دانشکده مهندسی کامپیوتر

سمینار کارشناسی ارشد گرایش رایانش امن

عنوان:

ارائه رویکرد تطبیق‌پذیر با تنوع ترافیکی شبکه‌های پهن‌باند برای شناسایی حملات منع خدمت توزیع‌شده

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:

باتوجه به گسترش روزافزون شبکه‌های کامپیوتری و متداول‌شدن استفاده از آنها، حجم تبادل اطلاعات نیز بالاتر رفته و امروزه نرخ‌گذر اطلاعات در بسیاری از تجهیزات شبکه به بیش از ۱۰۰ گیگابیت در ثانیه رسیده است. با توجه به گسترش شبکه‌ها، پروتکل‌های مختلفی برای کاربردها و برنامه‌های مختلف ارائه شده است و علاوه بر آن بسیاری از برنامه‌های کاربردی[۱] امروزه با این که از پروتکل‌های استاندارد و متداولی برای ارتباط با یکدیگر و تبادل اطلاعات در شبکه استفاده می‌کنند، شیوه استفاده از این پروتکل‌ها متفاوت می‌باشد. به عنوان مثال 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. از طرفی حملات منع خدمت توزیع‌شده[۳] به‌عنوان یکی از متداول‌ترین و پرهزینه‌ترین حملات در سطح شبکه شناخته می‌شود که موجب بروز اختلال در ارائه خدمات سطح سازمانی و حتی ملی شده است. یکی از اساسی‌ترین نیازهای امنیتی در سطح شبکه بحث دسترس‌پذیر[۴] بودن کامل شبکه می‌باشد. حملات منع خدمت به‌عنوان تهدیدی جدی برای قابلیت دسترس‌پذیری شبکه‌ها شناخته می‌شود.

در شبکه‌های پهن‌باند[۵] با افزایش نرخ ترافیک، چالش‌های امنیتی نظیر تشخیص حملات منع خدمت، که به دلیل سادگی در پیاده‌سازی و تاثیر بسیارمخرب یک تهدید جدی به حساب می‌آیند، افزایش پیدا کرده است. همچنین در این شبکه‌ها با تنوع پروتکلی زیادی روبرو هستیم و سیستم‌های تشخیص نفوذ در این شبکه ها به تعداد زیادی خط قوانین مبتنی بر امضا، نیاز خواهند داشت و در نتیجه توانایی پایش[۶] ترافیک به طور کامل و تشخیص حملات را نخواهند داشت. از این رو، مقابله با حملات منع خدمت در این شبکه‌ها، به یک بستر مهم تحقیقاتی در سال‌های اخیر تبدیل شده است. در دهه‌های گذشته محققان روش‌های شناسایی بسیاری را برای حملات منع خدمت توزیع‌شده پیشنهاد کرده‌اند. عدم تطبیق‌پذیری و مقیاس‌پذیری برای استفاده در شبکه‌های پهن‌باند، از متداول‌ترین مشکلات این روش‌ها هستند. لذا برای شناسایی صحیح حملات منع خدمت در شبکه‌های پهن‌باند نیاز به یک رویکردی است که شامل دو ویژگی پردازش جامع[۷] به معنای پردازش تمامی بسته‌ها و تطبیق‌پذیری[۸] 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 پهن‌باند 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 برمبنای 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 نظرگرفتن معیارهایی نظیر میزان استفاده از پردازشگر و حافظه، نرخ دور انداختن[۹] بسته‌ها، و میزان تاخیر در شناسایی حملات بررسی می‌شود.

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

۱. سرآغاز

امروزه با افزایش حجم تبادلات داده‌ای در بستر اینترنت، برقراریِ 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[۱۰] موظف هستند خدمات خود را به‌صورت شبانه‌روزی و بدون اختلال و وقفه در اختیار کارخواهان[۱۱] قرار دهند. درصورتی‌که 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[۱۲]، دسته‌ای از حملات در شبکه هستند که با هدف از بین بردن دسترس‌پذیری شبکه سعی در ممانعت از ارائه و انجام یک خدمت[۱۳] 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[۱۴] 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[۱۵] 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[۱۶] و ترافیک حمله‌کننده، پیدا کردن فرد مهاجم اصلی نیز به‌مراتب دشوارتر می‌شود.

ازسویی دیگر امروزه با شبکه‌های پهن‌باندی 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 پهن‌باند محسوب می‌شوند.

این گزارش در ۵ بخش تدوین شده است. در بخش ۲ مفاهیم پایه مورد نیاز در این پژوهش معرفی می‌شوند. ابتدا شبکه‌های پهن‌باند 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[۱۷] و انگاره‌ها[۱۸]، راهکارهای افزایش سرعت پردازش بسته‌ها و معرفی راه‌گزین[۱۹] 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 پهن‌باند، 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 پهن‌باند 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.

۲.۱ شبکه‌های پهن‌باند

امروزه نرخ تبادل اطلاعات در شبکه‌های کامپیوتری بالا رفته و مفهومی به عنوان شبکه‌های پهن‌باند مطرح می‌باشد. شبکه‌های پهن‌باند 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[۲۰]، 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[۲۱] می‌شوند. به عبارتی دیگر بسته‌هایی با محتوا[۲۲] و حجم زیادی از سرایندها[۲۳] 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 برگذاری یک کلاس مجازی باشد. در سال ۲۰۰۳، حجم کل داده‌های تولید شده در اینترنت حدود ۵ اگزابایت بود که این میزان در سال ۲۰۰۸ سه برابر شد و به ۱۴.۷ اگزابایت رسید. امروزه تقریباً ۵ اگزابایت داده در هر دو روز توسط کاربران تولید می شود[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.

