

باسمه تعالی

گزارش پروژه درس امنیت در اینترنت اشیاء

**دانشگاه صنعتي شريف**

عنوان:

تحلیل سامانه‌های تشخیص نفوذ مبتنی بر یادگیری عمیق در شبکه‌های اینترنت اشیاء

انجام‌دهنده:

پوریا دادخواه  
401201381

# زمستان 1402

چکیده

در شبکه‌ها و سیستم‌های کامپیوتری یکی از مهم‌ترین چالش‌های قابل تامل حفظ امنیت در برابر مهاجمان داخلی و خارجی با اهداف گوناگون است که برای دستیابی به این هدف از ابزار‌ها و الگوریتم‌های مختلف در سامانه‌های تشخیص نفوذ، فایروال‌ها و سایر ابزار‌ها استفاده می‌کنیم. یکی از این ابزار‌ها در سال‌های اخیر با گسترش روش‌های بهینه‌سازی و پردازش‌های موازی، الگوریتم‌های یادگیری عمیق و پردازش داده‌های ورودی هستند که بدون دخالت انسان توانایی تصمیم‌گیری و دسته‌بندی داده‌های ورودی یک سیستم‌ را برای ما فراهم می‌کنند. در این پروژه قصد داریم با معرفی حملات متخاصم به تحلیل این سامانه‌‌های بپردازیم.

# فهرست مطالب

[1 مقدمه 1](#_Toc99753053)

[2 تعاریف و مفاهیم اولیه 3](#_Toc99753056)

[3 مروری بر سامانه‌های تشخیص نفوذ موجود 6](#_Toc99753063)

[4 اعمال حمله متخاصم 12](#_Toc99753065)

[5 جمع‌بندی و مراجع 17](#_Toc99753067)

# فهرست اشکال

[شکل ‏1‑1 توضیحات مربوط به شکل نمونه [1] 3](#_Toc99753047)

# فهرست جداول

[جدول ‏1‑1 توضیحات مربوط به جدول نمونه [1] 3](#_Toc99753044)

# مقدمه

فصل 1

در سال‌های اخیر و با پیشرفت شبکه‌های کامپیوتری و کاربرد‌های روزافزون آن‌ها برای ارتباط انسان‌ها و ماشین‌ها با یکدیگر، اهمیت حفظ امنیت داده‌‌ها و درستی کارکرد این شبکه‌ها در برابر وجود مهاجمان مختلف با انگیزه‌های گوناگون اعم از دسترسی غیرمجاز، تغییر داده‌ها، تزریق بدافزار‌، ازکارانداختن شبکه و ... روز‌به‌روز جایگاه‌مهم‌تری به خود می‌گیرد.

این مسئله از قدیم مورد توجه بوده و الگوریتم‌های مختلفی برای جلوگیری از ورود و اعمال مهاجمان در شبکه‌ها معرفی و استفاده شده‌اند. هرچند همواره میزان دقت تشخیص و کارایی آن‌ها مورد بحث بوده و بسته به شبکه هدف و میزان ترافیک بسته‌ها و کاربرد شبکه، رویکرد‌های گوناگونی اتخاذ می‌شوند. یکی از این شبکه‌ها در سالیان اخیر شبکه اینترنت اشیاء است که در آن با حجم زیادی از داده‌ها و منابع محدودی در دستگاه‌ها طرف هستیم که آن را طعمه مناسبی برای اهداف متخاصم می‌کند و باید روشی ارائه دهیم که با مصرف بهینه منابع، در زمان معقول حمله‌های مختلف را شناسایی و از وقوع آن جلوگیری کند.[1]

در این مقاله یک مدل یادگیری عمیق مبتنی بر گراف ارائه می‌دهیم که با استفاده از آن بتوان ساختار شبکه را در کنار خصوصیات آماری بسته‌های شبکه به مدل آموزش داد تا با حفظ رکورد تعاملات هر جفت دستگاه در شبکه بتواند علاوه بر تشخیص وقوع ناهنجاری، نوع حمله را نیز تفکیک کند.[2] به‌این منظور در بازه زمانی مدنظر، تصاویر لحظه‌ای متوالی از ساختار شبکه و جریان‌های بسته‌های آن‌ها می‌گیریم و هر یک را پس از تبدیل به گرافی معادلی که رئوس آن متناظر با ترافیک شبکه هستند، به مدل یادگیری عمیق گرافی می‌دهیم تا برچسبی متناظر-عادی و یا نوع حمله- را به هر راس نسبت دهد و درنهایت با اعمال یک مدل یادگیری بازگشتی روی این تصاویر لحظه‌ای، اثر تعاملات زمانی جفت دستگاه‌های شبکه را اعمال کرده و ترافیک‌های موجود را به یکی از دسته‌بندی‌های موجود نظیر می‌کنیم. [2]

در سال‌های اخیر و با پیشرفت و معرفی مدل‌های جدید یادگیری ماشین، الگوریتم‌های بهتری نسبت به الگوریتم‌های سابق، که تنها قاعده محور بودند و دقت و عملکرد آن‌ها وابستگی زیادی به داده‌ ورودی و انسان داشته و در برابر حملات جدید کارایی مناسبی نداشتند، به کار گرفته شدند ولی بنا به مدل به کار گرفته شده یا در حفظ اطلاعات زمانی ناتوان بوده و تنها ویژگی‌های آماری بسته‌ها را درنظر می‌گیرند [3] و یا یک جریان زمانی را وارد مدل یادگیری می‌کنند و از حفظ تعامل یک جفت دستگاه باز می‌مانند.[4] در این مدل با استفاده از اعمال یک گراف به صورت پویا در طول زمان، علاوه بر ویژگی‌های آماری ترافیک، اثر ساختار و تغییرات هر جفت دستگاه‌ها را نیز درنظر می‌گیریم که باعث می‌شود برخلاف مدل‌های قبلی که در تشخیص نوع حمله عملکرد نامطلوبی دارند، بتوانیم در دسته‌بندی چندگانه نیز به‌صورت بلادرنگ عملکرد خوبی داشته باشیم. همچنین تفاوت دیگر این مدل، نیمه نظارت‌شده بودن آن است که باعث می‌شود در داده‌های بزرگ با صرف هزینه و زمان بسیار کمتر و کاهش اثر خطای جانبی، دقت مشابهی داشته‌باشیم.[5]

در ادامه مقاله، ابتدا به چند تعریف و مفهوم اولیه که برای روش‌های ارائه‌شده نیاز داریم می‌پردازیم سپس در بخش بعد ایده‌های معروف و مشابه قبلی برای تشخیص نفوذ را معرفی می‌کنیم و در بخش چهارم روش خود را بیان می‌کنیم و در بخش پنج با تبیین نحوه پیش‌پردازش و نحوه پیاده‌سازی، نتایج روش را با روش‌های قبلی مقایسه می‌کنیم و در آخر به جمع‌بندی کار‌ انجام شده می‌پردازیم.

