



دانشکده مهندسی کامپیوتر
سمینار کارشناسی ارشد گرایش رایانش امن

عنوان:

ارائه رویکرد تطبیق پذیر با تنوع ترافیکی شبکه‌های پهن‌بند برای شناسایی حملات منع خدمت توزیع شده

An Adaptive Approach with Variety Characteristic of High-Bandwidth Networks for Distributed Denial of Service Attacks Detection

نگارش:

روح‌الله جهان‌افروز

۴۰۰۲۱۰۷۵۵

استاد راهنما:

دکتر رسول جلیلی

استاد ممتحن داخلی:

دکتر امیرحسین جهانگیر

بهمن ۱۴۰۱

چکیده:

بسیاری از برنامه‌های کاربردی^۱ امروزی از پروتکل‌های یکسان و مشترکی برای تبادل اطلاعات استفاده می‌کنند. برنامه‌های پیام‌رسان و مرورگرهای وب از بسته‌های مبتنی بر پروتکل اچ.تی.تی.پی/اس^۲ برای تبادل اطلاعات استفاده می‌کنند، با این تفاوت که در برنامه‌های پیام‌رسان با ارسال تعداد معینی از بسته‌های اچ.تی.تی.پی/اس در مقایسه با مرورگرهای اینترنتی، نرخ متفاوتی از بسته‌ها را در پاسخ دریافت خواهیم کرد. لذا با ظهور برنامه‌های کاربردی مختلف شاهد بروز تنوع ترافیکی بر روی پروتکل‌های مختلف و رفتارهای متفاوت در ترافیک شبکه هستیم. در شبکه‌های پهن‌بند^۳ با افزایش نرخ ترافیک و وجود تنوع پروتکلی زیاد، چالش‌های امنیتی نظیر تشخیص حملات منع خدمت^۴، که به دلیل سادگی در پیاده‌سازی و تاثیر بسیار مخرب یک تهدید جدی به حساب می‌آیند، افزایش پیدا کرده‌است. در دهه‌های گذشته محققان روش‌های شناسایی بسیاری را برای حملات منع خدمت توزیع‌شده^۵ پیشنهاد کرده‌اند. عدم تطبیق پذیری و مقیاس‌پذیری برای استفاده در شبکه‌های پهن‌بند، از متداول‌ترین مشکلات این روش‌ها هستند. لذا برای شناسایی صحیح حملات منع خدمت در شبکه‌های پهن‌بند نیاز به یک رویکردی است که شامل دو ویژگی پردازش جامع^۶ به معنای پردازش تمامی بسته‌ها و تطبیق‌پذیری^۷ به معنای قابلیت تطبیق‌پذیری با تنوع ترافیکی باشد. در این پژوهش ضمن بررسی کارهای مشابه صورت گرفته در این زمینه، قصد داریم رویکردی تطبیق‌پذیر با تنوع ترافیکی موجود در شبکه‌های پهن‌بند برای شناسایی حملات منع خدمت توزیع‌شده معرفی نماییم که ویژگی پردازش جامع ترافیک را نیز شامل شود. روش پیشنهادی جریان‌ها را بر اساس اینکه برای کدام کاربرد می‌باشند دسته‌بندی کرده و بر مبنای رفتار عادی ترافیک هر برنامه کاربردی، ترافیک‌های متخاصم را تشخیص می‌دهد. به دلیل اینکه از الگوریتم‌ها و داده‌ساختارهای فشرده و سبک با قابلیت جستجوی سریع استفاده می‌شود، سرعت بالا و استفاده بهینه از حافظه تضمین می‌شود. همچنین در روش پیشنهادی از ابزارهای تسریع عملیات پردازش بسته که در سالیان اخیر بسیار مورد استقبال قرار گرفته‌است، استفاده می‌شود و بدین صورت می‌توان سرعت پردازش بسته‌ها را تسریع بخشید که منجر به پردازش جامع تمامی بسته‌های ترافیک عبوری شبکه خواهد شد. در انتها کارایی روش ارائه شده در مقایسه با برخی دیگر از راهکارهای موجود و با در نظر گرفتن معیارهایی نظیر میزان استفاده از پردازشگر و حافظه، نرخ دور انداختن^۸ بسته‌ها، و میزان تاخیر در شناسایی حملات بررسی می‌شود.

کلیدواژه: حملات منع خدمت توزیع‌شده، شبکه‌های پهن‌بند، تطبیق‌پذیری با تنوع ترافیکی، سامانه‌های تشخیص نفوذ

۱. مقدمه

امروزه با افزایش حجم تبدلات داده‌ای در بستر اینترنت، برقراری ارتباطی امن و پایدار در سطح شبکه به یکی از چالش‌های اساسی پیش روی هر سازمانی تبدیل شده است. با توجه به رشد روزافزون کاربران شبکه‌های کامپیوتری، حجم درخواست‌های آن‌ها بزرگ‌تر و پیچیده‌تر می‌شود. از طرف دیگر اینترنت به جز جدایی‌ناپذیری در زندگی و تعاملات کاربران تبدیل شده و بحث دسترس‌پذیری آسان به خدمات بستر اینترنت بیش از پیش مورد توجه قرار می‌گیرد، بدین معنا که ارائه‌دهندگان خدمات ارتباطی^۹ موظف هستند خدمات خود را به‌صورت شبانه‌روزی و بدون اختلال و وقفه در اختیار کارخواهان^{۱۰} قرار دهند. در صورتی که این سازمان‌ها به هر دلیلی در ارائه خدمات خود دچار مشکل شوند و نتوانند به نحو مطلوب خدمات موردنظر را ارائه دهند، با چالش‌های جدی از قبیل از بین رفتن اعتماد مشتریان، خسارات سنگین مالی و از بین رفتن اعتبار سازمان مواجه می‌شوند.

حملات منع خدمت، دسته‌ای از حملات در شبکه هستند که با هدف از بین بردن دسترس‌پذیری شبکه سعی در ممانعت از ارائه و انجام یک خدمت^{۱۱} در شبکه دارند. حملات منع خدمت، پهنای باند یا ظرفیت لینک شبکه را مصرف کرده و یا باعث از کار افتادن و اختلال عملکرد در یک کارپذیر^{۱۲} یا هر دستگاه حیاتی دیگر در شبکه خواهند شد. گونه‌های مختلفی از این حملات وجود دارد که هرکدام به طریقی سعی می‌کنند دسترس‌پذیری شبکه را هدف قرار داده و یا با مصرف منابع کارپذیر، مانع از ارائه خدمت به‌صورت کامل و باکیفیت به کارخواهان و کاربران قانونی شوند. حملات منع خدمت توزیع‌شده یک‌گونه مخرب‌تر از حملات منع خدمت هستند که در آن‌ها حمله‌کننده^{۱۳} از طریق سیستم‌هایی که تحت کنترل خود می‌آورد، حمله را انجام می‌دهد. بدین ترتیب علاوه بر حجم ترافیک سنگین حملات و دشواری‌های تمییز قائل شدن بین ترافیک بالا در عین حال قانونی شبکه^{۱۴} و ترافیک حمله‌کننده، پیدا کردن فرد مهاجم اصلی نیز به‌مراتب دشوارتر می‌شود.

ازسویی دیگر امروزه با شبکه‌های پهن‌بندی مواجه هستیم که منجر به بالا رفتن نرخ‌گذر اطلاعات به میزان بیش از ۱۰۰ گیگابیت در ثانیه در بسیاری از تجهیزات شبکه شده است. برای شناسایی مهاجمین در چنین شرایطی نیاز به راهکاری است که با سرعت بالایی بتواند تمامی بسته‌ها را بررسی کند. همچنین به دلیل ظهور پروتکل‌ها و برنامه‌های کاربردی مختلف با حجم زیادی از داده‌ها و تنوع زیادی از پروتکل‌ها مواجه هستیم. لذا چالش بعدی تطبیق معیار تشخیص حملات با توجه به کاربرد ترافیک می‌باشد. با توجه به دلایل مطرح شده، همچنان حملات منع خدمت (توزیع‌شده) یکی از تهدیدهای بزرگ در شبکه‌های پهن‌بند محسوب می‌شوند.

این گزارش در پنج بخش تدوین شده است. در بخش ۲ مفاهیم پایه مورد نیاز در این پژوهش معرفی می‌شوند. ابتدا شبکه‌های پهن‌بند و ویژگی‌های آن‌ها بیان می‌شود. سپس انواع حملات منع خدمت، از نقطه نظرهای مختلف مورد بررسی قرار می‌گیرند و در انتهای این بخش به توضیح مفاهیم داده جریان^{۱۵} و انگاره‌ها^{۱۶}، راهکارهای افزایش سرعت پردازش بسته‌ها و معرفی راه‌گزین^{۱۷} های برنامه‌پذیر اختصاص می‌یابد. بخش ۳ به بررسی کارهای پیشین انجام‌شده برای تشخیص حملات منع خدمت توزیع‌شده پرداخته می‌شود. در بخش ۴ روش پیشنهادی به منظور بهبود تشخیص حمله در شبکه‌های پهن‌بند، بیان می‌شود و سرانجام در بخش ۵ نتیجه‌گیری، مراحل انجام پروژه و زمان‌بندی آن بیان خواهد شد.