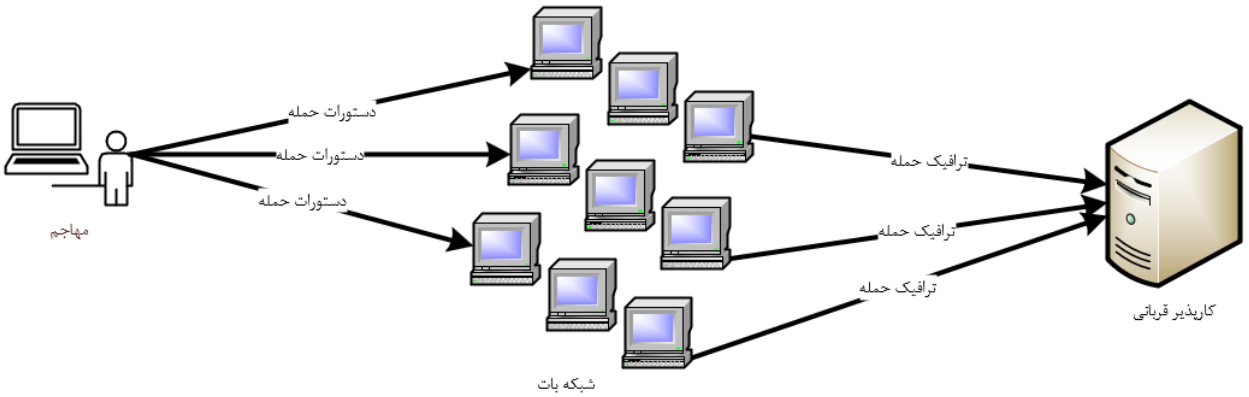
باتوجه به ویژگی‌های ذکر شده برای شبکه‌های پهن‌باند، مدیریت و کنترل ترافیک در این شبکه‌ها به یکی از چالش‌های اصلی در زمینه شبکه‌های کامپیوتری تبدیل شده است.

۲.۲ حملات منع خدمت (توزیع‌شده)

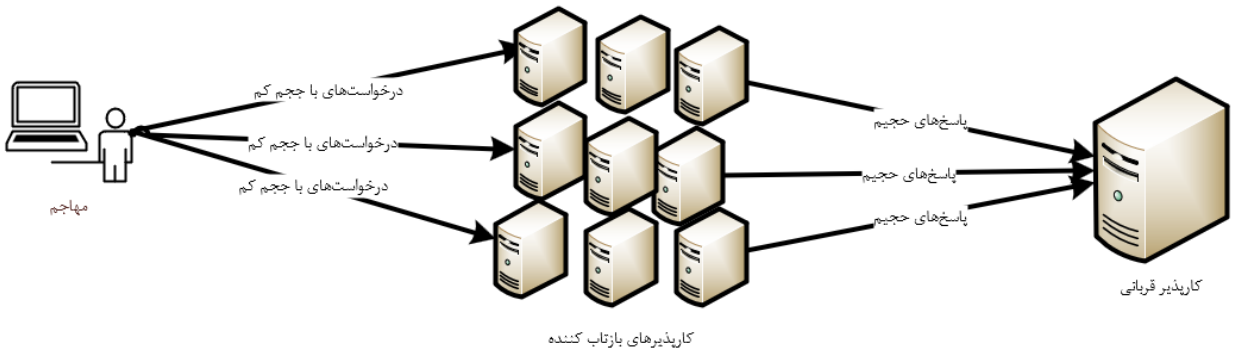
به مجموعه حملاتی که با هدف ممانعت از انجام یک خدمت صورت می‌پذیرند، حملات منع خدمت گفته می‌شود. این حملات با انگیزه‌های مختلفی نظیر ایجاد اختلال یا ممانعت از ارائه یک خدمت، از بین بردن اعتبار و مقبولیّت یک سازمان، آسیب زدن مالی و هدر دادن منابع یک سازمان، دستاوردهای سیاسی و ملی، انگیزه مالی و یا قدرت‌نمایی 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 این حملات می‌باشند که در آن فرد مهاجم ابتدا با پایش آسیب پذیری‌های دستگاه‌های مختلف موجود در شبکه اینترنت، شروع به نفوذ به ماشین‌های عامل[۲۴] 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[۲۵] گفته می‌شود، دستورات را از شخص مهاجم دریافت می‌کنند. 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.

مشکل دیگر دفاع در برابر حملات منع خدمت توزیع‌شده، بروز حملات تقویت بازتاب[۲۶] است. در سال ۲۰۱۸، گیتهاب[۲۷] با استفاده از آسیب‌پذیری پروتکل ممکچ[۲۸]، با انعکاس چند برابر بیش از ۵۰۰۰ بار و ترافیک پیک ۱.۳۵ ترابیت بر ثانیه، قربانی یک حمله منع خدمت توزیع‌شده از نوع تقویت بازتابی قرار گرفت. در فوریه ۲۰۲۰، ارایه دهندگان خدمات وب آمازون[۲۹] حمله‌ای با حجم ترافیک پیک ۲.۳ ترابایت بر ثانیه را تجربه کردند. در ژوئیه ۲۰۲۱، شرکت ارائه‌دهنده خدمات تحویل محتوا[۳۰]ی کلودفلر[۳۱] در گزارشی به محافظت از یکی از مشتریان خود در برابر حمله منع خدمت توزیع‌شده 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[۳۲]، ارائه‌دهنده خدمات پیشگیری از حملات منع خدمت توزیع‌شده روسیه گفت که در تاریخ ۵ سپتامبر ۲۰۲۱ یک حمله منع خدمت توزیع‌شده پروتکل اچ.تی.تی.پی را که از تجهیزات شبکه میکروتک[۳۳] بروزنشده[۳۴] سرچشمه می‌گرفت، مسدود کرده است.

