مجموعه مقالات پانزدهمین کنفرانس بین‌المللی انجمن رمز ایران
دانشکده مهندسی برق، دانشگاه تربیت دبیر شهید رجایی، تهران، ۶-۷ شهریور ۱۳۹۷

ردگیری IP مبتنی بر نشانه‌گذاری جریان در سطح AS

مرتضی ارجمندپناه ، سبحان علی‌آبادی ، حمیدرضا محروقی

گروه مهندسی کامپیوتر دانشگاه امام رضا (ع) ، مشهد
Arjmand0morteza@gmail.com

گروه مهندسی کامپیوتر دانشگاه امام رضا (ع) ، مشهد
sobhan.aliabady@gmail.com

استادیار ، گروه مهندسی کامپیوتر دانشگاه امام رضا (ع) ، مشهد
Mahrooghi@imamreza.ac.ir

چکیده

امروزه یکی از مسائل چالش برانگیز در دنیای امنیت شبکه، مقابله با حملات منع سرویس توزیع شده است، این نوع از حملات با ارسال حجم زیادی از ترافیک به سمت یک میزبان یا سرویس‌دهنده، دسترسی‌پذیری آن را محدود و یا سرویس‌دهی به دیگر کاربران شبکه را مختل می‌کند. مهم‌ترین چالش موجود برای مقابله با این نوع از حملات، پیدا کردن مبدا سیل ترافیکی بسته‌ها می‌باشد که بدین منظور روش‌های ردگیری IP ارائه شده‌اند.

روش ردگیری IP ارائه شده از سازوکار توزیع امضای مسیر برای شناسایی ترافیک‌های جعل شده استفاده می‌کند. در این روش، سیستم‌های خودمختار همسایه‌های مشابه خود را که از روش پیشنهادی پشتیبانی می‌کنند را، شناسایی و اطلاعات مسیریابی، احراز هویت و همچنین امضای مسیر بین یکدیگر را به اشتراک می‌گذارند. برای انتقال داده‌های ردگیری از سازوکار نشانه‌گذاری در سطح جریان استفاده شده است. در مقایسه با روش‌های موجود نتایج نشان می‌دهد که، روش ارائه شده دارای کم‌ترین نرخ نشانه‌گذاری، سربار پردازشی در گره‌های میانی و سربار محاسباتی در گره مقصد است، همچنین بالاترین میزان دقت در ردگیری مبدا حملات را داراست.

کلمات کلیدی

ردگیری IP، نشانه‌گذاری جریان، جعل آدرس IP، حملات منع سرویس توزیع شده، سیستم خودمختار

۱- مقدمه

امروزه با گسترش روزافزون شبکه‌های کامپیوتری و زیرساخت‌های ارائه دهنده آن‌ها و همچنین تولید میلیون‌ها برنامه‌ی کاربردی، حملات منع سرویس توزیع شده[۱] (DDoS) یکی از اصلی‌ترین و چالش برانگیزترین تهدیدات شبکه‌های کامپیوتری محسوب می‌شود. با توجه به افزایش تعداد و حجم این حملات در سال‌های آینده، مقابله با این نوع از حملات یک چالش اساسی برای محققان می‌باشد زیرا: ۱. مسیریابی بسته‌های پروتکل اینترنت[۲] (IP) مبتنی بر آدرس مقصد می‌باشد و ۲. بسته‌های IP احراز هویت نمی‌شوند و مهاجم قادر است آدرس مبدا بسته‌ها[۳] را جعل[۴] و از ردگیری آن جلوگیری کند. در نتیجه روش‌های ردگیری IP[۵] راه‌حلی برای شناسایی و ردگیری مبدا بسته‌ها و مسیر پیموده شده توسط آن ارائه می‌دهند.

روش ردگیری ارائه شده از تکنیک نشانه‌گذاری برای انتقال داده‌های ردگیری استفاده می‌کند. ایده اصلی این روش، تزریق داده‌های ردگیری به چند فیلد[۶] از سرآیند بسته‌های IP و انتقال آن‌ها به همراه بسته‌های IP به گره مقصد است. در نتیجه در گره مقصد با دریافت داده‌های ردگیری و

پردازش آن‌ها می‌توان گراف شبکه‌ی[۷] پیموده شده (در مقاله از آن به عنوان امضای مسیر[۸]یاد می‌شود) را صرف نظر از اینکه بسته‌ها جعل شده‌اند، بازسازی کرد. بدیهی است نشانه‌گذاری در سطح بسته[۹]برای تمام ترافیک‌های جریان، موجب ایجاد سربار در گره‌های میانی می‌شود، لذا نشانه‌گذاری در روش پیشنهادی در سطح جریان[۱۰]ارائه شده است که در عمل سربار پردازشی گره‌های میانی و گره مقصد که داده‌های ردگیری در آن جمع‌آوری و پردازش می‌شود را به صورت قابل توجهی کاهش می‌دهد. همچنین روش پیشنهادی در سطح سیستم خودمختار[۱۱](AS) ارائه شده است، به اینصورت که تنها مسیریاب‌های مرزی سیستم‌های خودمختار[۱۲](ASBR) جریان‌های عبوری را نشانه‌گذاری می‌کنند و از آنجایی که نشانه‌گذاری در سطح جریان است، تعداد محدودی بسته به نمایندگی از تمام بسته‌های جریان نشانه‌گذاری می‌شوند، در نتیجه روش ارائه شده دارای کم‌ترین نرخ نشانه‌گذاری نیز می‌باشد.

سهم علمی مقاله، در ادامه بیان می‌شود:

۱. توپولوژی شبکه‌ی تحت پوشش که امری بسیار با اهمیت است، آشکار نمی‌شود، در حالی که روش‌های سطح مسیریاب، توپولوژی شبکه‌ی تحت پوشش را فاش می‌کنند، به اینصورت که اگر مسیریاب‌های موجود در AS، در روند نشانه‌گذاری حضور داشته باشند، آنگاه با مشاهده نشانه‌ها، معماری شبکه به راحتی قابل تشخیص خواهد بود و به همین دلیل مدیران شبکه از این روش‌ها استقبال نمی‌کنند.

۲. تعداد دفعات نشانه‌گذاری به مراتب کاهش پیدا می‌کند، به اینصورت که تعداد بسته‌های کم‌تری برای نشانه‌گذاری نیاز داریم و در عین حال فضای بیشتری در سرآیند بسته‌ها در اختیار خواهیم داشت.