۲. مفاهیم پایه

در این بخش به شرح مختصری از مفاهیم پایه مرتبط با این پژوهش پرداخته خواهد شد. ابتدا شبکه‌های پهن‌بند را معرفی می‌کنیم. سپس به معرفی حملات منع خدمت و حملات منع خدمت توزیع‌شده می‌پردازیم و در پایان این بخش مفهوم داده جریان و انگاره‌ها را شرح داده می‌شود.

۱.۲ شبکه‌های پهن‌بند

امروزه نرخ تبادل اطلاعات در شبکه‌های کامپیوتری بالا رفته و مفهومی به عنوان شبکه‌های پهن‌بند مطرح می‌باشد. شبکه‌های پهن‌بند دارای سه ویژگی زیر می‌باشند[۱]:

- **سرعت بالا:** داده‌ها و بسته‌ها با سرعت و نرخ بالایی تولید می‌شوند. برای مثال در شبکه‌های نسل پنجم اینترنت همراه^{۱۸}، هر کاربر از قابلیت تبادل اطلاعات با سرعت ۱۵ گیگابیت بر ثانیه برخوردار می‌باشد.

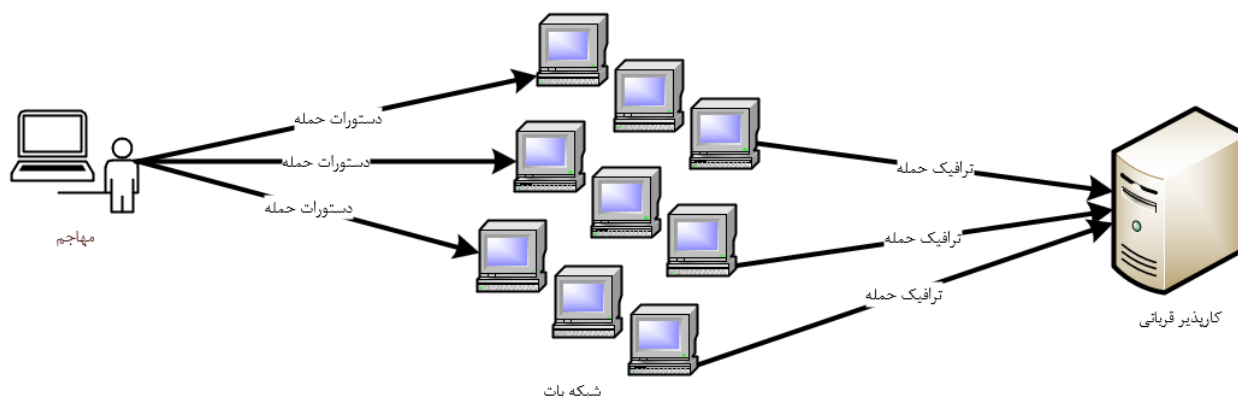
- **حجم بالا:** اطلاعات عبوری از شبکه و داده‌های در حال تبادل باعث تولید حجم زیادی از فراداده^{۱۹} می‌شوند. به عبارتی دیگر بسته‌هایی با محتوا^{۲۰} و حجم زیادی از سرایندها^{۲۱} را خواهیم داشت. به دلیل ظهور کاربردهای مختلف و به دنبال آن پروتکل‌های مختلف و لزوم استفاده از الگوریتم‌های رمزنگاری، حجم زیادی از سرایندها برای برقراری ارتباط الزامی می‌باشد که نگهداشت فراداده‌های تولید شده آن‌ها هزینه زیادی را شامل می‌شود. همچنین اطلاعاتی که کاربران در بستر اینترنت تبادل می‌کنند، می‌تواند طیف وسیعی از داده‌ها شامل فایل‌هایی حجیم و یا جریانی بی‌وقفه از بسته‌ها در هنگام مشاهده یک ویدئوی برخط یا در هنگام برگزاری یک کلاس مجازی باشد. در سال ۲۰۰۳، حجم کل داده‌های تولید شده در اینترنت حدود پنج اگزابایت بود که این میزان در سال ۲۰۰۸ سه برابر شد و به ۱۴.۷ اگزابایت رسید. در سال ۲۰۱۹، ترافیک تولیدی کاربران به میزان پنج اگزابایت داده در هر دو روز می‌رسید [۲].
- **تنوع بالا:** علاوه بر ظهور پروتکل‌های مختلف که هر کدام برای کاربردی خاص می‌باشند، نحوه انتقال و دریافت بسته‌ها بین کارخواه-کارپذیر و استفاده از این پروتکل‌ها وابسته به وضعیت و نوع کاربرد می‌تواند متنوع باشد. برای مثال با اینکه بیشتر برنامه‌های مستقر بر بستر اینترنت، داده‌ها و تبادلات خود را در قالب بسته‌های اچ.تی.تی.پی/اس لایه کاربرد انتقال می‌دهند، اما محتویات این بسته‌ها و نحوه تفسیر آنها برای برنامه‌های مختلف می‌تواند متفاوت باشد.

باتوجه به ویژگی‌های ذکر شده برای شبکه‌های پهن‌بند، مدیریت و کنترل ترافیک در این شبکه‌ها به یکی از چالش‌های اصلی در زمینه شبکه‌های کامپیوتری تبدیل شده است.

۲.۲ حملات منع خدمت (توزیع‌شده)

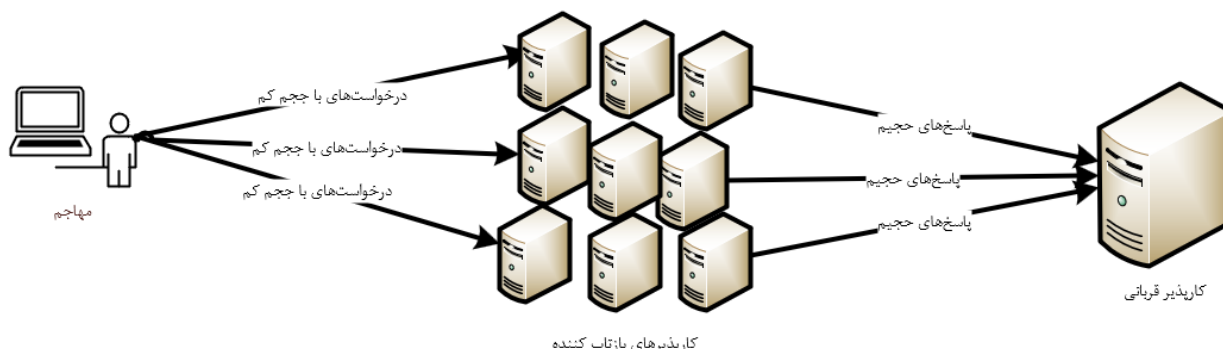
به مجموعه حملاتی که با هدف ممانعت از انجام یک خدمت صورت می‌پذیرند، حملات منع خدمت گفته می‌شود. این حملات با انگیزه‌های مختلفی نظیر ایجاد اختلال یا ممانعت از ارائه یک خدمت، از بین بردن اعتبار و مقبولیت یک سازمان، آسیب زدن مالی و هدر دادن منابع یک سازمان، دستاوردهای سیاسی و ملی، انگیزه مالی و یا قدرت‌نمایی مهاجمین و مواردی از این دست می‌تواند صورت پذیرد. هدف اصلی در حملات منع خدمت تولید ازدحام و اختلال در مصرف منابع پردازشی سیستم (پردازشگر سیستم) یا منابع شبکه (پهنای باند) می‌باشد.

حملات منع خدمت توزیع‌شده گونه خطرناک‌تر از حملات منع خدمت می‌باشند که در آن فرد مهاجم ابتدا با پایش پذیری‌های دستگاه‌های مختلف موجود در شبکه اینترنت، شروع به نفوذ به ماشین‌های عامل^{۲۲} متعددی می‌کند و سعی می‌کند این دستگاه‌ها را تحت کنترل خود کند. به این سیستم‌هایی که توسط فرد مهاجم از راه دور کنترل می‌شوند، ربات گفته می‌شود و این مجموعه ربات‌ها که به آنها شبکه بات^{۲۳} گفته می‌شود، دستورات را از شخص مهاجم دریافت می‌کنند (شکل ۱). مهاجم می‌تواند در مدت زمان کوتاهی حجم زیادی از ترافیک را به سمت کارپذیر و منابع آن هدایت کند که خدمت‌دهی آن یا رویکرد شبکه را برای پاسخگویی به کاربران قانونی با اختلال مواجه می‌کند. در صورت بروز حملات منع خدمت توزیع‌شده، رهگیری مبدأ حمله یعنی نقطه‌ای که حمله از آنجا شروع شده است، دشوارتر و همچنین ترافیک ایجاد شده در اثر حمله بزرگتر و مخرب‌تر می‌باشد.