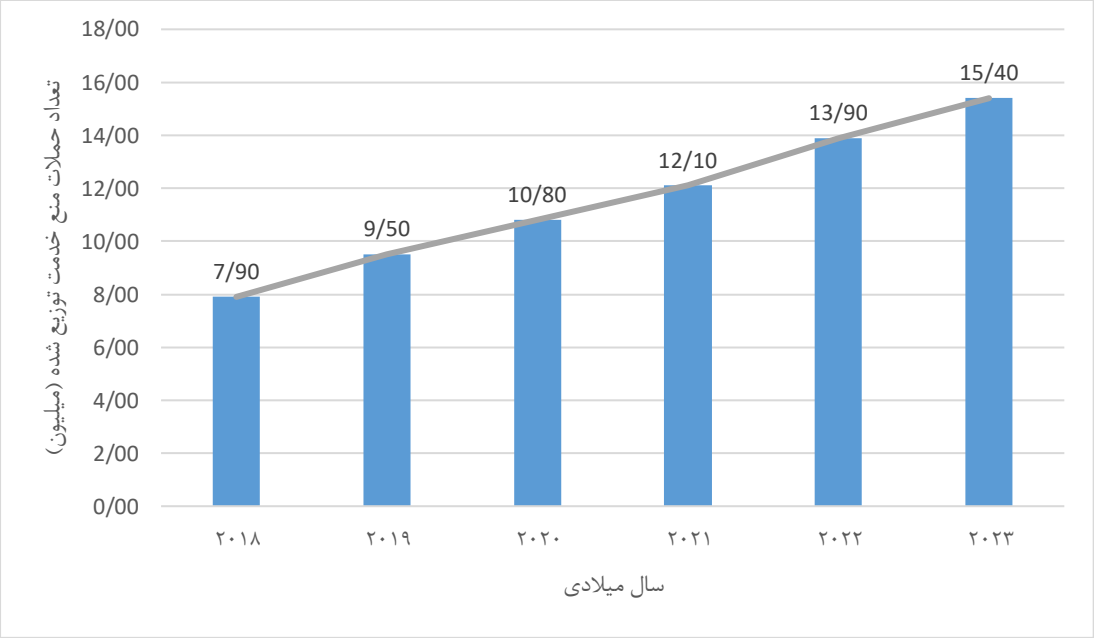


شکل ۱: حملات منع خدمت توزیع‌شده با استفاده از شبکه بات



شکل ۲: حملات منع خدمت توزیع‌شده از نوع تقویت بازتابی

طبق خلاصه سالانه آکامی[۳۵] برای سال ۲۰۲۰، تعداد حملات منع خدمت توزیع‌شده در مقیاس بزرگ به طور قابل توجهی افزایش یافته است. در بزرگترین حمله منع خدمت توزیع‌شده رخ داده تا به حال، ترافیک حمله به ۱.۴۴ ترابیت 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[۳۶] هشداری صادر کرد مبنی بر اینکه پروتکل برنامه‌های محدود شده[۳۷] و سایر پروتکل‌ها ممکن است برای انجام حملات منع خدمت توزیع‌شده مورد سوء استفاده قرار گیرند. حملات منع خدمت توزیع‌شده بر اساس بردارهای حمله[۳۸] 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].



جدول ۱: گزارش و پیش‌بینی سیسکو[۳۹] از مجموع حملات منع خدمت توزیع‌شده[۴]

۲.۳ داده جریان

همانطور که در ویژگی‌های شبکه‌های پهن‌باند 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[۴۰]: 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[۴۱]: اکثر راهکارهای ارائه شده که در قسمت بعد بررسی می‌شوند، مبتنی بر این رویکرد می‌باشند. این الگوریتم‌ها دو مشخصه زیر را درنظر 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[۵]. این خصیصه‌ها همان چالش‌هایی هستند که برای پردازش ترافیک در شبکه‌های پهن‌باند 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[۴۲] 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 ترنستیل[۴۳] می‌باشد. در این مدل یک داده جریان ورودی به نام I درنظر گرفته می‌شود که شامل مجموعه‌ای از تاپل‌های دوتایی می‌باشد:

$$I = \alpha_1, \alpha_2, \alpha_3, \alpha_4, \ldots$$

$$\forall E \alpha_i = \{(a_1, v_1) \mid a_1 \in \{0, 1, \ldots, u-1\}, v_1 \in R\}$$

$$[u] = key\ space$$

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:

$A[a_x] \mathrel{+}= v_x$

این پارامترها وابسته به مسأله داده جریانی که مطرح می‌شود، می‌توانند متفاوت باشند. دربحث پردازش بسته‌های دریافتی شبکه، جریان همان جریان ورودی و تاپل‌ها همان بسته‌ها می‌باشند که برای مثال کلیدشان ۵ خصیصه‌ی آدرس آی.پی[۴۴] مبدأ، آدرس آی.پی مقصد، شماره درگاه مبدأ، شماره درگاه[۴۵] مقصد و پروتکل و به‌روزرسانی نیز می‌تواند اندازه بسته باشد. درنتیجه برای شناسایی حملات منع خدمت، باید آدرس‌هایی که بسته‌هایی با حجم نامتعارف ارسال می‌کنند شناسایی کرد[۶].

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

$(srcIP, srcport, dstIP, dstport, protocol)$=شناسه جریان تعداد کل جریان‌های متمایز $=F$

داده جریان ورودی $= (f_1,c_1),\ldots,(f_t,c_t),\ldots$

پاسخ‌های مسائل یکی از انواع زیر می‌باشند[۷]:

- سایز هر جریان[۴۶]: خواسته این مسائل، یافتن سایز جریان یا تعداد بسته‌های دریافت شده متعلق به جریان f می‌باشد که با nf نشان داده می‌شود. سایز تمامی بسته های دریافتی نیز $n = \sum_{1<i\leq F} n_f$ می‌باشد.
- لحظه جریان[۴۷]: در لحظه دلخواهی، وضعیت جریان با استفاده از تابع g در لحظه g[۴۸] 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 این‌که به چه صورت تعریف شده باشد، می‌تواند وضعیت کلی از ترافیک شبکه را به صورت عددی بیان کند.