۳. از آنجایی که روش ارائه شده در سطح جریان است، نرخ نشانه‌گذاری باز هم کاهش پیدا خواهد کرد، ولی نرخ ردگیری افزایش خواهد داشت.

۴. مسیرهای شناسایی شده در سطح AS بسیار پایدارتر نسبت به مسیرهای شناسایی شده در سطح مسیریاب هستند.

۵. تعداد AS ها به مراتب کم‌تر از تعداد مسیریاب‌ها می‌باشد. در فوریه ۲۰۱۸، تعداد AS ها ۶۰۱۷۸ عدد شمارش شده است [27] که در مقایسه با میلیون‌ها مسیریاب تعداد کمی می‌باشد. همان‌طور که در [26] بیان شده است، ۹۹ درصد از بسته‌ها برای رسیدن به مقصدشان کم‌تر از هشت AS و بیشتر از ۳۰ مسیریاب را پشت سر می‌گذارند. بنابراین ردگیری در سطح AS نسبت به مسیریاب، عملیاتی‌تر می‌باشد. به علاوه اینکه، مسیریابی بین AS ها با پروتکل مسیریابی BGP انجام می‌شود و تعداد AS های بین مبدا تا مقصد قابل شناسایی است. دانستن تعداد AS ها باعث بهبود نشانه‌گذاری و کاهش تعداد بسته‌های مورد نیاز، برای بازسازی مسیر می‌شود.

۲- پیشینه‌ی تحقیق

۲-۱- ردگیری IP مبتنی بر شبکه

ایده اصلی این روش، بررسی ارتباطات شبکه‌ای بین مسیریاب‌ها[۱۳]تا ردگیری مبدا حمله است. در این روش الگوی خاصی از حمله، توسط مدیر شبکه، مشخص و با آن به ردگیری حمله می‌پردازد. بعد از تشخیص حمله، با استفاده از اتصالات جریان بالا[۱۴]گره به گره به مبدا حمله نزدیک‌تر شده تا مبدا اصلی حمله شناسایی شود. روش‌های ردگیری با توجه به ساختار شبکه‌ای که در آن ردگیری صورت می‌پذیرد ارائه می‌شوند. این روش‌ها دارای ایرادات فراوانی هستند و باید به صورت دستی کنترل شوند. در ابن بخش می‌توان به Input Debugging [2] و Control Flooding [3] اشاره کرد.

۲-۲- ردگیری IP مبتنی بر نشانه‌گذاری

ایده اصلی این روش تزریق داده‌های ردگیری درون چند فیلد از سرآیند بسته‌های IP است و هدف آن بازسازی مسیر حمله و شناسایی گره حمله یا نزدیک‌ترین گره به آن است. به طور کلی به دو دسته‌ی نشانه‌گذاری احتمالی[۱۵]و نشانه‌گذاری قطعی[۱۶]تقسیم می‌شود. Savage و همکارانش [4] اولین بار روش نشانه‌گذاری احتمالی بسته[۱۷](PPM) را معرفی کردند. در این روش، هر مسیریاب در طول مسیر حرکت بسته، با احتمال P بخشی از اطلاعات مربوط به شبکه را درون سرآیند بسته قرار می‌دهد. این اطلاعات نشانه‌گذاری شده می‌تواند شامل آدرس IP مسیریاب فرستنده، آدرس IP مسیریاب بعدی و یا تعداد پرش‌های طی شده باشد. iTrace [6] با عبور هر ۲۰,۰۰۰ بسته، یک بسته‌ی جدید ICMP تولید کرده و به مقصد بسته‌های عبوری ارسال می‌کند. پیام‌های ICMP تولید شده، شامل داده‌های ردگیری، مهرزمانی[۱۸]و داده‌های احراز هویت بین مبدا و مقصد هستند. قربانی با توجه به داده‌های موجود در پیام‌های ICMP دریافت شده، مسیر حمله را بازسازی می‌کند.Blenky و همکارانش روش نشانه‌گذاری قطعی بسته[۱۹]DPM [7] را ارائه دادند. در این روش، تنها مسیریاب‌های مرزی ورودی شبکه، در مقایسه با تمام مسیریاب‌ها در روش PPM، بسته‌ها را نشانه‌گذاری می‌کنند. اطلاعات نشانه‌گذاری شامل بخشی از داده‌های شناسایی رابط[۲۰]ورودی است. شبکه‌ی جدیدی وارد شود، توسط مسیریاب مرزی ورودی آن شبکه، نشانه‌گذاری می‌شود. در نهایت گره قربانی قادر خواهد بود آدرس مسیریاب‌های مرزی شبکه‌هایی که بسته از آن عبور کرده است را شناسایی کند.

۲-۳- ردگیری IP مبتنی بر ثبت

در مقابل روش‌های نشانه‌گذاری بسته، روش‌های مبتنی بر ثبت، فضایی را در گره‌های میانی مسیر برای ثبت داده‌های نشانه‌گذاری به خود اختصاص می‌دهند. در این روش، نشانه‌ای از همه یا بخش زیادی از بسته‌های عبوری در هر گره ثبت و از این داده‌ها برای ردگیری استفاده می‌شود. اصلی‌ترین ضعف این روش‌ها، نیاز به فضای ذخیره‌سازی زیاد، پردازش سنگین آن و نیاز به سخت‌افزارهای خاص است. Snoeren و همکارانش [8] روش SPIE را ارائه دادند. ایده اصلی این روش ذخیره کردن نشانه‌ای از بسته‌های عبوری در ساختار داده‌ای به نام فیلتر بلوم[۲۱]در مسیریاب‌های طول مسیر است. در این روش نیاز به فضای ذخیره‌سازی زیاد در مسیریاب‌های عبوری داریم و همچنین این نوع ذخیره‌سازی مثبت کاذب بالایی[۲۲]دارد. از معایب این روش می‌توان به این مورد اشاره کرد که اگر تعداد AS های کمی از روش پیشنهادی پشتیبانی کنند، نرخ ردگیری در این روش به شدت کاهش پیدا خواهد کرد و داده‌های ردگیری زیادی تولید می‌کند که نیاز به فضای ذخیره‌سازی زیادی دارد.