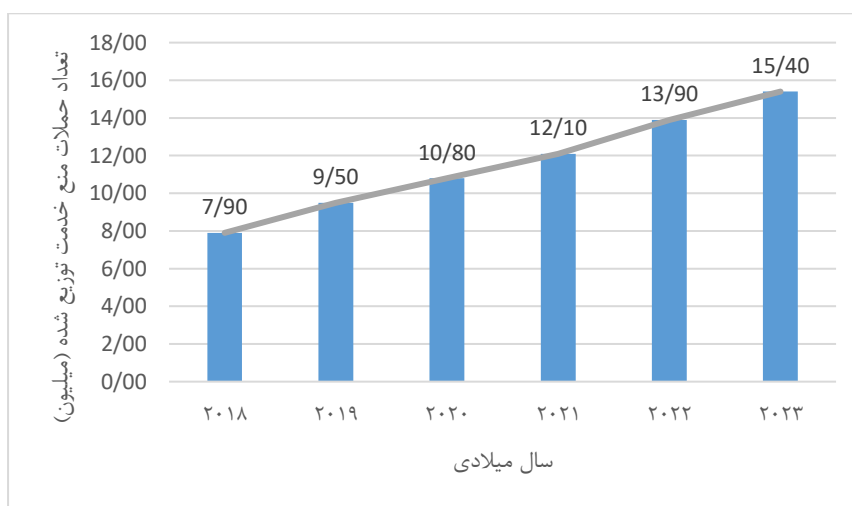
شکل ۱: حملات منع خدمت توزیع‌شده با استفاده از شبکه بات [۴]

مشکل دیگر دفاع در برابر حملات منع خدمت توزیع شده، بروز حملات تقویت بازتاب^{۲۴} است (شکل ۲). در سال ۲۰۱۸، گیت‌هاب^{۲۵} با استفاده از آسیب‌پذیری پروتکل ممکج^{۲۶}، با انعکاس چند برابر بیش از ۵۰۰۰ بار و اوج ترافیک^{۲۷} ۱.۳۵ ترابایت بر ثانیه، قربانی یک حمله منع خدمت توزیع شده از نوع تقویت بازتابی قرار گرفت.



شکل ۲: حملات منع خدمت توزیع شده از نوع تقویت بازتابی [۴]

در فوریه ۲۰۲۰، ارائه‌دهندگان خدمات وب آمازون^{۲۸} حمله‌ای با حجم ترافیک پیک ۲.۳ ترابایت بر ثانیه را تجربه کردند. در ژوئیه ۲۰۲۱، شرکت ارائه‌دهنده خدمات تحویل محتوا^{۲۹} آی‌کلودفلر^{۳۰} در گزارشی به محافظت از یکی از مشتریان خود در برابر حمله منع خدمت توزیع شده نشأت گرفته شده از یک شبکه بات در ابعاد جهانی توسط بدافزار میرای با ترافیک پیک ۱۷.۲ میلیون درخواست در ثانیه، اشاره کرد. یاندکس^{۳۱}، ارائه‌دهنده خدمات پیشگیری از حملات منع خدمت توزیع شده روسیه گفت که در تاریخ ۵ سپتامبر ۲۰۲۱ یک حمله منع خدمت توزیع شده پروتکل اچ.تی.تی.پی را که از تجهیزات شبکه میکروتک^{۳۲} بروز نشده^{۳۳} سرچشمه می‌گرفت، مسدود کرده است. طبق پیش‌بینی شرکت سیسکو^{۳۴}، تعداد حملات منع خدمت توزیع شده به حدود ۱۶ میلیون در سال ۲۰۲۳ میلادی خواهد رسید (شکل ۳). در بزرگترین حمله منع خدمت توزیع شده رخ داده تا به حال، ترافیک حمله به ۱.۴۴ ترابایت در ثانیه رسیده است. از طرفی در سالیان اخیر، این حملات با استفاده از پروتکل‌های جدیدتری ظاهر خود را تغییر می‌دهند. به عنوان مثال، در پایان ژوئیه ۲۰۲۰، پلیس فدرال آمریکا^{۳۵} هشدار صادر کرد مبنی بر اینکه پروتکل برنامه‌های محدود شده^{۳۶} و سایر پروتکل‌ها ممکن است برای انجام حملات منع خدمت توزیع شده مورد سوء استفاده قرار گیرند. حملات منع خدمت توزیع شده بر اساس بردارهای حمله^{۳۷} جدید ممکن است تغییرات زیادی در ویژگی‌های آماری مانند سرعت بسته‌ها و فاصله بسته‌های مورد استفاده در مقایسه با روش‌های سنتی داشته باشند، که این امر باعث می‌شود روش‌های سنتی مقابله در برابر حملات مختلف کارایی لازم را نداشته باشند [۳].



شکل ۳: گزارش و پیش‌بینی سیسکو از مجموع حملات منع خدمت توزیع شده [۵]

۳.۲ داده جریان

همانطور که در ویژگی‌های شبکه‌های پهن‌بند ذکر شد، نرخ بالای تولید اطلاعات یکی از شاخصه‌های این شبکه‌ها می‌باشد. برای پردازش بسته‌ها در این حالت، دو رویکرد متفاوت وجود دارد:

- **پردازش دسته‌ای**^{۳۸}: در این رویکرد تمامی بسته‌ها در یک پنجره زمانی را ضبط کرده و سپس در زمان‌های بعدی پردازش می‌شوند. از مشکلات پردازش دسته‌ای می‌توان به تأخیر در ارسال و پردازش و نیز هزینه بسیار زیاد (برای ذخیره‌سازی) به دلیل ذخیره‌سازی اطلاعات در ابتدای کار و سپس ارسال آن به مراکز دیگر، اشاره کرد.
- **پردازش جریانی**^{۳۹}: اکثر راهکارهای ارائه شده که در قسمت بعد بررسی می‌شوند، مبتنی بر این رویکرد می‌باشند. این الگوریتم‌ها دو مشخصه زیر را در نظر می‌گیرند: اول این که اطلاعات به صورت جریانی از داده‌ها (بی وقفه و با سرعت بالا) در حال ارسال می‌باشند و دوم اینکه از نظر زمانی و حافظه محدودیت وجود دارد [۶]. این خصیصه‌ها همان چالش‌هایی هستند که برای پردازش ترافیک در شبکه‌های پهن‌بند مطرح می‌شوند. برای تشخیص حملات در این شبکه‌ها باید تمامی بسته‌ها را ضبط و پردازش کرده و این کار باید با همان سرعت ورود اطلاعات^{۴۰} و با کمترین میزان استفاده از حافظه انجام شود. الگوریتم‌های پردازش جریانی در بحث پردازش اطلاعات مختلف بسیار کاربردی هستند. الگوریتم‌های مبتنی بر پردازش جریانی، ابتدا مسئله را به یکی از چندین روش موجود مدل می‌کنند. یکی از این مدل‌های بسیار متداول و کاربردی ترنسیتیل^{۴۱} می‌باشد. در این مدل یک داده جریان ورودی به نام I در نظر گرفته می‌شود که شامل مجموعه‌ای از تاپل‌های دوتایی می‌باشد:

$$I = \alpha_1, \alpha_r, \alpha_r, \alpha_f, \dots$$

$$\forall E \alpha_i = \{(a_i, v_i) \mid a_i \in \{0, 1, \dots, u-1\}, v_i \in R\}$$

$$/u/ = \text{key space}$$

تاپل‌ها، دوتایی‌هایی هستند که شامل مقدار کلید و به‌روزرسانی می‌باشند. آرایه‌ای به نام A وجود دارد که تعداد خانه‌های آن برابر $/u/$ و دارای مقادیر متناظر به‌روزرسانی برای هر کلید می‌باشد. هرگاه یک تاپل جدید (a_x, v_x) دریافت شود مقدار به‌روزرسانی آن با مقدار $A[a_x]$ جمع می‌شود:

$$A[a_x] += v_x$$

این پارامترها وابسته به مسئله داده جریانی که مطرح می‌شود، می‌توانند متفاوت باشند. در بحث پردازش بسته‌های دریافتی شبکه، جریان همان جریان ورودی و تاپل‌ها همان بسته‌ها می‌باشند که برای مثال کلیدشان ۵ خصیصه‌ی آدرس آی‌پی^{۴۲} مبدأ، آدرس آی‌پی مقصد، شماره درگاه مبدأ، شماره درگاه^{۴۳} مقصد و پروتکل و به‌روزرسانی نیز می‌تواند اندازه بسته باشد. در نتیجه برای شناسایی حملات منع خدمت، باید آدرس‌هایی که بسته‌هایی با حجم نامتعارف ارسال می‌کنند شناسایی کرد [۷].

در مسائل داده جریان، چندین نوع پاسخ برای مسائل اندازه‌گیری مختلف مطرح می‌باشد و پس از مدل‌سازی مسئله، الگوریتم‌هایی استفاده می‌شود که بر مبنای مدل سعی در یافتن این پاسخ دارند. برای تحلیل بهتر این نوع مسائل، ابتدا مفاهیم اولیه باید توضیح داده شود.

جریان ورودی را توالی از بسته‌هایی به شکل تاپل شامل شناسه جریان متناظر و اندازه آن بسته در نظر گرفته می‌شود

$$F = (\text{srcIP}, \text{srcport}, \text{dstIP}, \text{dstport}, \text{protocol}) = \text{شناسه جریان}$$

$$F = (f_1, c_1), \dots, (f_t, c_t), \dots$$

پاسخ‌های مسائل یکی از انواع زیر می‌باشند [۸]:

- سائز هر جریان^{۴۴}: خواسته این مسائل، یافتن سائز جریان یا تعداد بسته‌های دریافت شده متعلق به جریان f می‌باشد که با nf نشان داده می‌شود. سائز تمامی بسته‌های دریافتی نیز $nf = \sum_{1 \leq i \leq F} n_f$ می‌باشد.
- لحظه جریان^{۴۵}: در لحظه دلخواهی، وضعیت جریان با استفاده از تابع g در لحظه g^{46} مطلوب است، که به صورت زیر می‌تواند تعریف شود:

$$L_g = \sum_{1 \leq f \leq F} g(nf), f \in [1, F]$$

این تابع وابسته به این که به چه صورت تعریف شده باشد، می تواند وضعیت کلی از ترافیک شبکه را به صورت عددی بیان کند.