# تعاریف و مفاهیم اولیه

فصل 2

1. **تعاریف و مفاهیم اولیه**

**2-1 شبکه عصبی گرافی[[1]](#footnote-1)**

این شبکه عصبی، نوعی مدل یادگیری عمیق است که به عنوان ورودی یک گراف را گرفته که هر راس آن حاوی ویژگی‌های مشخص است و در طول یادگیری، با نوعی اثرگذاری ویژگی‌های رئوس همسایه، وضعیت هر راس را به‌روز کرده و به لایه بعدی انتقال می‌دهد و در لایه آخر بسته به مقادیر نهایی ویژگی‌ها، به یکی از n دسته مدنظر ورودی خود را برچسب می‌زند.[6] این نوع یادگیری باعث می‌شود که علاوه بر داده‌های منظم ، بتوانیم داده‌های نامنظم را نیز دسته‌بندی کنیم و علاوه بر توزیع ویژگی‌های هر راس، تاثیر ساختار [[2]](#footnote-2)ارتباطی ورودی‌ها را نیز منظور کنیم.

**2-1-1 شبکه عصبی گرافی پیچشی[[3]](#footnote-3)**

یک روش ممکن برای اعمال اثر ویژگی‌های رئوس همسایه بر مقدار ویژگی‌ به‌روز‌شده راس فعلی، انجام یک میانگین‌گیری وزن‌دار بر روی مقادیر ویژگی‌ متناظر رئوس همسایه است ( که با ماتریس همانی نیز جمع می‌شود که مقدار ویژگی خود راس نیز اثر مستقیم داشته‌باشد) تا پس از عبور دادن از یک تابع فعالسازی، به لایه بعدی انتقال یابد، به این نوع شبکه گرافی، شبکه گرافی پیچشی گفته می‌شود. مدلی که در این مقاله می‌خواهیم استفاده کنیم، بهبود یافته این روش است.[3]

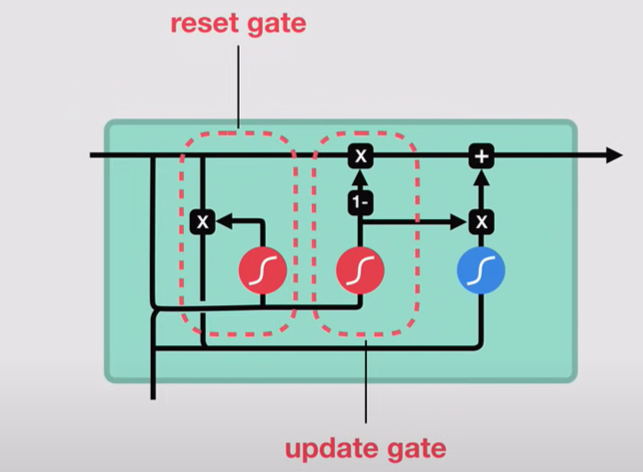
**2-2 گراف خطی[[4]](#footnote-4)**

همان‌گونه که در قسمت قبل اشاره شد، تمرکز یک شبکه گرافی، دسته‌بندی بر اساس رئوس گراف و ویژگی‌های آن‌هاست اما هدفی که ما در شبکه خود داریم، مدل کردن ساختار شبکه کامپیوتری با یک گراف است که مولفه‌های اصلی تشخیص نفوذ ترافیک‌های شبکه‌اند که معادل با یال‌های بین رئوس یا همان سیستم‌های شبکه هستند. در همین راستا از گراف اصلی، گرافی معادل به نام گراف خطی به‌دست‌می‌آوریم که یال‌های گراف اصلی، رئوس گراف خطی بوده و هر راس جدید، همسایه‌هایی دارد که در گراف قبلی، رئوس مشترک آن یال بوده‌اند[7]؛ به‌این صورت می‌توانیم ورودی مدل یادگیری خود را روی گراف جدید پیاده‌ کرده و بر اساس ویژگی‌های بسته‌های ترافیک مدل را به‌روز‌رسانی کنیم.[2]

**2-4 واحد بازگشتی دروازه‌ای[[5]](#footnote-5)**

مدل‌های یادگیری عمیق بازگشتی برای مقاصدی استفاده می‌شوند که ورودی مدل به صورت دنباله بوده و از ورودی قبلی و کنونی، قصد پیش‌بینی حالت بعدی را داریم و در دسته مناسب قرار دهیم؛ به این منظور ساده‌ترین مدل استفاده از یک بازخورد از خروجی مدل با ضریب مشخص در کنار پارامترهای عادی یک شبکه عصبی ANN است[[6]](#footnote-6)[1]. مشکل اصلی این روش، محو شدن/ انفجار اثر ورودی‌های اولیه در لایه‌های زیاد است چراکه درصورت کوچکتر/بزرگتر بودن ضریب اثرگذاری، در توان‌های بالا به صفر/بی‌نهایت میل می‌کند،[[7]](#footnote-7) به همین جهت مدل‌های بهبودیافته‌تری ارائه‌شدند که از وقوع این مشکل جلوگیری کند و دقت بالاتری داشته باشد.[4]

یک مدل معرفی شده، واحد بازگشتی دروازه‌ای است که ایده این روش به این صورت است که از بین ورودی‌های مدل، آن‌هایی را که ذخیره و در ورودی‌های بعدی اثر بدهد که حائز اهمیت بوده و سایر داده‌های حاشیه‌ای را محو می‌کند.[8] طرز کار کلی این مدل را می‌توان در شکل1 مشاهده کرد که همان‌طور که مشاهده می‌کنیم در هر واحد دو ورودی سیگنال فعلی و حالت قبلی را دریافت می‌کند و بسته به تابع فعال‌سازی درون دروازه، در صورت براورده شدن شرط به دریچه به‌روزرسانی وارد می‌شود و حالت بعدی را تولید می‌کند و درغیر این صورت به قسمت بازیابی رفته و اثرش ازبین می‌رود. ما از این مدل برای به دست‌آوردن دسته‌بندی نهایی بر اساس خروجی‌های مدل‌های تصاویر لحظه‌ای [[8]](#footnote-8)که از شبکه در طول زمان گرفته‌ایم استفاده می‌کنیم.[2]



شکل1: واحد بازگشتی دروازه‌ای

# بدنه اصلی (عنوان مناسب قرار یببگیرد)

فصل 3

در این قسمت برخی از مهم‌ترین کارهای اخیر انجام‌شده در این حوزه را معرفی می‌کنیم که به صورت کلی در چهار دسته پیچشی، بازگشتی، گرافی و ترکیبی[[9]](#footnote-9) تقسیم می‌کنیم.

