی واقعی و سناریوی غلط طبقه بندی شده، در توالی حالات همپوشانی داشت. این همپوشانی همراه با تنوع در داده های مشاهده شده به دلیل سیستم قدرت و پویایی سیستم اندازه گیری می توانید منجر به مثبت کاذب شود.

ارزیابی اضافی برای طبقه بندی زیر گروه های سناریو Q1، یک گسل SLG در خط L1 انجام شد. مسیر برای Q1 به زیر گروه های محل خطا و بار مدار همان طور که قبلا ذکر شد، طبقه بندی شد. گسل SLG با نرخ دقت گروه بندی ۸۴,۶٪ بود در

حالی که ۱۱,۳۵ درصد از مسیرهای غلط طبقه بندی شده بود. تجزیه و تحلیل بیشتر نشان داد که اکثریت طبقه بندی های غلط زمانی رخ داده است که گروه های گسل SLG به عنوان اعضای یک همسایه یا گروه نزدیک گسل طبقه بندی شده بودند. آزمایش گروه بندی نشان می دهد که الگوریتم گروه بندی واکاوی مسیر متداول، یافتن مسیرهای منحصر به فرد حتی برای سناریو های مشابه را قدرت می بخشد.

ده برابر اعتبار متقاطع برای ارزیابی دقت تشخیص روز صفر سناریوهای حمله مورد استفاده قرار گرفت که در جدول V نشان داده شده است. برای هر دور از تست چهار سناریو به صورت تصادفی انتخاب شدند و از آموزش کنار گذاشته شدند، اما در مجموعه داده های آزمون قرار داشتند. متوسط دقت تشخیص برای سناریوهای حمله روز صفر ۷۳,۴۳ درصد بود. با این حال، مواردی وجود داشت که میزان شناسایی حمله روز صفر کم بود. به عنوان مثال، تجزیه و تحلیل نتایج دور سوم نشان داد که سناریوی Q6 (تزریق فرمان حرکت ایستگاه های تقویت R1 و R2) همیشه به غلط به عنوان سناریوی Q3 (حمله بازپخش گسل در خط L1) طبقه بندی شد. این اتفاق به این دلیل می افتد که مسیر متداول مورد انتظار برای Q6 و Q3 مشابه هستند. بنابراین، هنگامی که Q6 در آموزش در دسترس نیست، نمونه های Q6 به عنوان نمونه های Q3 طبقه بندی می شوند که منجر به طبقه بندی غلط می شود. در این مورد، هر دو Q6 و Q3 حمله هستند و حمله روز صفر به عنوان یک حمله دیگر طبقه بندی می شود که بهتر از این است که به عنوان

یک غیر مهاجم طبقه بندی شود. به منظور بهبود دقت طبقه بندی بین سناریوهای مشابه سنسورهای اضافی مورد نیاز است تا رویدادهایی که بین دو سناریو، متفاوت هستند را روشن کند. البته، در مورد روز صفر دشوار است که پیش بینی کنیم چه سنسورهای اضافی ممکن است مورد نیاز باشد.

آموزش و زمان فرایند طبقه بندی و استفاده از حافظه با استفاده از یک ماشین مجازی لینوکس اوبونتو با ۳,۵ گیگاهرتز CPU و ۲ گیگابایت حافظه اندازه گیری شد. آموزش ۰,۳۳ و ۳۴ مگابایت حافظه نیاز داشت. طبقه بندی موارد آزمون ۰,۸۵ ثانیه در هر نمونه سناریو برای تکمیل و ۲۶,۲ MB حافظه نیاز است. دسته های چندتایی الگوریتم کاوش داده مبتنی بر پردازش برای طبقه بندی گسل های سیستم قدرت و حملات سایبری در [۲۲] مورد استفاده قرار گرفت و از داده های مشابه استفاده شده برای کار های ارائه شده در این مقاله، استفاده کرده است. نتایج در [۲۲] برای طبقه بندی با دسته های باینری (حمله و غیرحمله)، سه دسته (حملات، غیرحملات، و طبیعی)، و چنددسته (حفظ همه طبقات) بود. IDS مسیر متداول مبتنی بر کاوش تمام روش های سنتی در [۲۲] برای دقت کلی در مورد چند دسته بهتر عمل کرد. ترکیبی از الگوریتم JRipper و Adaboost دقتی نزدیک به ۹۰ درصد تولید می کنند، که شبیه به دقت IDS مسیر متداول مبتنی بر کاوش است. تمامی روش های آزمون دیگر دقت به طور قابل توجهی پایین تر از IDS ارائه شده در این مقاله داشتند. روش باینری و سه طبقه در [۲۲] منجر به بهبود دقت در هزینه طبقه بندی دقیق می شود. IDS مسیر متداول مبتنی بر

www.nn4e.com – nn4e@aol.com – 09360703858 – گروه نامیرا

کاوش طبقه بندی هر نوع سناریو را دقیق دقیق فراهم می کند. دقت طبقه بندی بر اساس نوع سناریو برای سرعت بخشیدن به درک حملات و برای فعال کردن پاسخ خودکار یا کتابچه راهنمای مورد نیاز است. IDS دودویی و سه دسته نیاز به عملیات پس از پردازش دارد تا قبل از پاسخ بتواند جزئیات بیشتری ارائه کند. مزیت اصلی IDS مبتنی برکاوش مسیرهای متداول نسبت به یک IDS پردازش دسته ای سنتی، توانایی پردازش داده ها به عنوان یک جریان به جای جمع آوری دسته از داده ها برای آنالیز آفلاین است. پردازش جریان مقدار حافظه مورد نیاز برای آموزش و طبقه بندی را به حداقل می رساند و در نتیجه بهتر برای IDS در مقیاس یک سیستم قدرت مناسب است.