۲-۴- ردگیری IP ترکیبی

ایده روش‌های ترکیبی، به کارگیری دو یا چند روش ردگیری ذکر شده است، از بارزترین نمونه‌های روش‌های ترکیبی، می‌توان به ترکیب روش‌های نشانه‌گذاری و ثبت بسته اشاره کرد که از مزایای هر دو روش استفاده شده است. اصلی‌ترین روش‌های ترکیبی ارائه شده عبارتند از: ردگیری لیست اتصالات توزیع شده[۳۳](DLLT) [10] و نشانه‌گذاری حلقه‌ای احتمالی بسته‌ها[۳۴](PPPM) [11]. در روش اول؛ اطلاعات نشانه‌گذاری شده، در مسیریاب‌های هسته[۳۵]نگه‌داری می‌شوند. روش دوم به بررسی بسته‌هایی با مقاصد یکسان، به منظور دستیابی به آدرس IP مسیریاب‌های درگیر در روند نشانه‌گذاری بسته‌ها می‌پردازد، این سازوکار نیاز به فضای زیادی برای ذخیره‌سازی در مسیریاب‌های هسته ندارد. مهاجم در این روش می‌تواند رشته نشانه‌گذاری در سرآیند بسته‌ها را جعل کند اما با بررسی مقدار موجود در رشته TTL[۳۶] می‌توان به این مهم پی‌برد. در مقایسه این روش با روش ثبت بسته‌ها، سربار پردازشی و محاسباتی و فضای ذخیره‌سازی کاهش پیدا می‌کند. روش RIHT [12] شماره‌ی رابط مسیریاب را نشانه‌گذاری می‌کند و در صورتی که فضای اختصاص داده شده به آن در سرآیند بسته فضای کافی را نداشته باشد، آن را در جدول مسیریاب، ثبت می‌کند.

۲-۵- ردگیری IP در سطح AS

در این بخش به روش‌های ردگیری که در سطح AS ارائه شده‌اند، می‌پردازیم. در [13] نویسنده روش ردگیری را برای شناسایی AS مهاجم ارائه کرده است. در این روش با احتمال نشانه‌گذاری P=1/6 شماره سیستم خودمختار[۳۷](ASN) نشانه‌گذاری می‌شود. این روش در زمانی که شماره AS ۱۶ بیت بوده است ارائه شده است. Gong و همکارانش در [14] روش AS-SPT که قابلیت اجرا در حالتی که بخشی از شبکه تحت پوشش از روش پیشنهادی پشتیبانی نمی‌کنند را دارد، ارائه داد و از روش ردگیری ترکیبی برای ردگیری AS مهاجم استفاده می‌کند. AS قربانی با دریافت ترافیک حمله از همسایه‌هایش درخواست بررسی داده‌های ثبت شده را دارد، در صورت بازخورد جواب مثبت از یکی از همسایه‌ها این درخواست تکرار شده تا AS مهاجم شناسایی شود. از مهم‌ترین معایب این روش می‌توان به این مورد اشاره کرد که به دلیل ارائه شدن این روش در سطح بسته، مقدار زیادی داده‌ی ردگیری تولید کرده که نیاز به فضای زیاد ثبت در گره‌های میانی برای دارد. Durresi و همکارانش روش FAST [15] را ارائه کردند که تنها قابلیت ردگیری پنج AS را دارد و در صورت وجود بیش از این تعداد AS در طول مسیر، حمله قابل شناسایی نمی‌باشد.

۳- مفاهیم پایه

روش پیشنهادی در سطح جریان ارائه شده است، به این صورت که به ازای تمام بسته‌های جریان که در پنج مشخصه‌ی آدرس مبدا و مقصد، شماره درگاه مبدا و مقصد و نوع پروتکل لایه ۴، یکسان هستند، تعداد محدودی از آن بسته‌ها به نمایندگی از تمام بسته‌های جریان نشانه‌گذاری می‌شود.

در ادامه این بخش، در مورد سیستم‌های خودمختار AS، نقش پروتکل BGP در روش پیشنهادی و سرآیند بسته‌های IP صحبت خواهیم کرد.

۳-۱- سیستم خودمختار(AS)

سیستم‌های خودمختار مجموعه‌ای از زیرشبکه‌های متصل به هم هستند که تحت کنترل یک یا چندین عملگر شبکه با یک مدیریت واحد قرار دارند. یک AS سیاست‌های مسیریابی کاملاً تعریف شده‌ای نسبت به اینترنت دارد، به عبارت دیگر AS می‌تواند داده‌هایی که مبدا و مقصد آن شبکه خودش باشد را به صورت داخلی جابجا کند و ارتباطات با دیگر شبکه‌ها را بر اساس پروتکل BGP تنظیم کند. در سال ۲۰۱۰، ASN توسط [16] از ۱۶ بیتی به ۳۲ بیتی تغییر پیدا کرد.

۳-۲- نقش پروتکل BGP در روش پیشنهادی

پروتکل شاهراه مرزی[۳۸](BGP)، پروتکل مسیریابی استانداردی است که برای انتقال اطلاعات درون AS و بین AS با استفاده از تبادل پیام‌های به‌روز رسانی بین آن‌ها استفاده می‌شود. پیام به‌روز رسانی می‌تواند شامل چندین ویژگی مسیر[۳۹]باشد. ویژگی مسیر extended community [17] که اختیاری[۴۰]و تعدی[۴۱]می‌باشد، برای انتقال اطلاعات اضافی بین دو همسایه راه دور می‌تواند مورد استفاده قرار گیرد و در نتیجه می‌تواند یک گروه منطقی بین AS های دارای وجه اشتراک ایجاد کند.

ما از ویژگی مسیر تعدی extended community، به منظور شناسایی AS هایی که از روش پیشنهادی در شبکه گسترش جزئی[۴۲] پشتیبانی می‌کنند، استفاده می‌کنیم. این ویژگی به عنوان بخشی از پیام به‌روز رسانی خروجی است که، توسط AS های پشتیبانی کننده از روش پیشنهادی ارسال می‌شود. AS هایی که از روش پیشنهادی پشتیبانی نمی‌کنند، هیچ مشکلی در انتقال پیام‌های به‌روز رسانی به وجود نخواهند آورد زیرا extended community ویژگی‌ای تعدی در پروتکل BGP است و بدون هیچ تغییری آن را ارسال خواهند کرد.