**3-1 شبکه‌های بازگشتی**

همان‌‌طور که در قسمت قبل با کارکرد شبکه‌های بازگشتی آشنا شدیم، آن‌ها گزینه ایده‌آلی برای استفاده‌ در روش‌های تشخیص نفوذ در شبکه می‌باشند چراکه در این کاربرد نیز با جریانی از بسته‌ها طرفیم که می‌خواهیم با دریافت بسته‌هایی که در یک بازه زمانی دریافت می‌شوند پیش‌بینی کنیم که آیا این جریان نرمال است یا قصد مخربی را درپی دارد؛ الگوریتم‌های متعددی از این دسته معرفی شده‌اند که یکی از الگوریتم‌هایی که در مقایسه با بقیه دقت عملکرد و کارایی زمانی بهتری دارد در [4]آمده‌است.

در این روش، برای عملکرد بهتر مدل بازگشتی اصلی، اولا برای پردازش داده‌ها، از یک روش دسته‌بندی داده‌های خام [[10]](#footnote-10) استفاده‌ می‌کند و سپس روی هر دسته که ویژگی‌های نزدیک به‌هم دارند، یک مدل جداگانه با پارامترهای مخصوص به‌خود را اجرا می‌کند در این صورت کارایی زمانی الگوریتم بهبود وی‍ژه‌ای می‌یابد چراکه مدل به صورت محلی به نقطه‌های اکسترمم خود می‌رسد. ثانیا بعد از اجرای مدل بازگشتی روی هر دسته، از روش بهینه‌سازی محاسبات تکاملی که در [9] توضیح‌داده‌شده استفاده می‌کنیم تا فراپارمتر‌های هر مدل را به‌روزرسانی کنیم که نتیجه پیش‌بینی به حداقل خطا برسد.

**3-2 شبکه‌های پیچشی**

یکی دیگر از شبکه‌های پرکاربرد عصبی هستند که با ادغام ویژگی‌های داده همسایه با یکدیگر، ورودی لایه‌های بعدی خود را می‌سازند و یکی از بهترین روش‌ها برای دسته‌بندی داده‌هایی با ساختار منظم مانند تصاویر و ماتریس‌ها دارند. کاربردی که از این شبکه‌ها در روش‌های تشخیص نفوذ می‌شود اکثرا برای دسته‌بندی ترافیک‌ها براساس خصوصیات آماری از جمله تعداد بایت‌های ورودی، طول هر بسته، پروتکل و ... استفاده می‌شود که از بهترین روش‌های پیاده‌سازی این روش که دقت بهتری نسبت به کارهای مشابه این دسته دارد می‌توان به [3] اشاره کرد.هرچند مشکل اصلی این الگوریتم‌ها در این است که اولا اثر زمانی بسته‌ها را درنظر نمی‌گیرد و ثانیا ساختار شبکه و عوامل نامنظم شبکه را در ورودی اعمال نمی‌کند.

**3-3 شبکه‌های گرافی**

پس از معرفی شبکه‌های گرافی و توانایی آن‌ها برای اعمال ساختار شبکه علاوه بر خصوصیات آماری بسته‌ها، الگوریتم‌های متفاوتی برای استفاده از آن‌ها برای کاربرد تشخیص نفوذ ارائه شدند و روش ما در این عرصه اولین مدل نیست، هرچند روش‌های پیشین از نقاط ضعفی برخوردار بودند که اکنون به چند نمونه آن‌ها اشاره می‌کنیم.

در [6] نوعی گراف به نام گراف شی امنیتی را برای پیوند انواع مختلف رویدادهای امنیتی پیشنهاد کرد و از رمزگذارهای خودکار برای تشخیص ناهنجاری‌های گراف استفاده کرد. علاوه بر این،[10] یک روش تشخیص نفوذ خودی را با استفاده از تعبیه‌های گراف برای رمزگذاری توصیفات معنایی ساختار سازمانی پیشنهاد کرد. با این حال، [10] و [6] توپولوژی های واقعی شبکه را در نظر نگرفتند و داده های نمودار تنها به عنوان ابزاری برای ادغام اطلاعات ناهمگن در نظر گرفته شدند.

در همین حال، مطالعات کمی برای بهره برداری از اطلاعات مکانی شبکه های واقعی به درجات مختلف آغاز شده است.[11] یک مدل گراف DL را پیشنهاد کرد که شبکه های ربات را از طریق تجزیه و تحلیل مبتنی بر نمودار رفتارهای ترافیک شبکه شناسایی می کند. با این وجود، این مدل فقط اطلاعات توپولوژیکی شبکه هدف را بدون ویژگی‌های گره یا لبه می‌آموزد.