- شاخص[۴۹]: شاخص‌ها جریان‌هایی هستند که اندازه آنها بر لحظه جریان L_g بیشترین تاثیر را می‌گذارد. به عبارتی دیگر:

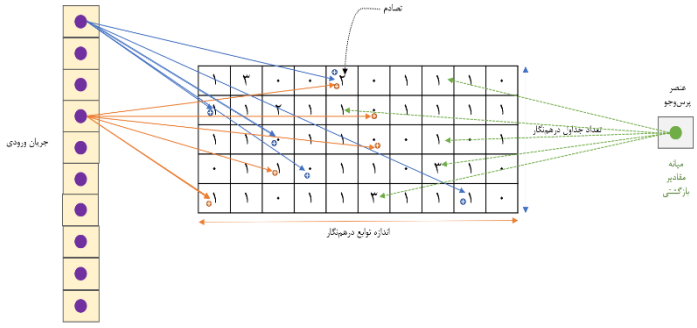
$$H_g = \{\, f \mid g(nf) \geq \alpha L_g \}$$

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

۲.۴ انگاره

برای حل مسائل داده جریان، راهکارهای متفاوتی را می‌توان استفاده کرد. یکی از راهکارها بدین صورت می‌باشد که به دلیل اینکه با حجم زیادی از اطلاعات روبرو هستیم، تنها بخشی از داده‌های ورودی به عنوان نمونه انتخاب شوند و عملیات پردازش تنها روی آن‌ها صورت گیرد. این روش نمونه برداری[۵۰] نامیده می‌شود[۸]. نمونه‌برداری دقت پایینی خواهد داشت. به منظور بالابردن دقت، پردازش تمامی بسته‌ها الزامی می‌باشد. اما بررسی همه بسته‌ها نیز نیازمند حجم زیادی از منابع پردازشی و زمان می‌باشد. برای حل این مشکل، الگوریتم‌هایی به نام انگاره ارائه شده‌اند که از یک داده ساختار فشرده برای ذخیره سازی اطلاعات داده‌های ورودی استفاده می‌کنند. انواع مختلفی از این الگوریتم‌ها در پژوهش‌های مختلف ارائه شده است که هر کدام سعی در حل یکی از انواع مسائل داده جریان دارند. در ذیل چند مورد از پراستفاده‌ترین آن‌ها معرفی خواهند شد:

- انگاره شمارشی[۵۱]: از یک جدول KxH تشکیل شده است که شامل K تابع درهم‌نگار[۵۲] می‌باشد. H نیز اندازه توابع درهم‌نگار در یک سطر می‌باشد. این ساختار را الگوریتم‌های انگاره دیگر نیز استفاده می‌کنند. چون از توابع درهم‌نگار استفاده می‌شود لذا امکان تصادم[۵۳] وجود خواهد داشت. مقادیر بازگشتی تخمینی خواهند بود و درنتیجه به آنها داده‌ساختارهای آماری احتمالاتی می‌گویند. اما بایستی نرخ خطای قابل قبول و کرانداری ارائه دهند. از این الگوریتم برای یافتن پاسخ مسائل شاخص استفاده می‌شود[۹].



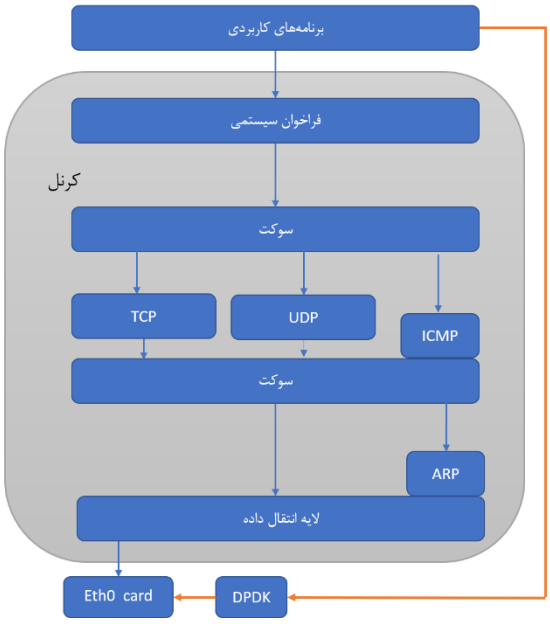
شکل ۳: انگاره شمارشی

- انگاره شمارشی کمینه[۵۴]: همانند انگاره شمارشی می‌باشد اما سعی دارد مرتبه فضایی را کاهش دهد[۱۰].
- انگاره عمومی[۵۵]: یک دسته جدیدی از انگاره‌ها با هدف ارائه داده‌ساختاری قابل استفاده برای حل تمامی انواع مسائل داده جریان می‌باشند. یونیومان[۵۶] یکی از این الگوریتم‌ها می‌باشد[۱۱].

۲.۵ پردازش سریع بسته‌ها

هنگامی که یک بسته از طریق واسطهای شبکه یک سیستم دریافت می‌شود تا پردازش آن، مراحل مختلفی را طی خواهدکرد. بنا به کاربرد، بسته‌ها از دستگاه‌های مختلفی عبور داده می‌شوند. برمبنای پشته پروتکل تی.سی.پی/آی.پی[۵۷] که کرنل تمامی سیستم‌های عامل لینوکس از آن پشتیبانی می‌کنند، بسته‌ها از دریافت تا پردازش بخش‌های مختلف آن‌ها مراحل مختلفی را پشت سر خواهند گذاشت و درنهایت در صورت نیاز بازارسال خواهند شد. رویدادهای مهم در هنگام دریافت یک بسته توسط ماشین بدین شرح می‌باشد:

- بسته توسط کارت شبکه[۵۸] ماشین دریافت می‌شود (وقفه کارت شبکه) .
- کارت شبکه از طریق دی.ام.ای[۵۹]، بسته را در فضای حافظه در یک بافر قرار می‌دهد.
- کارت شبکه یک سیگنال به پردازنده می‌دهد، 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 کرنل 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 کرنل به فضای کاربر انتقال داده خواهد شد. در غیر اینصورت، بسته در همان فضای کرنل خواهد ماند.