۳-۳- فضای ذخیره‌سازی در سرآیند بسته‌های IP

همانطور که اشاره شد در روش‌های نشانه‌گذاری داده‌های ردگیری در چند فیلد سرآیند بسته‌های IP تزریق می‌شوند. در برخی روش‌ها از فیلد اختیارات[۴۳] بسته IP برای این منظور استفاده شده است که علاوه بر سربار زیاد باعث دور انداخته شدن[۴۴]بسته می‌شود. محققان برای حل این موضوع در [18]، [4,5,6,7] از ۱۶ بیت، فیلد شناسایی[۴۵]و ۱۳ بیت، fragment offset استفاده کرده‌اند. ما نیز در روش پیشنهادی از این دو فیلد برای نگه‌داری داده‌های ردگیری در سرآیند بسته، استفاده کرده‌ایم

Savage و همکارانش در [4] بررسی کرده‌اند که تنها۰٫۲۵٪ از بسته‌ها تکه‌تکه شده‌اند و بیان می‌کند که با صرف نظر کردن از این بسته‌ها تنها ۰٫۰۶٪ از ترافیک‌های قانونی تحت تاثیر قرار می‌گیرد که مقدار زیادی نیست.

۴- روش پیشنهادی

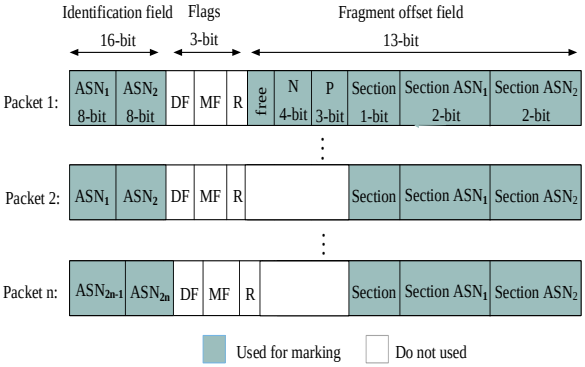
۴-۱- مرحله ۱: ساز و کار نشانه‌گذاری بسته‌های جریان

در روش پیشنهادی، به جای نشانه‌گذاری بسته‌ها با آدرس IP و یا شماره رابط مسیریاب، از ASN درون بسته‌های جریان نشانه‌گذاری، استفاده می‌کنیم. روش پیشنهادی در سطح جریان نشانه‌گذاری را انجام می‌دهد، به

طوری که چهار بسته اول هر جریان برای نشانه‌گذاری مورد استفاده قرار می‌گیرد و با توجه به تعداد بسته‌های موجود در جریان می‌توان تا شانزده AS را نیز ردگیری کرد. تعداد AS هایی که جریان از آن عبور می‌کند رابطه‌ی مستقیمی با تعداد بسته‌های نشانه‌گذاری شده در آن جریان دارد. برای مثال اگر جریان از دو AS عبور نماید، تنها بسته‌ی اول آن نشانه‌گذاری شده و اگر از شانزده AS عبور نماید، هشت بسته‌ی ابتدایی آن توسط ASBR ها نشانه‌گذاری خواهد شد. همانطور که بیان شد بر خلاف روش‌های دیگر [19] که تعداد ثابتی از بسته‌های جریان نشانه‌گذاری می‌شود، در روش پیشنهادی این مقدار متغییر می‌باشد، ASN که مقداری ۳۲ بیتی است، به چهار قسمت ۸ بیتی شکسته شده و در چهار مرحله، درون فیلد شناسایی بسته‌های جریان‌های عبوری، نشانه‌گذاری می‌شود. برای انتقال کامل شماره AS نیاز به چهار مرحله نشانه‌گذاری در چهار جریان مجزا است. در مقصد با وجود داده‌های کنترلی که در فیلد fragment offset به صورت همزمان نشانه‌گذاری می‌شود، می‌توان شماره AS های پیموده شده را شناسایی کرد. در ادامه الگوریتم نشانه‌گذاری بیان شده است.

- **۸ بیت تکه‌تکه شده:** برای نشانه‌گذاری شماره AS_m در فیلد شناسایی
- **۸ بیت تکه‌تکه شده:** برای نشانه‌گذاری شماره AS_{m+1} در فیلد شناسایی
- **۴ بیت رشته فاصله (N):** برای نشان دادن تعداد AS های عبور کرده (با ۴ بیت می‌توان تا شانزده AS عبوری را شمارش کرد)
- **۳ بیت تعیین بسته (P):** برای تعیین بسته‌ای که در هر مرحله باید نشانه‌گذاری شود. (با ۳ بیت می‌توان ۸ بسته را شمارش کرد)
- **۱ بیت تعیین بخش (section):** برای تعیین بخش (اول یا دوم) رشته شناسایی جهت نشانه‌گذاری.
- **۲ بیت تعیین تکه شماره AS ($section\ ASN_m$):** برای تعیین بخشی از شماره AS که توسط $ASBR_m$ نشانه‌گذاری شده است.(هر ۴ تکه‌ی شماره AS پس از دریافت تعیین مکان شده و شماره AS آشکار می‌شود)
- **۲ بیت تعیین تکه شماره AS ($section\ ASN_{m+1}$):** برای تعیین بخشی از شماره AS که توسط $ASBR_{m+1}$ نشانه‌گذاری شده است.

شکل ۱، سازوکار نشانه‌گذاری بسته‌های جریان را نشان می‌دهد.