جدول V. دقت تشخیص برای حملات چهار روز صفر تصادفی با اعتبار ۱۰ ×

دوره	سناریوهای محروم	Z.D. Acc. (%)
1	Q3, Q11, Q18, Q22	76.3
2	Q2, Q8, Q12, Q23	67.3
3	Q6, Q11, Q16, Q17	50.5
4	Q1, Q5, Q8, Q10	73.3
5	Q1, Q9, Q19, Q21	91.8
6	Q5, Q13, Q20, Q23	64.7
7	Q5, Q10, Q15, Q16	63.8
8	Q12, Q13, Q19, Q24	70.7
9	Q2, Q7, Q9, Q17	76.3
10	Q9, Q10, Q16, Q19	99.8

www.nn4e.com - گروه نامیرا

VII. نتیجه گیری

IDS مسیر متداول مبتنی بر کاوش نظارت stateful به یک سیستم حفاظت انتقال از راه دور الکتریکی با اعمال نفوذ تلفیقی داده synchrophasor و اطلاعات از ایستگاه تقویت، لگاریتم های امنیت شبکه، و لگاریتم های EMS، فراهم می کند. IDS با استفاده از یک الگوریتم کاوش مسیر متداول آموزش داده می شود. مسیرهای متداول، علائم و خصوصیات ترکیبی هستند که الگوهای رفتاری سیستم در ارتباط با وقایع و رویدادهای سیستم قدرت را توصیف می کند. الگوریتم یک روش تجزیه و تحلیل داده در حوزه زمان برای غلبه بر فرد گذرای حاضر در اندازه گیری ها را فراهم می کند. این کار با کاوش حالات مشترک از یک گروه از مسیرهای مشاهده شده، انجام می شود. مسیر متداول برای توصیف پاسخ سیستم به اختلالات سیستم قدرت، اقدامات کنترلی، و حملات سایبری، مورد استفاده قرار می گیرند. IDS حالت سیستم نظارت شده را نسبت به مسیر متداول پیمایش شده، تطابق می دهد تا تصمیم هایی برای طبقه بندی بگیرد. طبقه بندی برای هر سناریو آموزش دیده خاص است و به سادگی نشانه ای از فعالیت های طبیعی یا غیر طبیعی نیست.

در این مقاله، IDS برای ارزیابی یک سیستم انتقال دو خط سه اتوبوس که پیاده سازی یک طرح حفاظت از راه دور دو منطقه را انجام می دهد، آموزش دیده بود. بیست پنج سناریو متشکل از گسل های stockticker SLG، اقدامات کنترل، و حملات سایبری در یک بستر آزمون سخت افزار در حلقه اجرا شد. سناریوها در

www.nn4e.com – nn4e@aol.com – 09360703858 – گروه نامیرا

یک حلقه ۱۰۰۰۰ مرتبه، با پارامترهای سیستم تصادفی اجرا شد تا یک مجموعه داده برای آموزش و ارزیابی IDS ایجاد کند. IDS به درستی ۹۰,۴٪ از موارد سناریوی تست شده، طبقه بندی شده است. ارزیابی نیز شامل یک ده برابر اعتبار متقاطع برای ارزیابی دقت تشخیص روز صفر سناریوی حمله است. متوسط دقت تشخیص برای روز صفر سناریوی حمله ۷۳,۴۳ درصد بود. IDS مسیر متداول مبتنی بر کاوش بهتر از الگوریتم های یادگیری ماشین سنتی عمل کردند و برای حجم بالایی از اطلاعات موجود در سیستم های قدرت مناسب تر بودند.

در حال حاضر، IDS مسیر متداول مبتنی بر کاوش، از لگاریتم های دریافتی داده ها، مسیرهای متداول ساخته اند. گرفتن چنین لگاریتم داده ای برای سیستم های واقعی دشوار است. به این ترتیب، در کار های آینده مورد نیاز است تا میزان تعداد حالات نمونه سناریوی گرفته شده برای آموزش لگاریتم، را محدود کنیم. IDS توسط مجموعه داده های آزمون بررسی شده آفلاین، مورد آزمایش قرار گرفت. در کار آینده نیاز است برای انجام طبقه بندی زمان واقعی از ورودی های سیستم زنده، به روز رسانی IDS انجام شود و طبقه بندی کننده با یک چارچوب کنترل تطبیقی هوشمند [۲۳] برای دستیابی به افزایش اتوماسیون در سیستم های قدرت، ترکیب شود.

www.nn4e.com – nn4e@aol.com – 09360703858 – گروه نامیرا