شکل ۴: مراحل ضبط و پردازش بسته

تمامی این مراحل بایستی در سطح کرنل 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 پهن‌باند که با حجم زیادی از بسته‌ها مواجه هستیم، باعث اتلاف وقت زیادی خواهند شد.

دی.پی.دی.کی[۶۰] و ایکس.دی.پی[۶۱] از ابزارهای موجود برای تسریع عملیات پردازش بسته می‌باشند. دی.پی.دی.کی ابزار نرم افزاری می‌باشد که در سال ۲۰۰۹ توسط اینتل[۶۲] توسعه داده شد. ولی بعدها به صورت یک پروژه متن باز[۶۳] درآمد. به طور خلاصه یک ابزار دورزدن کرنل[۶۴] در هنگام دریافت بسته در شبکه می‌باشد که وقفه‌های مختلف مربوط به کرنل را حذف می‌کند و لذا تمام عملیات پردازش بسته را می‌توان در سطح کاربر انجام داد و درنهایت عملیات دریافت و پردازش بسته را تا حد خوبی می‌تواند تسریع بخشد. هدف این فناوری استفاده از قابلیت پردازش چندهسته‌ای پردازنده‌های معمولی ایکس۸۶[۶۵] برای بهبود سرعت پردازشی کارپذیرها می‌باشد. بدین صورت ما نرخ پردازشی برابر هنگام استفاده از پردازنده‌های مخصوص کارپذیرها و یا مدارهای مجتمع با کاربرد خاص[۶۶] و مدار مجتمع دیجیتال برنامه‌پذیر[۶۷] ، 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[۶۸] 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[۶۹]، 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[۱۲].

ایکس.دی.پی یکی از کامپوننت‌های جدید کرنل می‌باشد که پردازش بسته را به صورت خوبی بهبود می‌بخشد. روش‌هایی مثل دی.پی.دی.کی کرنل را دور می‌زنند و تمام عملیات پردازش بسته در فضای کاربر صورت می‌گیرد. همچنین کارت شبکه را باید توسط یک درایور 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 درایورهای مربوطه را خودشان توسعه دهند.
- برنامه‌های سطح کاربر می‌بایست درصورت نیاز عملکردهایی که توسط کرنل ایجاد می‌شد را پیاده‌سازی کنند.
- برنامه‌ها به صورت ایزوله اجرا می‌شوند که نحوه تعامل آنها با دیگر بخش‌های سیستم‌عامل را دشوار می‌کند.

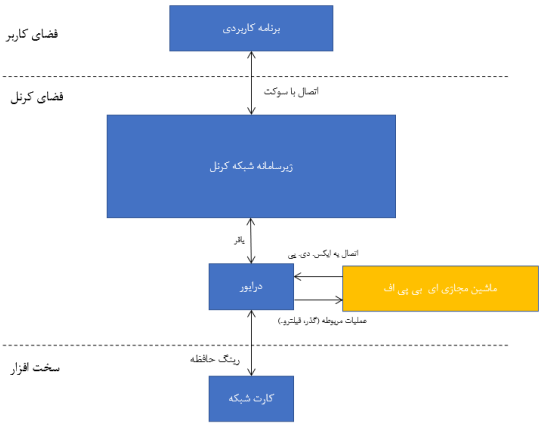
به طور خلاصه ایکس. دی. پی، برنامه‌های شبکه سطح کاربر (پالایش، نگاشت، مسیریابی و...) را به جای انتقال به سطح کاربر، به فضای کرنل می‌برد. ایکس.دی.پی امکان اجرای برنامه به محض ورود بسته به کارت شبکه و پیش از حرکت به سمت زیرسیستم شبکه‌ی هسته را فراهم می‌کند که منجر به افزایش قابل توجه سرعت پردازش بسته می‌شود. اجرای برنامه در سطح کرنل با استفاده از بی. پی. اف[۷۰] میسر می‌شود[۱۳].

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[۷۱] می‌باشد. عبارت پالایش مربوطه توسط یک کامپایلر به بایت‌کد[۷۲] 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[۷۳] (A و X یعنی انباشتگر[۷۴] و ثبات اندیس[۷۵]) ، یک حافظه موقت و یک شمارنده برنامه[۷۶] ضمنی نیز می‌باشد را تعریف می‌کند. کرنل لینوکس از نسخه ۲.۵ به بعد از بی.پی.اف پشتیبانی می‌کند. در سال ۲۰۱۱، مفسر بی.پی.اف به یک کامپایلر درجا[۷۷] تغییر داده شد. این کار باعث شد که کرنل به جای تفسیر برنامه‌های بی.پی.اف، قادر باشد که آن‌ها را به یک معماری هدف میپس[۷۸]، ایکس۸۶، آرم[۷۹] تبدیل کند. این امر به معرفی بی.پی.اف توزیع یافته[۸۰] 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 رجیسترها و تعداد آن‌ها و کدهای عملیاتی[۸۱] 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 کرنل مورد استفاده قرار می‌گیرند.
- استفاده از توابع کمکی که در سطح کرنل اجرا می‌شوند. امکان فراخوانی فراخوان سیستمی[۸۲] در برنامه‌های بی.پی.اف نیز وجود دارد.

- زنجیره سازی تعداد برنامه بیشتر بی.پی.اف نیز امکان‌پذیر خواهد بود.

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[۸۳] در پشته کرنل تعریف کرد که هرگاه بسته‌ای در کارت شبکه دریافت شد، آن را به فضای کاربر بفرستد و درآنجا تصمیم می‌گیرد که بسته دور انداخته شود[۸۴] یا اجازه عبور به لایه‌های بالاتر پشته را صادر کند. لذا نیاز به مکانیزمی بود که امکان اجرای کدهای سطح کاربر را در کرنل فراهم کند. به همین دلیل از بی.پی.اف توزیع یافته استفاده شد.