شکل (۱): سازوکار نشانه‌گذاری بسته‌های جریان

۴-۲- مرحله ۲ : تولید امضای مسیر با AS مبدا

در AS قربانی[۳۶]، جریان‌های دریافت شده از نظر جعلی بودن یا نبودن مورد بررسی قرار می‌گیرد. برای این منظور دو دسته اطلاعات مورد نیاز است:

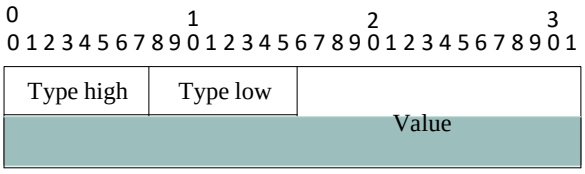
- شناسایی AS مبدا آدرس IP بسته‌ها (شناسایی محدوده آدرس IP متعلق به هر کدام از AS ها)
- شناسایی امضای مسیر آن AS

درون هر AS که از روش پیشنهادی پشتیبانی می‌کند، یک سرویس‌دهنده اختصاصی[۳۷]برای تبادل پیام‌ها با AS های دیگر و حفظ ارتباط با آن‌ها تعبیه شده است. AS ها، آدرس IP سرویس‌دهنده‌شان را، از طریق ویژگی مسیر extended community پیام‌های به‌روز رسانی پروتکل BGP به AS های دیگر منتقل می‌کنند. سپس ارتباط منطقی بین سرویس دهنده‌ها ایجاد شده که هر کدام از این سرویس دهنده‌ها اطلاعات مربوط به AS خود را به تمام سرویس‌دهنده‌های دیگر درون AS های آن‌ها منتقل می‌کند.

اولین وظیفه‌ی سرویس‌دهنده‌ها، نگه‌داری نگاشتی از آدرس‌های IP است که در اینترنت با AS شان مکاتبه دارند. این وظیفه می‌تواند بصورت برون‌خطی انجام شود. همچنین در [20]، تعدادی از ابزارها برای نگاشت آدرس IP ها ارائه شده است. مثلاً محدوده آدرس‌های موجود در هر AS با بررسی آگهی‌های BGP در مسیریاب‌ها بدست خواهد آمد.

یادگیری امضای مسیرِ متعلق به هر AS اصلی‌ترین وظیفه‌ی سرویس‌دهنده اختصاصی است. (AS های همسایه را شناسایی و مسیر بین آن‌ها را یادگیری می‌کند) در هر سرویس‌دهنده پایگاه‌داده‌ای از امضاهای مسیر بین AS مبدا با AS های دیگر ایجاد می‌شود.

ویژگی مسیر extended community پروتکل BGP، همانطور که در شکل ۲ نشان داده شده است، هشت بایت[۳۸]طول دارد و شامل فیلدهای نوع[۳۹]و مقدار[۴۰]است. فیلد مقدار، شش بایت طول دارد که از چهار بایت آن برای انتقال آدرس IP سرویس‌دهنده‌ی اختصاصی تعبیه شده در AS، استفاده می‌کنیم.



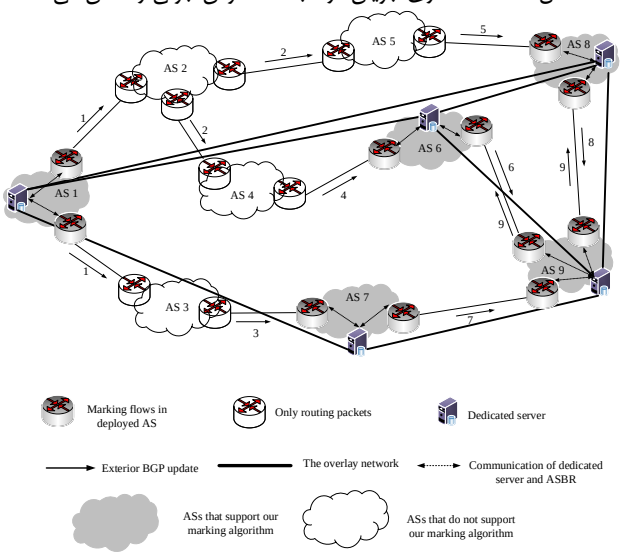
شکل ۲ : ویژگی مسیر extended community پروتکل BGP

ما از پروتکل BGP برای گسترش روش پیشنهادی استفاده می‌کنیم. شکل ۳، نحوه ساخت شبکه تحت پوشش را نشان می‌دهد. AS هایی که از روش پیشنهادی پشتیبانی می‌کنند، درون‌شان یک سرویس‌دهنده تعبیه شده است که با تمام مسیریاب‌های مرزی آن AS ارتباط داشته و اطلاعات مورد نیاز را دریافت و ارسال می‌کند. در ابتدا جدول شبکه تحت پوشش[۴۱] خالی است. AS1 پیام به‌روز رسانی را تولید و به همسایه‌هایش ارسال می‌کند (مرحله ۱)، از آنجایی که AS2 و AS3 از روش ما پشتیبانی نمی‌کنند، پیام رسیده را دریافت و بر طبق ویژگی تعدی، پیام به‌روز رسانی جدیدی تولید و آن را به همسایه‌ها منتقل می‌کنند (مرحله ۲ و ۳). AS7 با دریافت پیام به‌روز

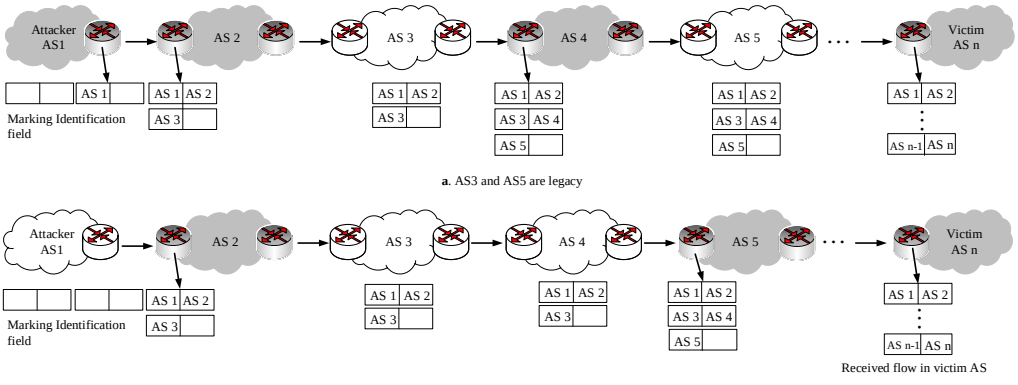
رسانی AS1 را به عنوان همسایه راه‌دور[۲]در جدول شبکه تحت پوشش ثبت می‌کند. این روند تا زمانی که شبکه تحت پوشش، همانطور که در شکل ۳ نمایش داده شده تکمیل شود، ادامه خواهد داشت.