تفاوت بین رویکردهای تشخیص نفوذ قبلی و روش پیشنهادی در این مقاله در این است که روش حاضر توپولوژی شبکه اضافی و اطلاعات همبستگی مکانی-زمانی را برای تقویت توانایی خود در تشخیص نفوذ شبکه و رفتار مخرب اتخاذ می‌کند.[2]

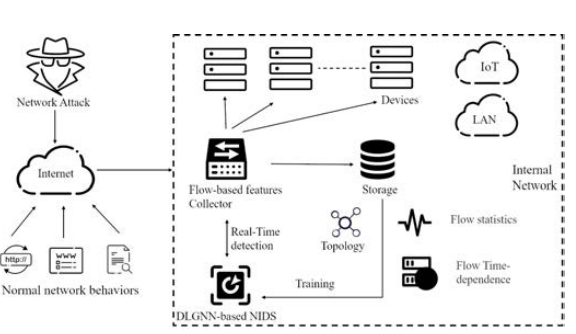
**3-4 روش‌های ترکیبی[[11]](#footnote-11)**

در اکثر کاربردهای واقعی، به دلیل اهمیت حوزه مدنظر نمی‌توان تنها به مدل‌های خودکار یادگیری ماشین اعتماد کامل داشت و حتی درصد خیلی کوچکی از خطا نیز ممکن هزینه بسیار بالایی درپی داشته باشد؛ در همین راستا در این موارد از ترکیب یک سیستم تشخیص نفوذ قاعده محور که انواع حملات به صورت دستی به‌وسیله انسان وارد شده‌اند در کنار یک مدل هوش مصنوعی پیاده‌ می‌شوند که حملات معمول را به صورت قطعی شناسایی و مسدود کند ( توسط بخش قاعده محور) و حملات جدید نیز توسط هوش مصنوعی شناخته و برای بهبود امنیت سیستم به قواعد اضافه شوند.در [12] یک سیستم ترکیبی کارا را برای شبکه‌های مخابرات بیسیم معرفی می‌کند.

**3-5 شبکه گرافی پیشنهادی**

اکنون با دانستن پیش‌نیاز‌های موردنیاز، روشی ارائه می‌دهیم که اهداف گفته‌شده را برآورده کند و یک مدل تشخیص نفوذ چندکلاسی را با دقت بالا و در زمان کارایی پیاده کند.

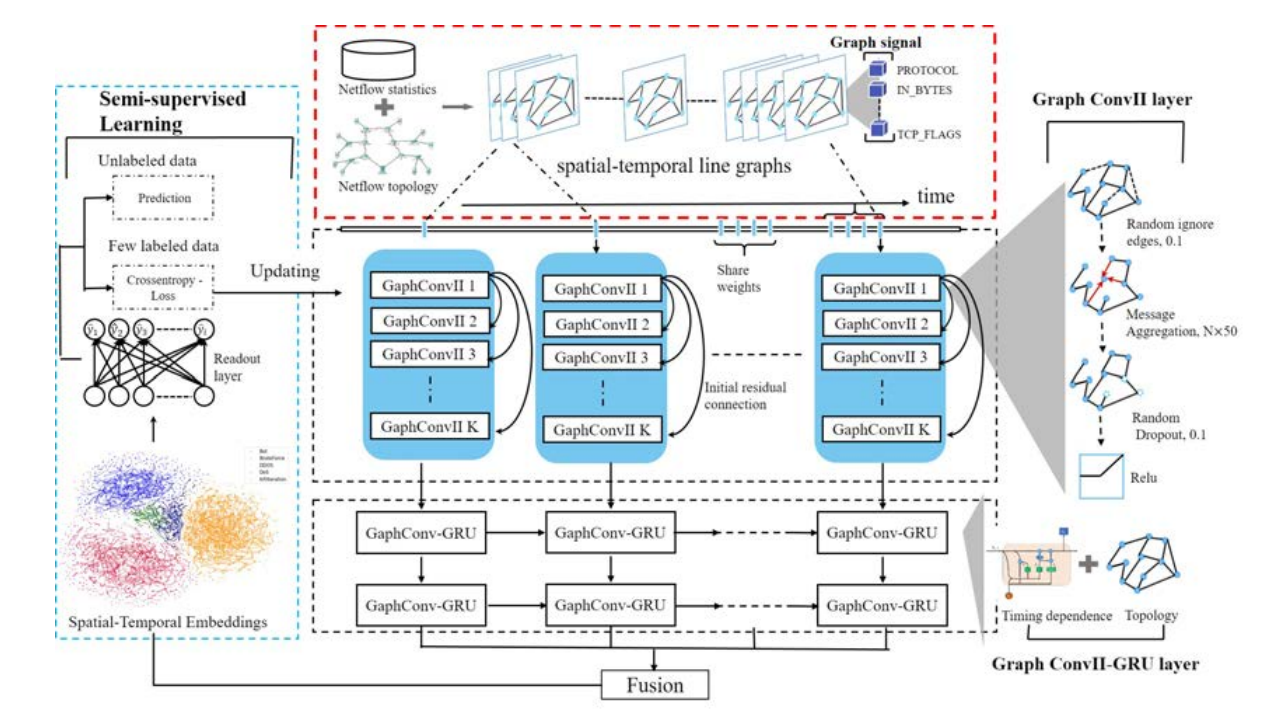
فرض تعریف مدل به این صورت است که تمام ترافیک شبکه محلی که قصد بررسی امنیت آن را داریم از یک دستگاه ناظر قابل مشاهده‌ است و ترافیک ورودی به سیستم، از جمله ترافیک‌‌های مهاجمان، از آن می‌گذرد و از آن‌جا الگوریتم خود را اجرا می‌کنیم و در صورت نیاز ترافیک خاصی را مسدود می‌کنیم. در شکل (1) ساختار کلی شبکه محلی و مراحل کار نشان داده شده‌اند.[2]



شکل1- ساختار شبکه هدف

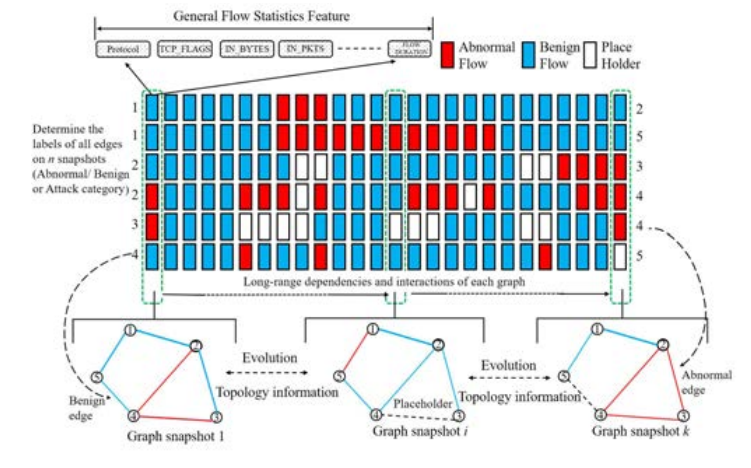
روند بررسی و تشخیص جریانی از بسته‌ها و تخصیص برچسب مناسب آن (نوع حمله و یا نرمال بودن بسته) در شکل (2) خلاصه شده است که مراحل آن در ادامه شرح داده می‌شود.