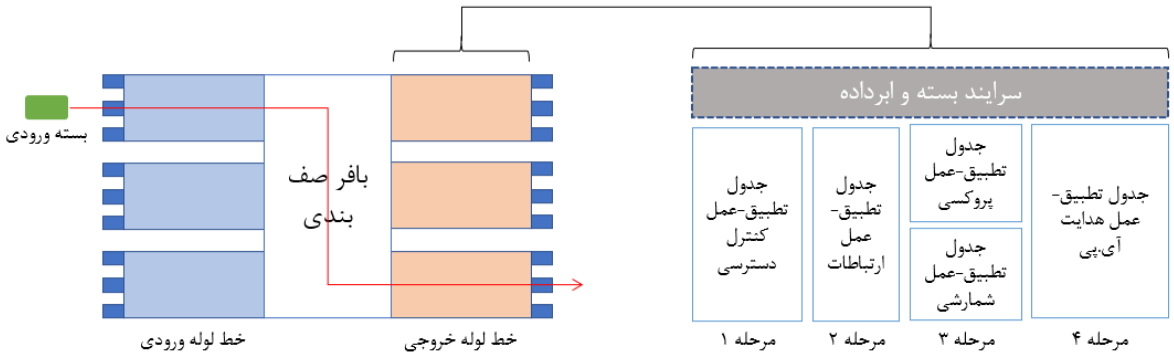


شکل ۵: ایکس.دی.پی بر روی کرنل

ایکس.دی.پی بسته‌های دریافتی را به برنامه بی.پی.اف هدایت می‌کند. در آنجا می‌توان بسته‌ها را ویرایش و یا هدایت[۸۵] 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 تعبیه‌شده در ایکس.دی.پی می‌توان عمل مورد نظر را بر روی بسته انجام داد[۱۴].

۲.۶ راه‌گزین‌های برنامه‌پذیر

استفاده از یک کنترلر 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 آسیب پذیر واحد برای مهاجمین فراهم کند. اما امروزه با معرفی راهگزین‌های برنامه پذیر[۸۶]، راهگزین‌های معمولی نیز، با استفاده از برنامه‌هایی که بر روی آنها با استفاده از زبان‌هایی مثل پی۴[۸۷] 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 واسطهای 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 پی۴ 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[۱۵].



شکل ۶: معماری راه‌گزین برنامه‌پذیر

۳. کارهای پیشین

به‌طورکلی پژوهش‌های انجام‌شده در حوزه حملات منع خدمت توزیع‌شده را می‌توان در سه دسته پیشگیری از وقوع حمله، تشخیص حمله و کاهش اثر حمله[۸۸] تقسیم‌بندی کرد. از آنجایی‌که 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 طبقه بندی 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[۸۹].

۳.۱ روش‌های مبتنی بر امضا

آنتروپی[۹۰] معیاری است که برای اندازه‌گیری میزان تصادفی بودن یک ویژگی در یک دوره زمانی معین استفاده می‌شود. روش‌های مبتنی بر آنتروپی به‌عنوان یک رویکرد مؤثر برای محاسبه تصادفی از یک مجموعه داده طراحی شده‌اند. به‌طورکلی مقادیر بالای آنتروپی نشان‌دهنده توزیع پراکنده‌تر ویژگی در دادگان[۹۱] موجود است و مقادیر پایین آنتروپی نشان‌دهنده نامتوازن بودن یک توزیع است. به عبارت دیگر برخی مقادیر ویژگی موردنظر، فراوانی بیشتری نسبت به سایر مقادیر دارند. از این معیار برای تشخیص ناهنجاری گسترده در سامانه‌های سنتی تشخیص نفوذ[۹۲] استفاده شده است. به‌منظور تشخیص حملات منع خدمت، آنتروپیِ جریان شبکه را می‌توان با استفاده از چندین ویژگی مانند جریان شبکه، آدرس آی.پی مبدأ و مقصد بسته‌ها و یا تعداد بسته‌های موجود در یک جریان محاسبه کرد. سپس با مقایسه با یک حد آستانه از پیش تعریف شده در مورد عادی یا غیرعادی بودن جریان بررسی‌شده، می‌توان تصمیم‌گیری کرد. یکی از مهم‌ترین مزیت‌های این روش داشتن سربار محاسباتی کم می‌باشد.

- 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 امضاهای مهاجم را بدست آورند و تعداد حداقل بهینه خط قوانین به منظور مقابله با آنها را تولید کنند[۱۶]. 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[۹۳]: در روش ارائه شده در سال ۲۰۲۰ توسط ارلاکر و همکاران، از دسته‌بندی جریان مبتنی بر آی.پی.فیکس[۹۴] 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[۱۷]. مشکل این روش عدم کارایی در شناسایی حملات مختلف می‌باشد.

۳.۲ روش‌های مبتنی بر مدل‌سازی

این دسته از روش‌ها با ضبط‌کردن و بررسی ترافیک عادی شبکه در یک بازه زمانی، رفتار عادی شبکه را شبیه سازی یا به اصطلاح مدل می‌کنند و هرگونه رفتار مغایر با این مدل یا اصطلاحاً آنومالی[۹۵] را به عنوان حمله در نظر می‌گیرند. یادگیری ماشین به‌عنوان یکی از روش‌های کارآمد مبتنی بر مدل‌سازی می‌باشد، که امروزه به صورت گسترده‌ای مورد استقبال پژوهشگران قرارگرفته 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 مورد استفاده قرار می‌گیرند.