○ شاخص^{۴۷}: شاخص ها جریان هایی هستند که اندازه آنها بر لحظه جریان L_g بیشترین تاثیر را می گذارد. به عبارتی دیگر:

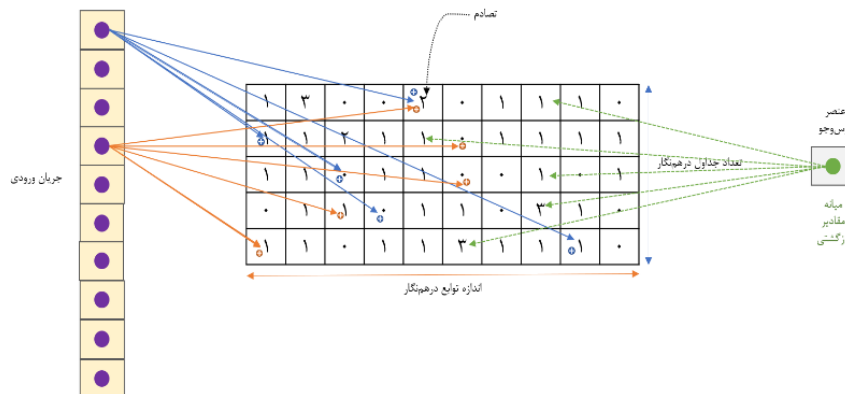
$$H_g = \{ f \mid g(nf) \geq \alpha L_g \}$$

که α مقدار آستانه از پیش تعریف شده بین صفر و یک می باشد.

۴.۲ انگاره

برای حل مسائل داده جریان، راهکارهای متفاوتی را می توان استفاده کرد. یکی از راهکارها بدین صورت می باشد که به دلیل اینکه با حجم زیادی از اطلاعات روبرو هستیم، تنها بخشی از داده های ورودی به عنوان نمونه انتخاب شوند و عملیات پردازش تنها روی آن ها صورت گیرد. این روش نمونه برداری^{۴۸} نامیده می شود [۹]. نمونه برداری دقت پایینی خواهد داشت. به منظور بالا بردن دقت، پردازش تمامی بسته ها الزامی می باشد. اما بررسی همه بسته ها نیز نیازمند حجم زیادی از منابع پردازشی و زمان می باشد. برای حل این مشکل، الگوریتم هایی به نام انگاره ارائه شده اند که از یک داده ساختار فشرده برای ذخیره سازی اطلاعات داده های ورودی استفاده می کنند. انواع مختلفی از این الگوریتم ها در پژوهش های مختلف ارائه شده است که هر کدام سعی در حل یکی از انواع مسائل داده جریان دارند. در ذیل چند مورد از پراستفاده ترین آن ها معرفی خواهند شد:

- انگاره شمارشی^{۴۹}: از یک جدول $K \times H$ تشکیل شده است که شامل K تابع درهم نگار^{۵۰} می باشد. H نیز اندازه توابع درهم نگار در یک سطر می باشد. این ساختار را الگوریتم های انگاره دیگر نیز استفاده می کنند. چون از توابع درهم نگار استفاده می شود لذا امکان تصادم^{۵۱} وجود خواهد داشت. مقادیر بازگشتی تخمینی خواهند بود و در نتیجه به آنها داده ساختارهای آماری احتمالاتی می گویند. اما بایستی نرخ خطای قابل قبول و کرانداری ارائه دهند. از این الگوریتم برای یافتن پاسخ مسائل شاخص استفاده می شود [۱۰].

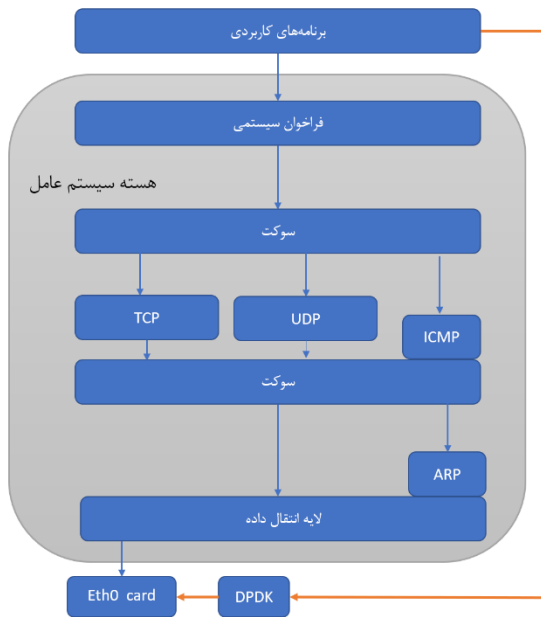


شکل ۴: انگاره شمارشی [۱۱]

- انگاره شمارشی کمینه^{۵۲}: همانند انگاره شمارشی می باشد اما سعی دارد مرتبه فضایی را کاهش دهد [۱۲].
- انگاره عمومی^{۵۳}: یک دسته جدیدی از انگاره ها با هدف ارائه داده ساختاری قابل استفاده برای حل تمامی انواع مسائل داده جریان می باشند. یونیومان^{۵۴} یکی از این الگوریتم ها می باشد [۱۳].

۵.۲ پردازش سریع بسته ها

هنگامی که یک بسته از طریق واسطه های شبکه یک سیستم دریافت می شود تا پردازش آن، مراحل مختلفی را طی خواهد کرد. بنا به کاربرد، بسته ها از دستگاه های مختلفی عبور داده می شوند. بر مبنای پشته پروتکل تی.سی.پی/آی.پی^{۵۵} که هسته^{۵۶} تمامی توزیع های مختلف سیستم عامل لینوکس از آن پشتیبانی می کنند، بسته ها از دریافت تا پردازش بخش های مختلف آن ها مراحل مختلفی را پشت سر خواهند گذاشت و در نهایت در صورت نیاز بازار سال خواهند شد. رویدادهای مهم در هنگام دریافت یک بسته توسط ماشین بدین شرح می باشد:



شکل ۵: مراحل ضبط و پردازش بسته (۱۳)

- بسته توسط کارت شبکه^{۵۷} ماشین دریافت می‌شود (وقفه کارت شبکه).
- کارت شبکه از طریق دی.ام.ای^{۵۸}، بسته را در فضای حافظه در یک بافر قرار می‌دهد.
- کارت شبکه سیگنالی به پردازنده ارسال می‌کند و آن را برای پردازش بسته بیدار می‌کند (وقفه نرم افزاری).
- پردازنده اطلاعات مورد نیازش را خوانده و در صورت نیاز در فضای بافر تعیین شده می‌نویسد.
- در صورت نیاز، بسته برای پردازش‌های بیشتر به پشته پروتکلی هسته برای انجام پردازش‌های مختلف (مثل بررسی آدرس آی.پی برای تطبیق با آدرس‌های متناظر لیست کنترل دسترسی) فرستاده می‌شود.
- در نهایت اگر برنامه کاربردی در سطح کاربر باشد، محتویات بسته از فضای هسته به فضای کاربر انتقال داده خواهد شد. در غیر اینصورت، بسته در همان فضای هسته خواهد ماند.

تمامی این مراحل بایستی در سطح هسته انجام شده ولی پردازش بسته توسط کاربر در لایه کاربرد صورت می‌گیرد. این مراحل به دلیل وقفه‌هایی که انجام می‌شود، سربار زیادی خواهند داشت و در شبکه‌های پهن‌بند که با حجم زیادی از بسته‌ها مواجه هستیم، باعث اتلاف وقت زیادی خواهند شد.

دی.پی.دی.کی^{۵۹} و ایکس.دی.پی^{۶۰} از ابزارهای موجود برای تسریع عملیات پردازش بسته می‌باشند. دی.پی.دی.کی ابزار نرم افزاری می‌باشد که در سال ۲۰۰۹ توسط اینتل^{۶۱} توسعه داده شد. ولی بعدها به صورت یک پروژه متن باز^{۶۲} درآمد. به طور خلاصه یک ابزار دورزدن هسته در هنگام دریافت بسته در شبکه می‌باشد که وقفه‌های مختلف مربوط به هسته را حذف می‌کند و لذا تمام عملیات پردازش بسته را می‌توان در سطح کاربر انجام داد و در نهایت عملیات دریافت و پردازش بسته را تا حد خوبی می‌تواند تسریع بخشد. هدف این فناوری استفاده از قابلیت پردازش چندهسته‌ای پردازنده‌های معمولی ایکس۸۶^{۶۳} برای بهبود سرعت پردازشی کاربردها می‌باشد. بدین صورت ما نرخ پردازشی برابر هنگام استفاده از پردازنده‌های مخصوص کاربردها و یا مدارهای مجتمع با کاربرد خاص^{۶۴} و مدار مجتمع دیجیتال برنامه‌پذیر^{۶۵}، با صرف هزینه‌ای بسیار کمتر، خواهیم داشت. از چندین پردازنده برای محاسبات مربوط به سطح داده و از بقیه هسته‌ها برای امور کنترلی و خدمات دیگر استفاده می‌کند. به صورت جزئی‌تر، چندین صف بر روی هر واسط شبکه تعریف می‌کند و هسته‌ها با حالت سرکشی^{۶۶} به این صف‌ها الصاق می‌شوند. از این ابزار در کاربردهای مختلفی در مواقعی که حجم زیادی از ورودی/خروجی مطرح می‌باشد از حیطه شبکه و امنیت آن، پردازش و راهگزینی در ابرها، بهبود کارایی حافظه‌ها، توابع مجازی شبکه^{۶۷}، مخابرات و تلکام استفاده می‌شود. البته به غیر از مورد اشاره شده که ویژگی اصلی این ابزار می‌باشد، امکانات مختلف دیگری مانند رمزگذاری و فشرده‌سازی به کمک رابط‌های برنامه‌نویسی‌اش نیز ارائه می‌دهد[۱۴].