ابتدا در یک بازه زمانی مشخص، k تصویر لحظه‌ای در فاصله‌های یکسان از شبکه می‌گیریم که حاوی ساختار شبکه و اطلاعات آماری بسته‌های آن لحظه می‌باشد. هدف این است که ابتدا هرکدام از این تصاویر را که حاوی ویژگی‌های آماری بسته‌ها و ساختار ارتباطی نقاط شبکه‌هستند را دسته‌بندی کنیم و برچسب بزنیم و سپس با استفاده از یک شبکه بازگشتی دروازه‌ای که این خروجی‌های لحظه‌ای را به عنوان ورودی می‌گیرد برچسب‌نهایی هر جریان بسته را مشخص کنیم.[2]



شکل2 – مراحل الگوریتم تشخیص نفوذ

در مرحله اول برای هر تصویر لحظه‌ای i که ، ساختار شبکه را به گراف معادل آن (G) تبدیل می‌کنیم که رئوس، دستگاه‌های میزبان شبکه و یال‌های آن، ترافیک‌های ارسالی موجود در آن لحظه بین جفت دستگاه‌ها می‌باشد. سپس برای ایجاد ورودی مناسب به مدل عصبی گرافی پیچشی، از گراف فعلی، گراف خطی معادل L(G) را تشکیل می‌دهیم که رئوس آن بسته‌های ترافیک شبکه و ویژگی‌های آماری آن‌ها هستند و این گراف را به عنوان ورودی به شبکه عصبی GCNII می‌دهیم که نوعی شبکه گرافی پیچشی است تا به هر ترافیک برچسب مناسب را اختصاص دهد (نرمال یا غیر نرمال و نوع حمله). شکل 3 روند اجرای این مرحله را نمایش می‌دهد.[2]



شکل3 – دسته‌بندی بسته‌های یک لحظه از شبکه

درهرلایه از شبکه عصبی گرافی باید مجموعه ویژگی‌های هر راس، با استفاده از حالت لایه فعلی خود راس و رئوس مجاور آن به‌روزرسانی شود تا ویژگی‌های لایه بعدی به دست‌آیند. رابطه کلی این فرایند را برای یک شبکه پیچشی عادی در معادله (1) می‌بینیم که اثر رئوس همسایه و خود راس در یک جمع وزن دار ( معادل با ضرب در ماتریس ضرایب W ) و اعمال بایاس مدل کرده و خروجی را از تابع فعالسازی که طی فرایند آموزش برای تعیین دسته‌ها تعیین می‌شود عبور داده می‌شود.[6]

رابطه 1- به‌روز رسانی‌لایه‌های شبکه گرافی

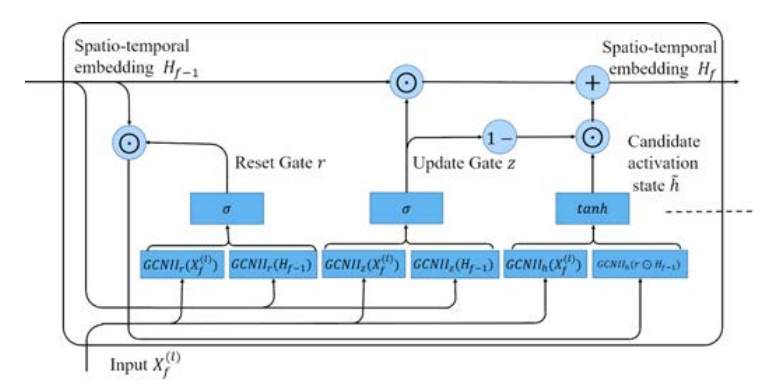
در شبکه GCNII که بهبود یافته‌شده GCN است، مشکل oversmoothing را برطرف می‌کند؛ در شبکه پیچشی عادی به دلیل نحوه به‌روز‌رسانی رئوس، پس از تعداد لایه‌های زیاد، مقادیر همه رئوس به سمت یکدیگر همگرا می‌شوند چراکه ویژگی‌های همه رئوس، با ضرایب تقریبی یکسانی روی یکدیگر اثر می‌گذارند. در مدل پیچشی عمیق GCNIIبا تعریف دو مفوهم اتصالات باقی‌مانده‌ای و نگاشت هویت، اثر ویژگی‌های ذاتی هر راس را در طول آموزش حفظ کرده و در عین حال، ویژگی‌های رئوس مجاور را با ضریب مناسب جمع می‌کند. در رابطه (2) نحوه به‌روز‌رسانی ویژگی‌های لایه بعد هر راس را مشاهده می‌کنیم.[2]

رابطه2- به روزرسانی لایه‌های شبکه پیچشی عمیق

که در این رابطه ترم اول اتصالات باقی‌مانده‌ای و جمع ضریب‌دار ویژگی سایر رئوس است که در آن A ماتریس مجاورت گراف و ماتریس نرمالایز شده ضرایب است که و ماتریس قطری متناظر و فراپارامتری است که در طول آموزش تعیین می‌گردد که با چه ضریبی این ترم را اثر دهیم و ترم دوم مربوط به حفظ ویژگی ذاتی رئوس بوده که در آن فراپارمتر میزان تاثیر ماتریس همانی ویژگی‌ها را کنترل می‌کند.