- رویکرد مبتنی بر 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 برروی این راهگزین‌ها می‌توان پیاده کرد. به دلیل این که با استفاده از زبان پی۴، الگوریتم‌های تشخیص و رفع مخاطره را پیاده می‌کنیم، لذا این روش مبتنی بر معماری خاصی از راهگزین‌ها نمی‌باشد[۱۸]. عدم وارسی محتوای کامل بسته‌ها و استفاده از وارسی‌کننده عمیق بسته[۹۶]، 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 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]. اما در مورد نحوه انتخاب این ویژگی‌ها برای برنامه‌های کاربردی مختلف و البته انجام این‌کار به صورت پویا صحبتی نمی‌کند. در مورد محل پیاده‌سازی این الگوریتم‌ها نیز توضیحی نمی‌دهد. به دلیل استفاده از انگاره‌ها، آر.تی.سد از نظر مرتبه فضایی بسیار بهینه می‌باشد.
- دفاع هوشمند: 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 فراهم کننده اینترنت برای شناسایی حملات استفاده می‌کند[۱۹]. اما مشکل این روش، عدم ارائه راهکاری بهینه به منظور آموزش مجدد شبکه‌ها در شبکه‌های پهن‌باند می‌باشد.
- روش بلادرنگ و قابل اعتماد مبتنی بر آنومالی برای تشخیص رخنه در شبکه‌های با سرعت بالا: با درنظرگرفتن بسته‌ها به عنوان جریان، ویژگی‌های آن‌ها را استخراج می‌کند و سپس از روی آنها تشخیص می‌دهد. یک بخش اعتمادسازی دارد که میزان قابل اعتمادبودن 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 پهن‌باند به مشکل خواهد خورد و سربار بالا و دقت پایینی خواهد داشت.

۴. روش پیشنهادی

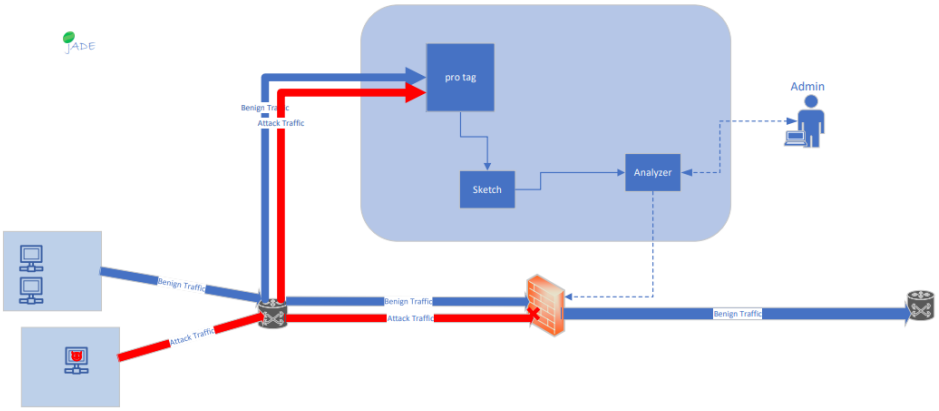
در بخش قبل برخی روش‌های مبتنی بر امضا و مدل‌سازی به منظور تشخیص حملات منع خدمت توزیع‌شده معرفی شدند. به عنوان نتیجه می‌توان گفت، روش‌های شناسایی و مقابله با حملات منع خدمت توزیع‌شده، ویژگی‌های ترافیک را از سه منظر بررسی می‌کنند (ترافیک را از سه منظر مشاهده می‌کنند) و سعی در مقابله دارند:

- بسته: یعنی جلوی یک بسته خاص مثلاً اچ. تی.تی.پی را می‌گیرند.
- جریان: یعنی جلوی یک جریان که مثلاً اندازه آن بیش از ۱۰۰ کیلوبایت باشد را می‌گیرند.
- رفتار کاربر: اگر رفتار ترافیک کاربری نامتعارف بود، جلوی آن را می‌گیرند. مثلاً در یک دقیقه، بیش از ۱۰ درخواست به منابع مختلف یک سایت ارسال کند.

در هر یک ازمنظرها طبق اطلاعاتی که جمع‌آوری می‌کنند، ترافیک نامتعارف را با استفاده از رویکردهای مبتنی بر محاسبات آماری یا با استفاده از رویکردهای مدل‌سازی و تشخیص آنومالی، می‌توانند تشخیص دهند. اما عدم سازگاری با تنّوع 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 هر کاربردی نمی‌توان یک الگو، مرز و شناسه برای حالت متعارف آن تعریف نمود. برای حل این مشکل به شناسایی کاربردهای مختلف می‌پردازند که از روش‌هایی مانند وارسی 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 پهن‌باند 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 وارسی‌کننده عمیق بسته، سربار محاسباتی زیاد دارند.

همانطور که گفته شد، نکته‌ای که در پژوهش های پیشین نادیده گرفته می‌شد، مربوط به مولفه سوم شبکه‌های پهن‌باند یا همان تنوع ترافیکی می‌باشد. در روش‌های پیشین مولفه‌های 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 درنظر گرفته بودند. لذا روشی که ارائه می‌دهیم تمامی این سه مورد را با تمرکز بیشتر بر روی ویژگی سوم به

عنوان مسئله اصلی را هدف قرار می‌دهد. روشی که ارائه می‌دهیم از ویژگی پردازش جامع برخوردار می‌باشد، یعنی تمامی بسته‌ها را یک و تنها یکبار بررسی می‌کند. و بدین صورت روشی بسیار سریع و با دقت بالا و تطبیق‌پذیر با مشخصات ترافیکی شبکه‌های پهن‌باند ارائه می‌دهیم (علاوه بر معیارهای متداولی مثل سرعت-نرخ‌گذر بالا و تأخیر کم که خواسته همه روش‌های قبلی بوده است) که راه‌حل نوینی می‌باشد.