۴-۳- مرحله ۳ : نشانه‌گذاری جریان

هرگاه ASBR جریانی را دریافت کند، ابتدا داده‌های کنترلی موجود در فیلد fragment offset بسته‌ها را بررسی کرده. فیلد P تعیین می‌کند که کدام بسته از این جریان باید نشانه‌گذاری شود. فیلد بخش[۴۳]تعیین کننده‌ی بخش اول یا دوم فیلد شناسایی برای نشانه‌گذاری داده ردگیری است. هرگاه جریان از AS ای عبور کرده و نشانه‌گذاری شود یک واحد به مقدار N اضافه می‌شود. در نهایت فیلد section ASN، نشان دهنده این است که کدام بخش از شماره AS نشانه‌گذاری شده است. همانطور که بیان شد با استفاده از پروتکل BGP می‌توان روش پیشنهادی را گسترش و از وضعیت همسایه‌ها[۴۴] با خبر شد. در روش پیشنهادی اگر AS بالادست[۴۵]یا پایین‌دست[۴۶]از روش پیشنهادی پشتیبانی نکند، AS قبل یا بعد از آن باید نشانه‌گذاری را جبران کند. شکل ۴، نشانه‌گذاری جریان در شبکه گسترش جزئی را نشان می‌دهد.



شکل (۳) : ساخت شبکه تحت پوشش



a. AS3 and AS5 are legacy

b. AS1, AS3 and AS4 are legacy

شکل (۴) : نشانه‌گذاری جریان در شبکه گسترش جزئی

۴-۴- مرحله ۴ : تولید امضای مسیر

اگر AS قربانی جریانی را دریافت و در پایگاه‌داده‌ی سرویس‌دهنده‌اش، امضای مسیر آن جریان، برای AS ادعا شده را نداشته باشد، در اینصورت یک پیام درخواست[۴۷]به AS مبدا ارسال می‌کند. این درخواست با کلید عمومی سرویس‌دهنده مقابل امضا شده تا هویت درخواست دهنده را تایید کند. سپس AS مبدا، درخواست دریافت شده را با امضای آن با کلید عمومی درخواست دهنده پاسخ می‌دهد. در این پیام‌ها از مهر زمانی نیز استفاده شده تا از حمله تکرار[۴۸]جلوگیری شود.

رمزنگاری مورد بحث در فوق با استفاده از معماری مسیریابی امن BGP قابل اجر خواهد بود که پیش‌تر مورد بحث قرار گرفت. جریان‌هایی که در پاسخ ارسال می‌شوند، دقیقا همان مسیری را طی می‌کنند که دیگر جریان‌ها عبور می‌کنند.

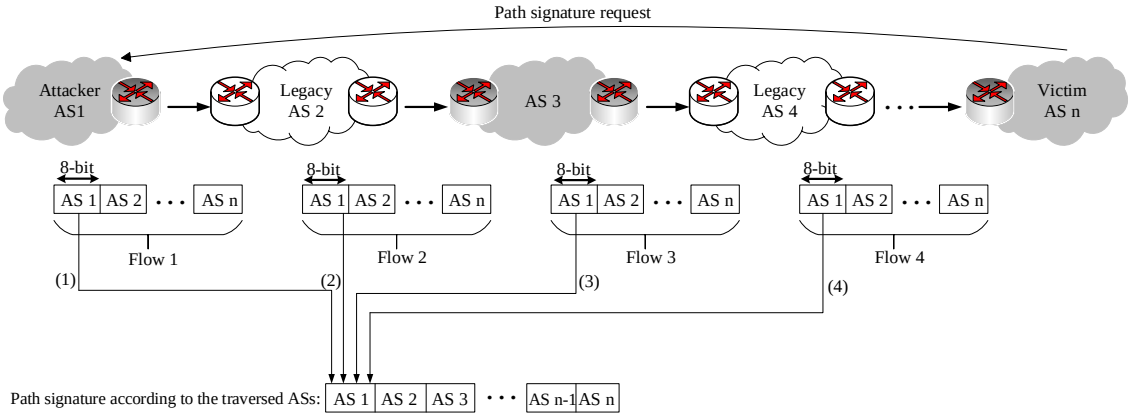
سرویس‌دهنده مقصد، پیام درخواست را به تمام AS های طول مسیر، که از روش ما پشتیبانی می‌کنند، ارسال می‌کند. هر کدام از سرویس‌دهنده‌ها، مسیر خود تا سرویس‌دهنده مقصد را، با تولید جریان‌های جدید و قرار دادن بخشی از شماره AS در آن‌ها، می‌سازد و ارسال می‌کند.

برای تولید امضای مسیر به صورت کامل نیاز به تولید و ارسال چهار جریان است. در هر جریان یک چهارم شماره AS های پیموده شده قرار داده می‌شود و در نهایت با کمک رشته‌های کنترلی می‌توان شماره AS های پیموده را به صورت کامل ساخت. شکل ۵، نحوه تولید امضای مسیر در شبکه گسترش جزئی را نمایش می‌دهد. همانطور که در این شکل مشاهده می‌کنید، تنها AS1 و AS3 از روش پیشنهادی پشتیبانی می‌کنند، بنابراین این AS ها پیام درخواست امضای مسیر را دریافت کرده و در جواب آن، چهار جریان جدید تولید و در مسیر حرکت تا درخواست کننده، ارسال می‌کنند. در مقصد، با ترکیب جریان‌های دریافت شده با استفاده از داده‌های کنترلی، می‌توان امضای مسیر را ساخت.

۴-۵- مرحله ۵ : ردگیری IP (شناسایی AS مبدا)

AS مبدا، با تطابق جریان‌های دریافت شده در مقصد و امضاهای مسیر موجود در پایگاه داده، قابل شناسایی است. در صورت دریافت ترافیک حمله، ابتدا یکی از جریان‌های آن را بررسی کرده و داده‌های نشانه‌گذاری شده در آن با امضاهای موجود در پایگاه‌داده‌ی سرویس‌دهنده مقایسه می‌شود. اگر تطابقی با هر کدام از امضاهای موجود در پایگاه‌داده پیدا شود، با داشتن ۳ امضای دیگر متعلق به آن مسیر، می‌توان شماره AS های عبوری را تشخیص داد.

نکته‌ی قابل توجه این که، تنها با داشتن ۱ جریان حمله و تطبیق آن با امضاهای موجود در پایگاه‌داده، در صورت وجود مشابهت می‌توان مبدا حمله را شناسایی و مسیر حمله را بازسازی کرد.