پس از دسته‌بندی‌ تصاویر لحظه‌ای، از مدل یادگیری بازگشتی GRU که تعریف شد، استفاده می‌کنیم که به عنوان ورودی، جریان خروجی‌های K دسته‌بندی به‌دست آمده را داده و یک برچسب نهایی برای هر یال از گراف به دست می‌آوریم. رابطه به‌روزرسانی حالت لحظه f از ماتریس مجاورت ساختار شبکه در آن لحظه ()، حالت گراف لحظه قبلی () و بردار ویژگی‌های آماری بسته‌ها در آن لحظه() را در رابطه3 می‌بینیم و ساختار تابع GRU در شکل 4 آمده‌است.[2]

رابطه3- حالت بعدی ساختار شبکه با اثر بازخورد فعلی



شکل4 – واحد بازگشتی به‌روزرسانی زمانی ویژگی‌ها

که مولفه‌های (z) update، (r) reset، فعالسازی () و خروجی واحد () به صورت زیر از ورودی‌ها به‌دست می‌آیند که GNII اپراتور شبکه گرافی مربوطه است.[2]

رابطه 5- روابط واحد بازگشتی دروازه‌ای

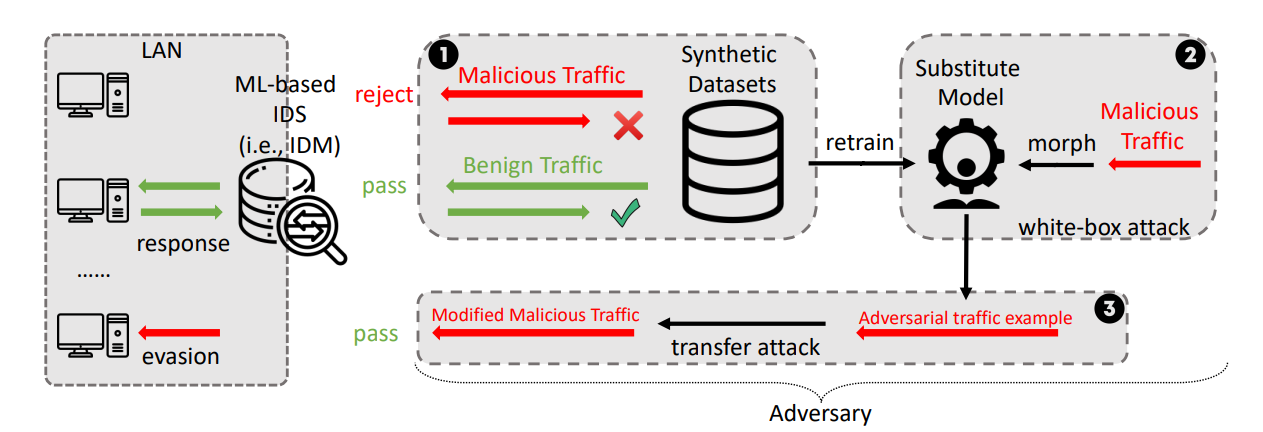
در آخر پس از به‌دست آوردن مدل کلی، به‌صورت نیمه‌نظارت شده، کافی است تنها 30 درصد داده‌های آموزش را برچسب بزنیم و مدل را آموزش دهیم و فراپارمتر‌های مراحل مختلف را تعیین کنیم که خروجی مدل (y) کمترین اختلاف را با خروجی مورد انتظار داشته باشد ()و برای مرحله تست آماده شود که نتایج آن را در قسمت بعدی می‌بینیم. برای مدل کردن این رابطه، خروجی را به ازای لحظه‌های متوالی (j) روی تمام ویژگی‌ها (i) جمع می‌کنیم.[2]

رابطه6- رابطه کمینه کردن خطای خروجی

# حمله متخاصم به IDS

فصل 4

در این قسمت نتایج یک حمله جعبه سفید با اجرای روش‌های تولی داده متخاصم متفاوت را عنوان کرده و نقاط ضعف آن را بیان می‌کنیم:



**مدل نهایی حمله**

**انواع روش‌های تولید داده متخاصم**

**۱**. \*\*روش علامت گرادیان سریع (FGSM

- \*\*مرور:\*\* FGSM یکی از ساده‌ترین و اولین روش‌ها برای تولید نمونه‌های تهاجمی است. این ورودی را با یک مقدار کوچک در جهت علامت گرادیان تابع از دست دادن نسبت به ورودی تحریک می‌کند.

- \*\*روند:\*\* با توجه به یک تصویر ورودی x، نمونه تهاجمی x\_{adv} به صورت زیر تولید می‌شود:

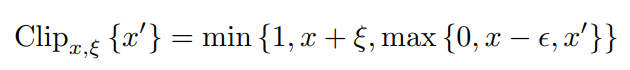


- \*\*مزایا:\*\* FGSM به لحاظ محاسباتی کارآمد و موثر برای تولید نمونه‌های تهاجمی با هزینه محاسباتی نسبتاً کم است.

۲. \*\*روش تکراری پایه (BIM

- \*\*مرور:\*\* BIM، همچنین به عنوان FGSM تکراری شناخته می‌شود، یک گسترش از FGSM است. به جای تولید نمونه تهاجمی در یک مرحله، BIM به صورت تکراری FGSM را با اندازه‌های گام کوچکتر برای یافتن اختلالی که منجر به طبقه‌بندی اشتباه می‌شود، اعمال می‌کند.

- \*\*روند:\*\* نمونه تهاجمی با اعمال به صورت تکراری FGSM با گام‌های کوچکتر تولید می‌شود تا زمانی که یک طبقه‌بندی اشتباه رخ دهد یا تا زمانی که تعدادی از تکرارها برسد.



- \*\*مزایا:\*\* BIM معمولاً به نمونه‌های تهاجمی قوی‌تری نسبت به FGSM منجر می‌شود زیرا فضای اختلال را به صورت کامل‌تری بررسی می‌کند.

۳. \*\*DeepFool:\*\*

- \*\*مرور:\*\* DeepFool یک روش حمله تکراری است که به حداقل رساندن اختلال مورد نیاز برای تغییر دسته‌بندی یک نمونه ورودی می‌پردازد.

- \*\*روند:\*\* DeepFool به صورت تکراری پارامترهای حاشیه‌ای تابع تصمیم را در ورودی فعلی خطی می‌کند و به سوی حداقل کردن اختلال با حرکت در جهت حاشیه تصمیم می‌پردازد.

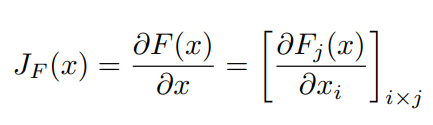


- \*\*مزایا:\*\* DeepFool در تولید نمونه‌های تهاجمی با اختلال‌های کوچک موفق است، اغلب با استفاده از تعداد کمتری از تکرارها موفقیت زیادی دارد.

۴. \*\*حمله نقشه سلیسیتی بر اساس یاکوبی (JSMA

- \*\*مرور:\*\* JSMA یک روش حمله هدفمند است که هدف آن بیشینه کردن اختلال در ویژگی‌های انتخاب شده ورودی با استفاده از ماتریس یاکوبی مدل است.