شکل ۷: شمای کلی از روش پیشنهادی

روال کاری ما بدین صورت خواهد بود که با استفاده از روش ارائه شده در مقاله آکوپی و همچنین استفاده از یک زیرسامانه وارسی‌کننده عمیق بسته، جریان‌ها را برچسب‌گذاری می‌کنیم و جریان‌های شبیه به هم از نظر رفتار را در یک گروه قرار می‌دهیم و این اطلاعات را در داده‌ساختارهای انگاره که بر روی راهگزین‌های برنامه‌پذیر می‌باشند و توسط مدیر شبکه کنترل می‌شوند، ذخیره می‌کنیم. این اطلاعات را برای هر برنامه کاربردی و پروتکل متناظر به صورت جدا ذخیره می‌کنیم. یک قسمت تشخیص داریم 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 امضای معادل آن را تولید کند و به عنوان خروجی به یک دیوار آتش ارسال کند.

۵. نتیجه‌گیری

در این نوشتار به مرور مفاهیم اولیه مرتبط با حملات منع خدمت توزیع‌شده و انواع آن، شبکه‌های پهن‌باند و ویژگی‌های این شبکه، روش‌ها و الگوریتم‌های پردازش و معرفی مفاهیم و واژه‌های به کاررفته در این زمینه، مانند محاسبات و مسائل داده جریان پرداخته شد. سپس برخی پژوهش‌های انجام‌شده درزمینه‌ی تشخیص حملات منع خدمت در شبکه‌های پهن‌باند مورد بررسی قرار گرفت و مشکلات پیاده‌سازی و عملکردی و چالش‌های حل‌نشده آنها بیان شد. در آخر 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 پهن‌باند به صورت مختصر توضیح داده شد.

جدول ۲: 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, doi: 10.1016/j.comnet.2019.106943.

[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, doi: 10.1016/j.comcom.2019.06.010.

[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, doi: 10.1155/2021/9409473.

[4] K. Machap and H. Qiang, ‘Evaluating firewall tools and techniques in enhancing network security’, vol. 6, pp. 1–4, Jan. 2022.

[5] 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, Aug. 2021, pp. 2285–2293. doi: 10.1145/3447548.3467353.

[6] 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, 2003, p. 234. doi: 10.1145/948205.948236.

[7] 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, Jul. 2020, pp. 974–983. doi: 10.1109/INFOCOM41043.2020.9155454.

[8] 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, Apr. 2017, pp. 164–176. doi: 10.1145/3050220.3063772.

[9] 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, 2002, pp. 693–703. doi: 10.1007/3-540-45465-9_59.

[10] 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, doi: 10.1016/j.jalgor.2003.12.001.

[11] 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, Aug. 2016, pp. 101–114. doi: 10.1145/2934872.2934906.

[12] H. Zhu, *Data Plane Development Kit (DPDK): A Software Optimization Guide to the User Space-based Network Applications*, 1st Edition. CRC Press, 2020.

[13] 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, Dec. 2018, pp. 54–66. doi: 10.1145/3281411.3281443.

[14]M. Fleming, 'A thorough introduction to eBPF', *LWN.net Linux Weekly News*, Dec. 02, 2017. https://lwn.net/Articles/740157/

[15]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. doi: 10.14722/ndss.2020.24007.

[16]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, doi: 10.1109/ACCESS.2021.3104115.

[17]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, doi: 10.1109/TDSC.2020.2973992.

[18]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)*, 2021, pp. 3829–3846.

[19]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, doi: 10.1016/j.comnet.2022.108874.

[20]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, Apr. 2019, doi: 10.1016/j.future.2018.09.051.

واژه‌نامه

[1] Application
[2] HTTP/S
[3] Distributed Denial Of Service attack (DDOS)
[4] Availability
[5] High-Bandwidth
[6] Monitor
[7] Comprehensive Processing
[8] Adaptive Learning
[9] Drop Rate
[10] Internet Service Provider
[11] Client
[12] Denial Of Service
[13] service
[14] Server
[15] Attacker
[16] Flash Coward
[17] Data Streaming
[1818] Sketch
[19] Switch
[20] 5G
[21] Metadata
[22] Payload
[23] Header
[24] Agent Machine
[25] Botnet
[26] Amplification DDoS Attack
[27] GitHub
[28] Memcached Distributed Caching Memory System
[29] Amazon Web Services
[30] Content Delivery Network (CDN)
[31] Cloudflare
[32] Yandex
[33] MikroTik
[34] Unpatched
[35] Akamai Technologies
[36] FBI
[37] Constrained Application Protocol(CAP)
[38] Attack Vector
[39] CISCO
[40] Batch Processing
[41] Stream Processing

[42] Line Rate Processing
[43] Turnstile Model
[44] Internet Protocol (IP)
[45] Port
[46] Per Flow Size
[47] Flow Moment
[48] Moment-g
[49] Heavy Hitter
[50] Sampling
[51] Count-Sketch
[52] Hash Function
[53] Hash Collision
[54] Count Min Sketch
[55] Universal Sketch
[56] Univmon
[57] TCP/IP Stack
[58] Network Interface Card (NIC)
[59] Direct Memory Access (DMA)
[60] Data Plane Development Kit(DPDK)
[61] eXpress Data Path(XDP)
[62] Intel
[63] Open Source Project
[64] Kernel Bypass
[65] X86 Processor Architecture
[66] Application Specific Integrated Circuit (ASIC)
[67] Field Programmable Gate Array (FPGA)
[68] Polling Mode
[69] Network Function Virtualization(NFV)
[70] Berkely Packet Filter (BPF)
[71] Tcpdump
[72] Bytecode
[73] Register
[74] Accumulator
[75] Index Register
[76] Program Counter
[77] Just-In-Time Compiler (JIT)
[78] MIPS Architecture
[79] ARM Architecture
[80] Extended BPF (EBPF)
[81] Opcode
[82] System Call
[83] Checkpoint
[84] Drop
[85] Forward
[86] Data-Plane Switch
[87] P4
[88] Attack Mitigation
[89] Middlebox
[90] Entropy
[91] Dataset
[92] Intrusion Detection System
[93] Snort Intrusion Detection System
[94] IPFIX
[95] Anomaly
[96] Deep Packet Inspection(DPI)