شکل (۵) : تولید امضای مسیر در شبکه گسترش جزئی

۵- ارزیابی روش پیشنهادی

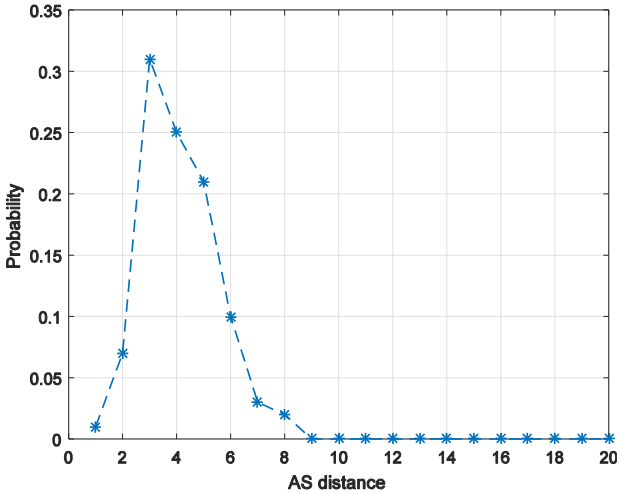
برای ارزیابی روش ارائه شده از مجموعه‌داده‌ی تولید شده در مرکز امنیت اطلاعات ISCX به نام UNB ISCXIDS [21] استفاده می‌کنیم. این ترافیک‌ها که شامل ترافیک‌های حمله و عاری از حمله می‌باشد، در هفت روز هفته ثبت شده است. از دو مجموعه داده حمله منع سرویس توزیع شده و ترافیک‌های معمولی، که به ترتیب ۲۳,۲ و ۴,۲۲ گیگابایت حجم دارند و نتیجه‌ی ترافیک‌های ثبت شده در دو روز هفته می‌باشد، استفاده می‌کنیم. حمله منع سرویس توزیع شده با استفاده از باتنت IRC اجرا می‌شود؛ این بات که ضمیمه پیام‌های به‌روز رسانی می‌شود، به سیستم هدف نفوذ کرده و سپس با بارگیری پایین[۴۹]مخفیانه‌ی ترافیک‌های مخرب، به هدف خصمانه خود می‌رسد.

برای تحلیل و استخراج جریان‌های موجود در مجموعه‌داده ذکر شده، از ابزار argues [22] استفاده می‌کنیم. این ابزار برای استخراج جریان‌ها به همراه ویژگی‌های ترافیکی آن مورد استفاده قرار می‌گیرد، با استفاده از آن تمام بسته‌های موجود در مجموعه‌داده که در پنج ویژگی؛ آدرس مبدا و مقصد، شماره درگاه مبدا و مقصد و نوع پروتکل لایه ۴، یکسان هستند را، به عنوان یک جریان استخراج می‌کنیم. حال با استفاده از قطعه کد پایتونی که برای تبدیل خروجی دریافت شده از ابزار Argus به فایل CSV نوشته‌ایم، فایل جریان دریافت شده از ابزار Argus را به فایل CSV تبدیل کرده و با بررسی داده‌های موجود در این فایل، اطلاعات آماری مربوط به این مجموعه‌داده‌ها در جدول ۱، بیان می‌شود.

جدول (۱) :اطلاعات آماری مجموعه داده مورد ارزیابی

مجموعه داده	تعداد بسته‌ها	تعداد جریان‌ها
UNB ISCX free attack	۶۳۸۰۶۰۸	۲۶۳۳۹۷
UNB ISCX DDoS attack	۲۴۴۷۸۹۸۷	۱۰۰۶۴۶۹

همانطور که بیان شد، نشانه‌گذاری تمام بسته‌های موجود در ترافیک باعث افزایش نرخ نشانه‌گذاری[۵۰]و همچنین افزایش سربار پردازشی در گره‌های میانی و گره مقصد می‌شود. در اکثر روش‌های ردگیری [4-7]، [18,19] حجم زیادی از داده‌های ردگیری (داده‌های نشانه‌گذاری شده) در مقصد دریافت می‌شود که پردازش این داده‌ها در برخی موارد تا چندین روز زمان صرف می‌کند، و عملاً این روش‌ها دیگر در حین اجرای حمله قادر به شناسایی مبدا حمله و مقابله با آن نیستند. همچنین در روش‌های ذکر شده، هدف از جمع‌آوری داده‌های ردگیری در مقصد، بازسازی مسیر حمله یا گراف شبکه است، اما در روش پیشنهادی تنها با تطبیق یک جریان حمله با امضاهای مسیر موجود در پایگاه‌داده می‌توان مسیر حمله را در کم‌ترین زمان و با حداقل سربار ممکن شناسایی کرد. در نتیجه برخلاف روش‌های ارائه شده تاکنون، روش پیشنهادی در زمان اجرای حمله قادر به شناسایی مبدا و مسیر حمله خواهد بود که در نهایت می‌تواند با حمله مقابله کند.

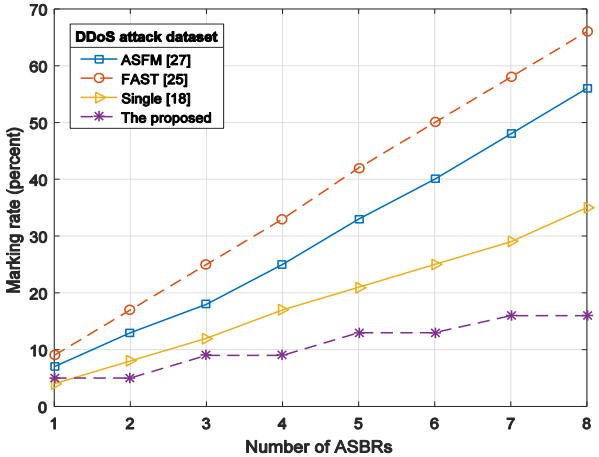


شکل (۶) : فراوانی تعداد AS ها بین مبدا و مقصد [26]

۵-۱- نرخ نشانه‌گذاری

نرخ نشانه‌گذاری برابر است با بسته‌های نشانه‌گذاری شده توسط ASBR ها به تمام بسته‌ها. همانطور که در شکل ۶ نشان داده شده است، ۹۹٪ بسته‌ها برای رسیدن به مقاصدشان از کم‌تر از هشت AS عبور کرده. همچنین AS سوم، بیشترین فراوانی عبور بسته‌ها را دارد که در روش پیشنهادی ترخ