- \*\*روند:\*\* JSMA به صورت تکراری گرادیان‌های تابع از دست دادن نسبت به ویژگی‌های ورودی را محاسبه کرده و ویژگی‌هایی را که بیشترین مشارکت را در اشتباه کردن دارند انتخاب می‌کند. سپس آن ویژگی‌ها را به گونه‌ای تغییر می‌دهد که احتمال اشتباه طبقه‌بندی بیشینه شود.

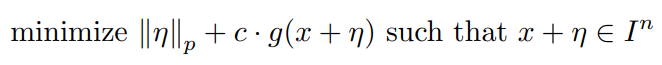


- \*\*مزایا:\*\* JSMA در تولید نمونه‌های تهاجمی که به طور خاص برای هدف‌های خاص طراحی شده‌اند مؤثر است، این ویژگی را برای حملات هدفمند مفید می‌کند.

۵. \*\*حمله کارلینی و واگنر (C&W

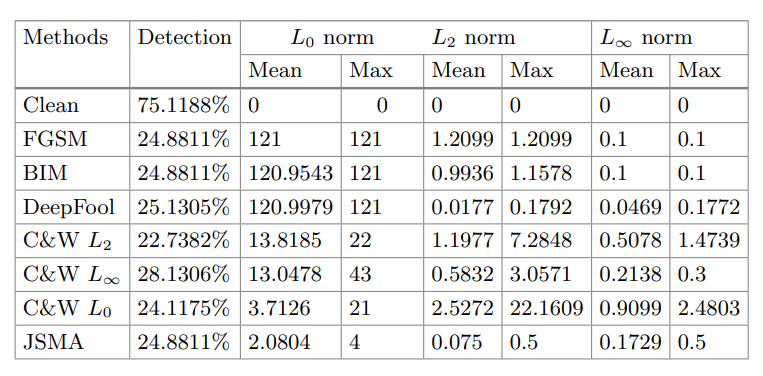
- \*\*مرور:\*\* حمله C&W یک روش حمله مبتنی بر بهینه‌سازی است که هدف آن یافتن حداقل اختلالی است که منجر به اشتباه کردن طبقه‌بندی می‌شود در حالی که همچنین به اندازه و قابل توجه بودن اختلال را در نظر می‌گیرد.

- \*\*روند:\*\* C&W تولید نمونه تهاجمی را به عنوان یک مسأله بهینه‌سازی فرموله می‌کند، جایی که هدف کمینه کردن ترکیبی از اندازه اختلال و یک عبارت است که انگیزه اختلال را برای پنهان بودن انسان‌ها تشویق می‌کند. معمولاً از تکنیک‌هایی مانند نزول گرادیان یا دیگر روش‌های بهینه‌سازی برای حل این مسأله استفاده می‌شود.

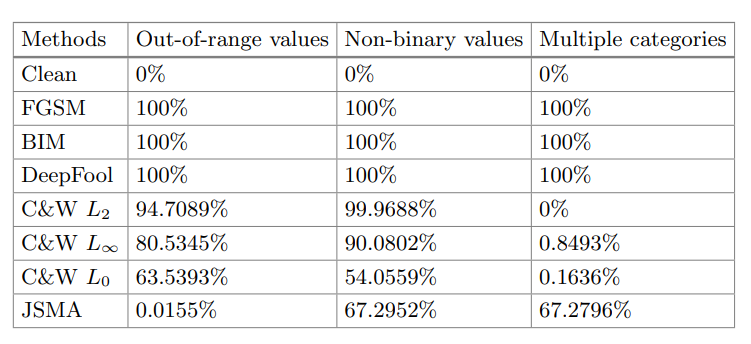


- \*\*مزایا:\*\* حمله C&W یکی از مؤثرترین و چندکاره‌ترین روش‌های حمله تهاجمی است که قادر به تولید نمونه‌های تهاجمی بسیار مؤثر هستند در حالی که اطمینان حاصل می‌کند که اختلال‌ها برای انسان‌ها قابل تشخیص نیستند.

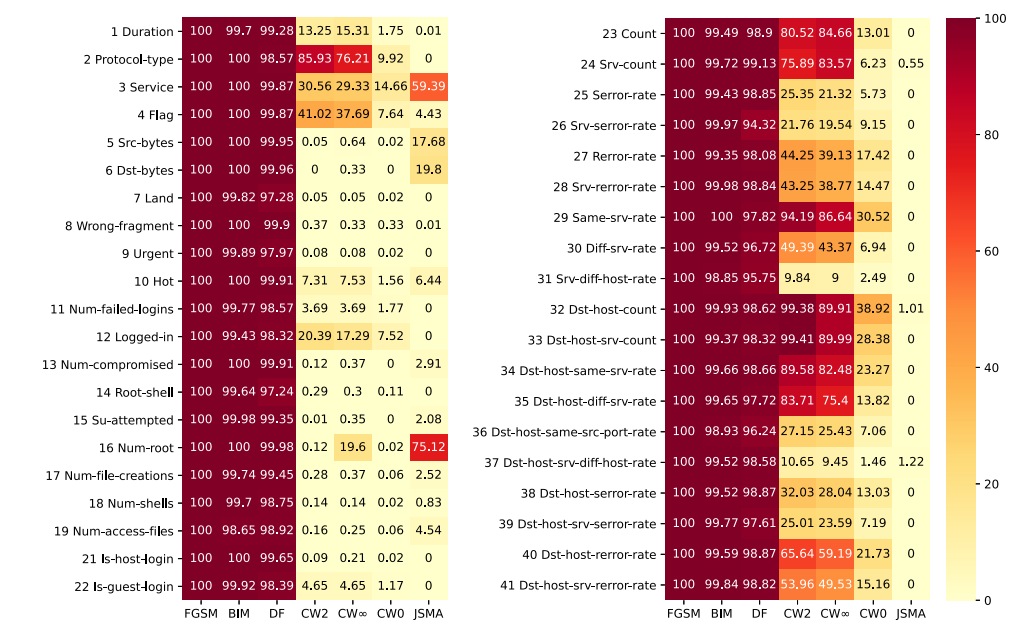
**نتایج پیاده‌سازی:**



تغییر تشخیص سامانه



درصد نامعتبر بودن بسته‌های تولید شده



میزان تغییرات اعمال شده و فاصله از بسته اولیه

همان‌گونه که انتظار داشتیم این نوع حمله نیاز به وابسته شدن به شبکه دارد که علاوه بر نتایج، معتبر نیز باشد.