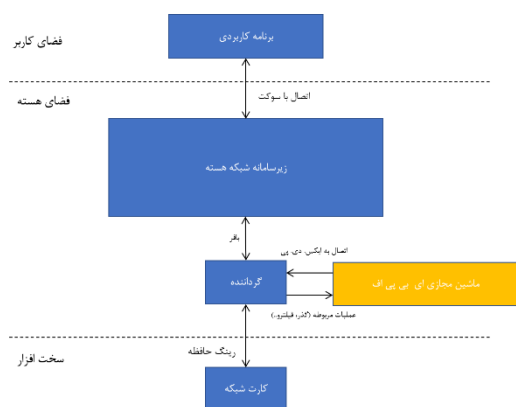
ایکس.دی.پی.کی از مؤلفه‌های جدید هسته می‌باشد که پردازش بسته را به صورت خوبی بهبود می‌بخشد. روش‌هایی مثل دی.پی.دی.کی هسته را دور می‌زنند و تمام عملیات پردازش بسته در فضای کاربر صورت می‌گیرد. همچنین کارت شبکه را باید توسط یک گرداننده^{۶۸} سطح کاربر کنترل نمود. پردازش شبکه در سطح کاربر با وجود مزایای زیادش، معایب زیر را نیز به همراه خواهد داشت:

- به دلیل اینکه سیستم عامل یک لایه انتزاعی برای ارتباط با منابع سخت‌افزاری می‌باشد لذا برنامه‌های سطح کاربر برای تعامل با آنها بایستی گرداننده‌های مربوطه را خودشان توسعه دهند.
- برنامه‌های سطح کاربر می‌بایست در صورت نیاز عملکردهایی که توسط هسته ایجاد می‌شد را پیاده‌سازی کنند.
- برنامه‌ها به صورت ایزوله اجرا می‌شوند که نحوه تعامل آنها با دیگر بخش‌های سیستم عامل را دشوار می‌کند.

به طور خلاصه ایکس.دی. پی، برنامه‌های شبکه سطح کاربر (پالایش، نگاشت، مسیریابی و...) را به جای انتقال به سطح کاربر، به فضای هسته می‌برد. ایکس.دی.پی امکان اجرای برنامه به محض ورود بسته به کارت شبکه و پیش از حرکت به سمت زیرسیستم شبکه‌ی هسته را فراهم می‌کند که منجر به افزایش قابل توجه سرعت پردازش بسته می‌شود. اجرای برنامه در سطح هسته با استفاده از بی.پی.اف^{۶۹} میسر می‌شود [۱۵].

بی.پی.اف یک ماشین مجازی است که تنها مخصوص پردازش پالایش ترافیک می‌باشد. یکی از ابزارهایی که از بی.پی.اف استفاده می‌کند، تی.سی.پی دامپ^{۷۰} می‌باشد. عبارت پالایش مربوطه توسط یک کامپایلر به بایت‌کد^{۷۱} بی.پی.اف تبدیل خواهد شد. از آنجایی که بی.پی.اف یک ماشین مجازی می‌باشد، محیطی به منظور اجرای برنامه‌ها در آن که علاوه بر بایت‌کد شامل یک مدل حافظه مبتنی بر بسته (دستورالعمل‌های بارگذاری به طور ضمنی بر روی بسته موردنظر انجام می‌شود)، ثبات‌ها^{۷۲} (A و X یعنی انباشتگر^{۷۳} و ثبات اندیس^{۷۴})، یک حافظه موقت و یک شمارنده برنامه^{۷۵} ضمنی نیز می‌باشد را تعریف می‌کند. هسته لینوکس از نسخه ۲.۵ به بعد از بی.پی.اف پشتیبانی می‌کند. در سال ۲۰۱۱، مفسر بی.پی.اف به یک کامپایلر درجا^{۷۶} تغییر داده شد. این کار باعث شد که هسته به جای تفسیر برنامه‌های بی.پی.اف، قادر باشد که آن‌ها را به معماری هدف‌های مختلف مانند ایکس۸۶، میپس^{۷۷}، و آرم^{۷۸} تبدیل کند. این امر به معرفی بی.پی.اف توزیع یافته^{۷۹} در سال ۲۰۱۴ و کنار گذاشته شدن بی.پی.اف سنتی منجر شد. ویژگی‌های جدید شامل موارد زیر می‌باشد:

- از ویژگی‌های معماری ۶۴-بیتی مثل ثبات‌ها و تعداد آن‌ها و کدهای عملیاتی^{۸۰} بیشتر بهره می‌برد.
- از زیرسیستم شبکه جدا شده است و امکان استفاده در کاربردهای دیگر میسر می‌شود.
- نگاشت‌ها به عنوان راهی برای تبادل داده بین سطح کاربر و هسته مورد استفاده قرار می‌گیرند.
- استفاده از توابع کمکی که در سطح هسته اجرا می‌شوند. امکان فراخوانی فراخوان سیستمی^{۸۱} در برنامه‌های بی.پی.اف نیز وجود دارد.
- زنجیره سازی تعداد برنامه بیشتر بی.پی.اف نیز امکان‌پذیر خواهد بود.



شکل ۶: ایکس.دی.پی بر روی هسته سیستم عامل [۱۵]

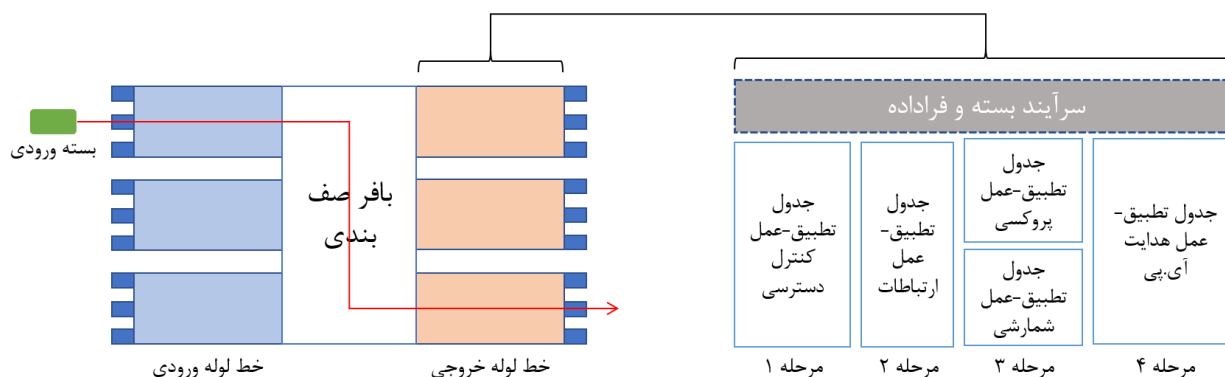
برخی از صف‌های کارت شبکه هنوز به هسته متصل هستند، در حالی که برخی دیگر به یک برنامه فضای کاربر متصل هستند که در مورد حذف شدن یا نشدن یک بسته تصمیم می‌گیرد. با این کار، میزان ترافیکی که به زیرسیستم شبکه هسته می‌رسد به میزان قابل توجهی کاهش می‌یابد. برای این کار بایستی یک نقطه بازرسی^{۸۲} در پشته هسته تعریف کرد که هرگاه بسته‌ای در کارت شبکه دریافت شد، آن را به فضای کاربر بفرستد و در آنجا تصمیم می‌گیرد که بسته دور انداخته شود^{۸۳} یا اجازه عبور به لایه‌های بالاتر پشته را صادر کند. لذا نیاز به مکانیزمی بود که امکان اجرای کدهای سطح کاربر را در هسته فراهم کند. به همین دلیل از بی.پی.اف توزیع یافته استفاده شد.

ایکس.دی.پی بسته‌های دریافتی را به برنامه بی.پی.اف هدایت می‌کند. در آنجا می‌توان بسته‌ها را ویرایش و یا هدایت^{۸۴} کرد. از توابع کمکی می‌توان برای انجام محاسبات و پردازش بسته‌ها بدون نیاز به فراخوان سیستمی استفاده کرد. همچنین با استفاده از داده‌ساختارهای نگاشت امکان ذخیره داده‌ها به صورت دائمی را خواهیم داشت. در نهایت با استفاده از ویژگی‌های از پیش تعبیه شده در ایکس.دی.پی می‌توان عمل مورد نظر را بر روی بسته انجام داد [۱۶].