# جمع‌بندی و مراجع

فصل 5

در این مقاله پس از معرفی مفهوم امنیت شبکه‌های کامپیوتری و اهمیت حفظ آن‌ها در برابر بسته‌های متخاصم به بررسی روش‌های نوین ارائه‌شده در این حوزه با استفاده‌ از روش‌های یادگیری ماشین و هوش مصنوعی پرداختیم و دیدیم که در اکثر این روش‌ها با این‌که نسبت به روش‌های قاعده محور دقت و کارایی قابل توجهی در تشخیص ناهنجاری داشتند ولی به دلیل ضعف در ساختار یادگیری مدل خود در تشخیص نوع حمله کارایی مطلوب را نداشتند. بنابراین ما یک مدل یادگیری عمیق مبتنی بر گراف ارائه دادیم که با استفاده از آن بتوان ساختار شبکه را در کنار خصوصیات آماری بسته‌های شبکه به مدل آموزش داد تا با حفظ رکورد تعاملات هر جفت دستگاه در شبکه بتواند علاوه بر تشخیص وقوع ناهنجاری، نوع حمله را نیز تفکیک کند.

به‌این منظور در بازه زمانی مدنظر، تصاویر لحظه‌ای متوالی از ساختار شبکه و جریان‌های بسته‌های آن‌ها گرفته و هر یک را پس از تبدیل به گرافی معادلی که رئوس آن متناظر با ترافیک شبکه هستند، به مدل یادگیری عمیق گرافی میدهیم تا برچسبی متناظر-عادی و یا نوع حمله- را به هر راس نسبت دهد و درنهایت با اعمال یک مدل یادگیری بازگشتی روی این تصاویر لحظه‌ای، اثر تعاملات زمانی جفت دستگاه‌های شبکه را اعمال کرده و ترافیک‌های موجود را به یکی از دسته‌بندی‌های موجود نظیر کردیم.

با بررسی پیاده‌سازی حمله متخاصم عادی و نتایج، ضعیف بودن کلی مدل را دیدیم و باید در ادامه حمله معتبری اعمال کنیم.

بهبود زمان آموزش مدل، ایمن بودن آن در برابر حملات مهاجمان به خود مدل یادگیری و شبیه‌سازی یک شبکه واقعی که در آن توان پردازش دستگاه‌ها نابرابر هستند از جمله مواردی است که در ادامه مسیر پژوهش می‌توان به آن پرداخت.

## مراجع

|  |  |
| --- | --- |
| [1] | [1] A. L. Buczak and E. Guven, “A survey of data mining and machine learning methods for cyber security intrusion detection,” IEEE Commun. Surveys Tuts., vol. 18, no. 2, pp. 1153–1176, 2nd Quart., 2016  [2] Application of a Dynamic Line Graph Neural Network for Intrusion Detection With Semisupervised Learning. Guanghan Duan. IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 18, 2023  [3] Applying convolutional neural network for network intrusion detection . R. Vinayakumar, K. P. Soman, and P. Poornachandran, in Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI), Sep. 2017, pp. 1222–1228  [4] Emergent Deep Learning for Anomaly Detection in Internet of Everything. Youcef Djenouri. IEEE INTERNET OF THINGS JOURNAL, VOL. 10, NO. 4, 15 FEBRUARY 2023  [5] F. Manessi, A. Rozza, and M. Manzo, “Dynamic graph convolutional networks,” Pattern Recognit., vol. 97, Jan. 2020, Art. no. 107000.  [6] L. Leichtnam, E. Totel, N. Prigent, and L. Mé, “Sec2Graph: Network attack detection based on novelty detection on graph structured data,” in Proc. Int. Conf. Detection Intrusions Malware, Vulnerability Assessment. Cham, Switzerland: Springer, 2020, pp. 238–258.  [7] T. S. Evans and R. Lambiotte, “Line graphs, link partitions, and overlapping communities,” Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top., vol. 80, no. 1, Jul. 2009, Art. no. 016105  [8] A. F. M. Agarap, “A neural network architecture combining gated recurrent unit (GRU) and support vector machine (SVM) for intrusion detection in network traffic data,” in Proc. 10th Int. Conf. Mach. Learn. Comput., Feb. 2018, pp. 26–30  [9] A. Karale, M. Lazarova, P. Koleva, and V. Poulkov, “A hybrid PSOMiLOF approach for outlier detection in streaming data,” in Proc. Int. Conf. Telecommun. Signal Process., 2020, pp. 474–479  [10] S. Zerhoudi, M. Granitzer, and M. Garchery, “Improving intrusion detection systems using zero-shot recognition via graph embeddings,” in Proc. IEEE 44th Annu. Comput., Softw., Appl. Conf. (COMPSAC), Jul. 2020, pp. 790–797.  [11] J. Zhou, Z. Xu, A. M. Rush, and M. Yu, “Automating BotNet detection with graph neural networks,” 2020, arXiv:2003.06344.  [12] An Efficient Hybrid IDS Deployment Architecture for Multi-Hop Clustered Wireless Sensor Networks. Da-Wen Huang. IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY, VOL. 17, 2022 |

# پیوست

فصل 6

در صورتی که مطالبی وجود دارد که فکر می‌کنید ذکر آن‌ها لازم است و نباید آن‌ها از قلم بیافتند، در اینجا مطرح کنید. در حجم و تعداد صفحات پیوست هیچ‌ محدودیتی وجود ندارد.

1. Graph Neural Network (GNN) [↑](#footnote-ref-1)
2. Topology [↑](#footnote-ref-2)
3. Graph convolutional neural network [↑](#footnote-ref-3)
4. Line Graph [↑](#footnote-ref-4)
5. Gated Recurrent Unit (GRU) [↑](#footnote-ref-5)
6. Recurrent neural network (RNN) [↑](#footnote-ref-6)
7. Vanishing/exploding Problem [↑](#footnote-ref-7)
8. Snapshot [↑](#footnote-ref-8)
9. Hybrid [↑](#footnote-ref-9)
10. K-means [↑](#footnote-ref-10)
11. Hybrid [↑](#footnote-ref-11)