۶.۲ راه‌گزین‌های برنامه‌پذیر

استفاده از یک کنترل‌کننده به عنوان مرکزی که تمام اطلاعات به آن فرستاده می‌شود و سپس در آنجا بر مبنای الگوریتم پیاده شده بر روی آن، تصمیم می‌گیرد که جلوی ترافیک را بگیرد یا نه، یکی از مشکلات برخی روش‌های تشخیص پیشین بود. این روش با تأخیر زیادی همراه است و همچنین می‌تواند یک نقطه آسیب‌پذیر واحد برای مهاجمین فراهم کند. اما امروزه با معرفی راه‌گزین‌های برنامه‌پذیر^{۸۵}، راه‌گزین‌های معمولی نیز، با استفاده از برنامه‌هایی که بر روی آن‌ها با استفاده از زبان‌هایی مثل پی.آر.آی^{۸۶} توسعه داده می‌شوند، توانایی پردازش داده را تا حد

زیادی خواهند داشت. یک راهگزین برنامه‌پذیر مبتنی بر مدارهای مجتمع با کاربرد خاص، چندین خط لوله شامل واسطه‌های ورودی و خروجی را شامل می‌شود و بسته‌ها مراحل مختلفی را در طول خط لوله برای پردازش سپری می‌کنند. هر کدام از این مراحل نیز منابع اختصاصی خود یعنی: ثبات‌ها برای ذخیره‌سازی، جداول تطبیق-عمل، و واحدهای منطق ریاضی به منظور پردازش را شامل می‌شوند. توسط زبان پی. ۴ امکان شخصی‌سازی جداول تطبیق-عمل به منظور انجام تغییر روی بسته‌ها میسر خواهد بود. در مجموع راهگزین‌های برنامه‌پذیر مبتنی بر مدارهای مجتمع با کاربرد خاص، دو برتری بهینه بودن سرعت پردازشی به نسبت هزینه مصرفی و مصرف برق و انعطاف‌پذیری در برابر حملات جدید را در مقایسه با سخت‌افزارهای دیگر ارائه می‌دهند [۱۷].



شکل ۷: معماری راه‌گزین برنامه‌پذیر [۱۷]

۳. کارهای پیشین

به‌طور کلی پژوهش‌های انجام‌شده در حوزه حملات منع خدمت توزیع‌شده را می‌توان در سه دسته پیشگیری از وقوع حمله، تشخیص حمله و کاهش اثر حمله^{۸۷} تقسیم‌بندی کرد. از آنجایی که تمرکز این گزارش بر پژوهش‌های موجود در حوزه تشخیص حملات منع خدمت توزیع‌شده می‌باشد، در ادامه به بررسی چند روش اخیراً معرفی‌شده تشخیص حملات منع خدمت توزیع‌شده در شبکه‌های کامپیوتری می‌پردازیم. علاوه بر شیوه‌ی دسته‌بندی‌ای که در ادامه استفاده می‌کنیم، الگوریتم‌های تشخیص را می‌توان بر اساس اینکه در کدام ناحیه از شبکه سعی به تشخیص مهاجم دارند نیز طبقه‌بندی کرد، که شامل سه گروه می‌شوند:

- شناسایی در مبدأ: از توانایی تشخیص همه حملات برخوردار نمی‌باشند.
- شناسایی در مقصد (قربانی): نیاز به منابع بیشتری دارند و ممکن است با تأخیر هم همراه باشند.
- شناسایی در مسیرهای میانی^{۸۸}.

۱.۳ روش‌های مبتنی بر امضا

طبقه‌بندی و فیلترکردن ترافیک شبکه با استفاده از امضا، معمولاً در سیستم‌های تشخیص نفوذ^{۸۹} استفاده می‌شود. در این روش، مقادیر فیلدهای مختلف بسته‌های عبوری با یکسری از امضاها (الگوها) ی بسته‌های مهاجم به منظور تشخیص حمله، مقایسه می‌شوند. اگرچه این روش‌ها الگوهای حمله‌ی مشاهده‌شده‌ی قبلی را با سرعت بالایی شناسایی می‌کنند، اما قادر به شناسایی حملات جدید نمی‌باشند.

- **تقسیم‌کننده و کاهنده ترافیک حملات منع خدمت توزیع‌شده مبتنی بر امضا** با استفاده از راهگزین‌های برنامه‌پذیر سطح داده: در روش ارائه‌شده در سال ۲۰۲۱، دیمولیانس و همکاران سعی می‌کنند امضاها را بدست آورند و تعداد حداقل بهینه خط قوانین به منظور مقابله با آنها را تولید کنند [۱۸]. مشکل روش، عدم کارایی در شناسایی حملات متنوع می‌باشد. همچنین در مورد نحوه یاددهی مجدد مدل‌های طبقه‌بندی کننده توضیحی ارائه نمی‌دهد.
- **روش تشخیص مبتنی بر جریان در شبکه‌های با سرعت بالا برای شناسایی حملات** با تولید امضای سازگار با اسنورت^{۹۰}: در روش ارائه شده در سال ۲۰۲۰ توسط ارلاکر و همکاران، از دسته‌بندی جریان مبتنی بر آی.پی. فیکس^{۹۱} استفاده می‌شود، که اطلاعات بیشتری

علاوه بر اطلاعات آماری متداول می‌تواند استخراج کند، همچنین با استفاده از برخی روش‌ها امکان بررسی محتوای داده‌ای نیز میسر خواهد بود [۱۹]. مشکل این روش عدم کارایی در شناسایی حملات مختلف می‌باشد.

۲.۳ روش‌های مبتنی بر ناهنجاری^{۹۲}

این دسته از روش‌ها با ضبط کردن و بررسی ترافیک عادی شبکه در یک بازه زمانی، رفتار عادی شبکه را شبیه‌سازی می‌کنند و هرگونه رفتار مغایر با این مدل را به عنوان ناهنجاری و بروز حمله در نظر می‌گیرند. یادگیری ماشین به‌عنوان یکی از روش‌های کارآمد مبتنی بر ناهنجاری می‌باشد، که امروزه به صورت گسترده‌ای مورد استقبال پژوهشگران قرار گرفته است. در حوزه تشخیص حملات منع خدمت توزیع‌شده نیز از این روش استفاده می‌شود. گونه‌های مختلفی از الگوریتم‌های یادگیری ماشین نظیر استفاده از ماشین بردار پشتیبان، بیز ساده، نزدیک‌ترین همسایه، شبکه عصبی و شبکه‌های عصبی ژرف، نگاشت خودسازمان‌ده و مواردی از این قبیل به‌منظور انجام طبقه‌بندی جریان ترافیک و تشخیص ناهنجاری در آن مورد استفاده قرار می‌گیرند.

- **رویکرد تشخیص ناهنجاری به کمک راهگزین‌های برنامه‌پذیر برای شناسایی و مقابله با حملات منع خدمت توزیع‌شده:** این روش به نام جاکن در سال ۲۰۲۱ توسط لیو و همکاران با استفاده از انگاره‌های عمومی و پیاده‌سازی آنها روی راهگزین‌های برنامه‌پذیر به منظور جمع‌آوری اطلاعات توسط همین دستگاه‌ها، ارائه شد. یک کنترل‌کننده مرکزی از این اطلاعات برای تشخیص حملات استفاده می‌کند. همچنین الگوریتم‌هایی به منظور رفع مخاطره بر روی این راهگزین‌ها می‌توان پیاده کرد. به دلیل این که با استفاده از زبان پی‌۴، الگوریتم‌های تشخیص و رفع مخاطره پیاده می‌شود، لذا این روش مبتنی بر معماری خاصی از راهگزین‌ها نمی‌باشد [۲۰]. عدم واریسی محتوای کامل بسته‌ها و عدم استفاده از واریسی‌کننده عمیق بسته^{۹۳}، مشکل اصلی جاکن می‌باشد. به همین دلیل به بحث تنوع پروتکلی و اینکه مقادیر آستانه برای برنامه‌های کاربردی مختلف می‌تواند متفاوت باشد، اشاره‌ای نکرده است.
- **روش بلادرنگ تطبیق‌پذیر مبتنی بر انگاره مخصوص شبکه‌های ارائه دهنده خدمات اینترنتی:** آر.تی.سد^{۹۴} در سال ۲۰۲۱ توسط شی و همکاران پیشنهاد شد. از نامتوازن بودن مقادیر یک ویژگی برای یک آدرس مقصد مشخص، قربانی بودن آن را تشخیص می‌دهد [۳]. به دلیل استفاده از انگاره‌ها، آر.تی.سد از نظر مرتبه فضایی بسیار بهینه می‌باشد. اما مسئله نحوه انتخاب این ویژگی‌ها برای برنامه‌های کاربردی مختلف به صورت پویا را در نظر نمی‌گیرد.
- **روش توزیع شده مقابله با حملات منع خدمت توزیع‌شده به کمک شبکه‌های عصبی:** روشی به نام دفاع هوشمند در سال ۲۰۲۲ توسط ماینی و همکاران ارائه شد که از شبکه‌های عصبی عمیق در سمت لبه مشتری و شبکه‌های عصبی عمیق با الگوریتم‌های پیشرفته‌تر در سمت فراهم‌کننده اینترنت برای شناسایی حملات استفاده می‌کند [۱۹]. اما مشکل این روش، عدم ارائه راهکاری بهینه به منظور آموزش مجدد شبکه‌ها در شبکه‌های پهن‌بند می‌باشد.
- **روش بلادرنگ مبتنی بر ناهنجاری برای تشخیص رخنه در شبکه‌های با سرعت بالا:** روشی توسط ویگاس و همکاران در سال ۲۰۱۹ ارائه شد که با در نظر گرفتن بسته‌ها به عنوان جریان، ویژگی‌های آن‌ها را استخراج می‌کند و سپس از روی آنها حمله را تشخیص می‌دهد. یک بخش اعتمادسازی دارد که میزان قابل اعتماد بودن گروه‌بندی ارائه شده توسط طبقه‌بندی‌کننده را بررسی می‌کند و اگر از مقدار حداقلی پایین بود، به کمک یک شخص مدیر آن را برچسب‌گذاری می‌کند و سپس مدل طبقه‌بندی‌کننده را به صورت افزایشی به روز می‌کند [۲۰]. به دلیل مداخله انسان برای برچسب‌گذاری برخی جریان‌ها، در شبکه‌های پهن‌بند با مشکل مواجه می‌شود و سربار بالا و دقت پایینی خواهد داشت.

۴. روش پیشنهادی

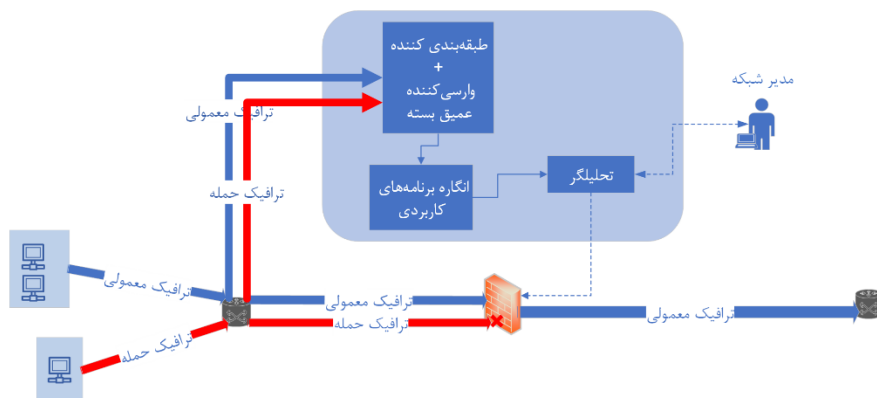
در بخش قبل برخی روش‌های مبتنی بر امضا و مدل‌سازی به منظور تشخیص حملات منع خدمت توزیع‌شده معرفی شدند. به عنوان نتیجه می‌توان گفت، روش‌های شناسایی و مقابله با حملات منع خدمت توزیع‌شده، ویژگی‌های ترافیک را از سه منظر بررسی می‌کنند (ترافیک را از سه منظر مشاهده می‌کنند) و سعی در مقابله دارند:

- **بسته:** با استفاده از ویژگی‌ها در سطح بسته (برای مثال یک بسته اچ. تی. پی با اندازه خاص) حملات را تشخیص می‌دهند.

- **جریان:** در صورت مشاهده ویژگی‌های نامتعارف برای یک جریان (برای مثال داشتن اندازه‌ای بیش از ۱۰۰ کیلوبایت) حملات را تشخیص می‌دهند.
- **رفتار کاربر:** با استفاده از نامتعارف بودن رفتار ترافیک کاربری سعی می‌کنند حملات را تشخیص دهند. به طور مثال کاربری که در مدت زمان مشخصی، بیش از ۱۰ درخواست به منابع مختلف یک سایت ارسال کند.

در هر یک از منظرها طبق اطلاعاتی که جمع‌آوری می‌کنند، ترافیک نامتعارف را با استفاده از رویکردهای مبتنی بر امضا یا با استفاده از رویکردهای مدل‌سازی و تشخیص ناهنجاری، می‌توانند حملات را تشخیص دهند. اما عدم سازگاری با تنوع پروتکلی برنامه‌های کاربردی مختلف، از مشکلات آن‌ها می‌باشد و در هنگام تشخیص حملات دچار خطا می‌شوند. در واقع ترافیک برای هر کاربرد می‌تواند الگوی مختلفی داشته باشد و برای همه کاربردها نمی‌توان یک الگو، رمز و شناسه برای حالت متعارف آن تعریف نمود. برای حل این مشکل به شناسایی کاربردهای مختلف می‌پردازند که از روش‌هایی مانند واریانس عمیق بسته یا یادگیری ماشین استفاده می‌شود و سپس با استفاده از نتایج آن‌ها ترافیک را دسته‌بندی کرده و در هر کدام برای تشخیص الگوهای نامتعارف، تنظیمات متفاوتی (مانند مقادیر آستانه متفاوت برای حجم بسته‌ها) به کار می‌برند. اما در شبکه‌های پهن‌بند با مشکلی به نام تنوع ترافیکی بالا مواجه هستیم و از طرفی با توجه به نرخ بالای تولید ترافیک بایستی در کمترین زمان ممکن، کم هزینه‌ترین راهکار را استفاده کنیم. راهکارهای مبتنی بر یادگیری ماشین و استفاده از واریانس عمیق بسته، سربار محاسباتی زیاد دارند.

همانطور که گفته شد، نکته‌ای که در پژوهش‌های پیشین نادیده گرفته می‌شد، مربوط به مؤلفه سوم شبکه‌های پهن‌بند یا تنوع بالای ترافیک می‌باشد. در روش‌های پیشین مؤلفه‌های اول و دوم یعنی اینکه داده‌ها با سرعت زیادی در حال تولید هستند و با حجم زیادی از سرآیندها و محتوا روبرو هستیم را تنها در نظر گرفته بودند. لذا روشی که ارائه می‌دهیم تمامی این سه مورد را با تمرکز بیشتر بر روی ویژگی سوم به عنوان مسئله اصلی، هدف قرار می‌دهد. روش ارائه شده، از ویژگی پردازش جامع برخوردار می‌باشد، یعنی تمامی بسته‌ها را یک و تنها یکبار بررسی می‌کند. بدین صورت روشی بسیار سریع و با دقت بالا و تطبیق‌پذیر با مشخصات ترافیکی شبکه‌های پهن‌بند می‌باشد (علاوه بر معیارهای متداولی مثل سرعت-نرخ‌گذر بالا و تأخیر کم که خواسته همه روش‌های قبلی بوده است) که راه‌حل نوینی نیز به شمار می‌رود.



شکل ۸: معماری کلی روش پیشنهادی

روش پیشنهادی ما بدین صورت خواهد بود که با استفاده از راهکارهای ارائه شده در پژوهش‌های پیشین و همچنین استفاده از یک زیرسامانه واریانس عمیق بسته، جریان‌ها را برچسب‌گذاری می‌کنیم و جریان‌های شبیه به هم از نظر رفتار را در یک گروه قرار می‌دهیم. این اطلاعات در داده‌ساختارهای انگاره که بر روی راهگزين‌های برنامه‌پذیر می‌باشند و توسط مدیر شبکه کنترل می‌شوند، برای هر برنامه کاربردی و پروتکل متناظر به صورت جدا ذخیره می‌شود. یک قسمت تحلیلگر وجود دارد که با استفاده از این ویژگی‌های آماری و مشاهده رفتار متداول هر پروتکل و یا برنامه کاربردی در بازه‌های زمانی مختلف، این اطلاعات را با مقادیر آستانه‌ای که از قبل به دست آورده و نشان‌دهنده حداکثر بی‌نظمی قابل چشم‌پوشی در شبکه می‌باشد، مقایسه می‌کند و در صورت مشاهده مغایرت آن جریان را به عنوان یک حمله تشخیص داده و سعی می‌کند خط‌مشی امنیتی معادل آن را تولید کند و به عنوان خروجی به یک دیوار آتش ارسال کند. برای بهبود سرعت پردازش بسته‌ها و حذف وقفه‌های

زمانبر نیز از ابزارهای مخصوص پردازش سریع بسته‌ها (مانند دی.پی.کی و ایکس.دی.پی) استفاده می‌شود. همچنین کارایی روش ارائه شده در مقایسه با دیگر راهکارها و با در نظر گرفتن معیارهایی نظیر میزان استفاده از پردازشگر و حافظه، نرخ دورانداختن بسته‌ها، و میزان تاخیر در شناسایی حملات در مواجهه با ترافیک‌های حجیم و ترافیک‌های تولیدی از انواع مختلف حملات منع خدمت، بررسی خواهد شد.

۵. نتیجه‌گیری

در این نوشتار به معرفی شبکه‌های پهن‌بند و ویژگی‌های این شبکه‌ها، حملات منع خدمت توزیع‌شده، روش‌ها و الگوریتم‌های پردازش داده جریان و معرفی مفاهیم و واژه‌های به کاررفته در این زمینه، مانند محاسبات و مسائل داده جریان و موضوعاتی دیگر پرداخته شد. سپس برخی پژوهش‌های انجام‌شده در زمینه تشخیص حملات منع خدمت توزیع‌شده در شبکه‌های پهن‌بند مورد بررسی قرار گرفت و مشکلات پیاده‌سازی و عملکردی و چالش‌های حل‌نشده آنها بیان شد. در پایان روش پیشنهادی سریع با دقت بالا و بهینه از نظر میزان مصرف منابع و سازگار با تنوع ترافیکی به منظور شناسایی حملات منع خدمت توزیع‌شده در بستر شبکه‌های پهن‌بند به صورت مختصر توضیح داده شد. در جدول ۱ زمان‌بندی لازم برای بخش‌های مختلف این پژوهش آورده شده است.

جدول ۱: مراحل انجام و پیشبرد پروژه

فعالیت	میزان پیشرفت	تخمین زمان باقی‌مانده
۱. مطالعه و بررسی مفاهیم	۹۰٪	۱ هفته
۲. تحلیل و بررسی کارهای پیشین	۷۰٪	۳ هفته
۳. ارائه و امکان‌سنجی روش پیشنهادی	۱۰٪	۶ هفته
۴. پیاده‌سازی روش پیشنهادی	۱۰٪	۸ هفته
۵. ارزیابی روش پیشنهادی	۰٪	۶ هفته
۶. جمع‌بندی و تدوین پایان‌نامه	۰٪	۶ هفته

کتاب‌نامه

- [1] M. Noferesti and R. Jalili, 'ACoPE: An adaptive semi-supervised learning approach for complex-policy enforcement in high-bandwidth networks', *Computer Networks*, vol. 166, p. 106943, Jan. 2020.
- [2] R. K. Deka, D. K. Bhattacharyya, and J. K. Kalita, 'Active learning to detect DDoS attack using ranked features', *Computer Communications*, vol. 145, pp. 203–222, Sep. 2019.
- [3] H. Shi, G. Cheng, Y. Hu, F. Wang, and H. Ding, 'RT-SAD: Real-Time Sketch-Based Adaptive DDoS Detection for ISP Network', *Security and Communication Networks*, vol. 2021, pp. 1–10, Jul. 2021.
- [4] R. Vishwakarma and A. K. Jain, 'A survey of DDoS attacking techniques and defence mechanisms in the IoT network', *Telecommun Syst*, vol. 73, no. 1, pp. 3–25, Jan. 2020.
- [5] 'Cisco Annual Internet Report (2018–2023) White Paper', Mar. 2020. [Online]. Available: <https://www.a10networks.com/blog/5-most-famous-ddos-attacks/>
- [6] B. Zhao, X. Li, B. Tian, Z. Mei, and W. Wu, 'DHS: Adaptive Memory Layout Organization of Sketch Slots for Fast and Accurate Data Stream Processing', in *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, Virtual Event Singapore, pp. 2285–2293, Aug. 2021.
- [7] B. Krishnamurthy, S. Sen, Y. Zhang, and Y. Chen, 'Sketch-based change detection: methods, evaluation, and applications', in *Proceedings of the 2003 ACM SIGCOMM conference on Internet measurement - IMC '03*, Miami Beach, FL, USA, p. 234, 2003.
- [8] Q. Xiao, Z. Tang, and S. Chen, 'Universal Online Sketch for Tracking Heavy Hitters and Estimating Moments of Data Streams', in *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications*, Toronto, ON, Canada, pp. 974–983, Jul. 2020.

- [9] V. Sivaraman, S. Narayana, O. Rottenstreich, S. Muthukrishnan, and J. Rexford, ‘Heavy-Hitter Detection Entirely in the Data Plane’, in *Proceedings of the Symposium on SDN Research*, Santa Clara CA USA, pp. 164–176, Apr. 2017.
- [10] M. Charikar, K. Chen, and M. Farach-Colton, ‘Finding Frequent Items in Data Streams’, in *Automata, Languages and Programming*, vol. 2380, P. Widmayer, S. Eidenbenz, F. Triguero, R. Morales, R. Conejo, and M. Hennessy, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 693–703, 2002.
- [11] S. Muthukrishnan, ‘Data Streams: Algorithms and Applications’, *FNT in Theoretical Computer Science*, vol. 1, no. 2, pp. 117–236, 2005.
- [12] G. Cormode and S. Muthukrishnan, ‘An improved data stream summary: the count-min sketch and its applications’, *Journal of Algorithms*, vol. 55, no. 1, pp. 58–75, Apr. 2005.
- [13] Z. Liu, A. Manousis, G. Vorsanger, V. Sekar, and V. Braverman, ‘One Sketch to Rule Them All: Rethinking Network Flow Monitoring with UnivMon’, in *Proceedings of the 2016 ACM SIGCOMM Conference*, Florianopolis Brazil, pp. 101–114, Aug. 2016.
- [14] H. Zhu, *Data Plane Development Kit (DPDK): A Software Optimization Guide to the User Space-based Network Applications*, 1st Edition. CRC Press, 2020.
- [15] T. Høiland-Jørgensen *et al.*, ‘The eXpress data path: fast programmable packet processing in the operating system kernel’, in *Proceedings of the 14th International Conference on emerging Networking EXperiments and Technologies*, Heraklion Greece, pp. 54–66, Dec. 2018.
- [16] M. Fleming, ‘A thorough introduction to eBPF’, *LWN.net Linux Weekly News*, Dec. 02, 2017. <https://lwn.net/Articles/740157/>
- [17] M. Zhang *et al.*, ‘Poseidon: Mitigating Volumetric DDoS Attacks with Programmable Switches’, in *Proceedings 2020 Network and Distributed System Security Symposium*, San Diego, CA, 2020.
- [18] M. Dimolianis, A. Pavlidis, and V. Maglaris, ‘Signature-Based Traffic Classification and Mitigation for DDoS Attacks Using Programmable Network Data Planes’, *IEEE Access*, vol. 9, pp. 113061–113076, 2021.
- [19] F. Erlacher and F. Dressler, ‘On High-Speed Flow-Based Intrusion Detection Using Snort-Compatible Signatures’, *IEEE Trans. Dependable and Secure Comput.*, vol. 19, no. 1, pp. 495–506, Jan. 2022.
- [20] Z. Liu *et al.*, ‘Jaqen: A High-Performance Switch-Native Approach for Detecting and Mitigating Volumetric DDoS Attacks with Programmable Switches’, in *30th USENIX Security Symposium (USENIX Security 21)*, pp. 3829–3846, 2021.
- [21] S. Myneni, A. Chowdhary, D. Huang, and A. Alshamrani, ‘SmartDefense: A distributed deep defense against DDoS attacks with edge computing’, *Computer Networks*, vol. 209, p. 108874, May 2022.
- [22] E. Viegas, A. Santin, A. Bessani, and N. Neves, ‘BigFlow: Real-time and reliable anomaly-based intrusion detection for high-speed networks’, *Future Generation Computer Systems*, vol. 93, pp. 473–485.

-
- ¹ Application
² HTTP/S
³ High-Bandwidth networks
⁴ Denial of Service
⁵ Distributed Denial Of Service(DDOS)
⁶ Comprehensive Processing
⁷ Adaptive Learning
⁸ Drop Rate
⁹ Internet Service Provider
¹⁰ Client
¹¹ Service
¹² Server
¹³ Attacker
¹⁴ Flash Coward
¹⁵ Data Streaming
¹⁶ Sketch
¹⁷ Switch
¹⁸ 5G
¹⁹ Metadata
²⁰ Payload
²¹ Header
²² Agent Machine
²³ Botnet
²⁴ Amplification DDoS Attack
²⁵ GitHub
²⁶ Memcached Distributed Caching Memory System
²⁷ Traffic peak
²⁸ Amazon Web Services
²⁹ Content Delivery Network (CDN)
³⁰ Cloudflare
³¹ Yandex
³² MikroTik
³³ Unpatched
³⁴ Cisco
³⁵ FBI
³⁶ Constrained Application Protocol(CAP)
³⁷ Attack Vector
³⁸ Batch Processing
³⁹ Stream Processing
⁴⁰ Line Rate Processing
⁴¹ Turnstile Model
⁴² Internet Protocol (IP)
⁴³ Port
⁴⁴ Per flow size
⁴⁵ Flow Moment
⁴⁶ Moment-g
⁴⁷ Heavy Hitter
⁴⁸ Sampling
⁴⁹ Count-Sketch
⁵⁰ Hash Function
⁵¹ Hash Collision
⁵² Count Min Sketch
⁵³ Universal Sketch
⁵⁴ Univmon sketch
⁵⁵ TCP/IP Stack
⁵⁶ Kernel
⁵⁷ Network Interface Card (NIC)
⁵⁸ Direct Memory Access (DMA)
⁵⁹ Data Plane Development Kit(DPDK)
⁶⁰ eXpress Data Path(XDP)
⁶¹ Intel
⁶² Open Source Project
⁶³ X86 Processor Architecture
⁶⁴ Application Specific Integrated Circuit (ASIC)
⁶⁵ Field Programmable Gate Array (FPGA)
⁶⁶ Polling Mode
⁶⁷ Network Function Virtualization(NFV)
⁶⁸ Driver
⁶⁹ Berkely Packet Filter (BPF)
⁷⁰ Tcpdump
⁷¹ Bytecode
⁷² Register
⁷³ Accumulator
⁷⁴ Index Register
⁷⁵ Program Counter
⁷⁶ Just-In-Time Compiler (JIT)
⁷⁷ MIPS Architecture
⁷⁸ ARM Architecture
⁷⁹ Extended BPF (EBPF)
⁸⁰ Opcode
⁸¹ System Call
⁸² Checkpoint
⁸³ Drop
⁸⁴ Forward
⁸⁵ Data-Plane Switch
⁸⁶ P4
⁸⁷ Attack Mitigation
⁸⁸ Middlebox
⁸⁹ Network Intrusion Detection
⁹⁰ Snort Intrusion Detection System
⁹¹ IPFIX
⁹² Anomaly
⁹³ Deep Packet Inspection(DPI)
⁹⁴ RT-SAD