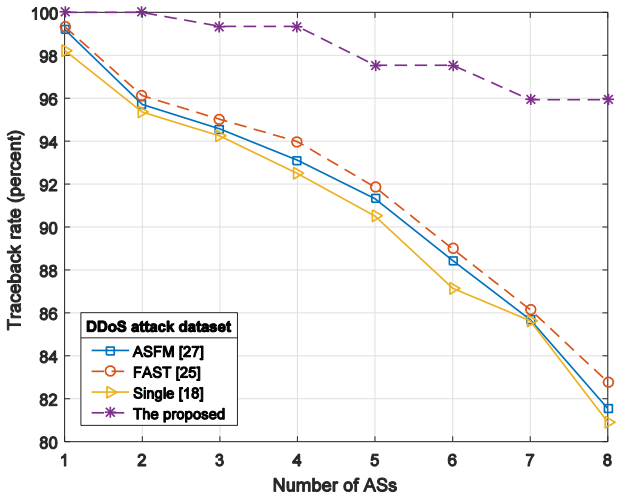
نشـانه‌گـذاری در ایـن AS بـه ترتیـب در مجموعـه‌داده عـاری از حملـه و مجموعه‌داده حمله برابر ۸٪ و ۹٪ است. با مقایسه روش پیشـنهادی بـا دیگـر روش‌ها در شکل ۷ متوجه این موضوع می‌شویم کـه روش پیشـنهادی دارای کم‌ترین نرخ نشانه‌گذاری است.



شکل (۷) : نرخ نشانه‌گذاری در روش پیشنهادی و سایر روش‌ها

۵-۲- نرخ ردگیری

نرخ ردگیری برابر است با تعداد بسته‌های ردگیری شده به تمام بسته‌ها. بـرای ردگیری AS مهاجم در روش پیشنهادی تنها به یک جریان نشانه‌گذاری شده برای تطبیق با پایگاه‌داده امضاهای مسیر نیاز داریم؛ بر خلاف روش‌های دیگر [4-9]، [18] و [23-25] که مسیر حملـه را بـا داده‌هـای دریافـت شـده در مقصد بازسازی کرده. تنها دلیلی که نرخ ردگیـری در روش پیشـنهادی ۱۰۰٪ نیست این است که امکان دارد جریان‌هایی در طول مسیر وجود داشته باشـند که تعداد بسته‌های موجود در آن کم‌تر از چهار بسته باشد، زیرا برای ردگیـری هشت AS نیاز چهار بسته درون جریان داریم. شـکل ۸ نـرخ ردگیـری روش پیشنهادی در مقایسه با دیگر روش‌ها را نشان می‌دهد.



شکل (۸) : نرخ ردگیری در روش پیشنهادی و سایر روش‌ها

۶- نتیجه‌گیری

در این مقاله روش ردگیری IP مبتنی بر نشانه‌گذاری جریـان در سـطح AS ارائـه شـده اسـت. بـرای تشـکیل شـبکه تحـت پوشـش از ویژگـی مسـیر extended community پروتکل BGP، استفاده کرده‌ایم. این ویژگی نیاز به گسترش شبکه تحت پوشش را بر روی تمام مسیریاب‌ها، حـذف کـرده است. سازوکار نشانه‌گـذاری در سـطح جریـان و رشـته‌هـای کنترلـی، بـرای بازسازی داده‌های دریافت شده در مقصد، ارائه شده است. همچنین با اسـتفاده از سازوکار تولید امضای مسیر، قادر به شناسایی ترافیک‌های جعـل شـده نیـز خواهیم بود و در کم‌ترین زمان ممکن با حداقل سربار در حین اجـرای حملـه می‌توان AS مهاجم را شناسایی و با حمله مقابله کرد.

از جمله کارهایی که در ادامه‌ی این تحقیق مـی‌تـوان انجـام داد، ارائـه روش ردگیری در پروتکل اینترنت نسخه ۶ [۵۱]است، زیرا با توجه بـه گسـترش این پروتکل در سطح اینترنت، نیازمند روش‌هایی برای ردگیری در آن هستیم.

مراجع

[1] Arbor Networks, 12th *Worldwide Infrastructure Security Report*, 21-22, 2016.

[2] R. Stone: Centertrack: *An IP Overlay Network for Tracking DoS Floods*, Proceedings of the 9th conference on USENIX Security Symposium, Berkeley, USA, 2000, pp. 199-212

[3] H. Burch, B. Cheswick: *Tracing Anonymous Packets to Their Approximate Source*, Proceedings of the 14th USENIX Conference on System Administration, New Orleans, LA, USA, 2000, pp. 319-328

[4] S. Savage, D. Wetherall, A. Karlin, and T. Anderson. *Practical Network Support for IP Traceback*. In SIGCOMM '00: Proceedings of the conference on Applications,Technologies, Architectures, and Protocols for Computer Communication, pages 295-306, New York, NY, USA, 2000. ACM.

[5] A. Yaar, A. Perrig, and D. Song. *FIT: Fast Internet Traceback*. In INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE, pages 1395-1406, 2005.

[6] S.M. Bellovin, M. Leech, and T. Taylor. *ICMP Traceback Messages*,March 2000.

[7]] A. Belenky and N. Ansari. *Tracing Multiple Attackers with Deterministic Packet Marking (DPM)*. Communications, Computers and signal Processing, 2003. PACRIM. 2003 IEEE Pacific Rim Conference on, 1:49-52, Aug. 2003.

[8] AC Snoeren, C. Partridge, LA Sanchez, CE Jones, F. Tchakountio, B. Schwartz, ST Kent, and WT Strayer. *Single-Packet IP Traceback*. Networking, IEEE/ACM Transactions on, 10(6):721-734, 2002

[9] C. Gong, T. Le, T. Korkmaz, and K. Sarac. *Single Packet IP Traceback in ASlevel Partial Deployment Scenario*. In Global Telecommunications Conference,2005. GLOBECOM'05. IEEE, pages 1817-1821, Nov. 2005.

[10] Duwairi, B., Chakrabarti, A., & Manimaran, G. (2004, May). *An efficient probabilistic packet marking scheme for IP traceback. In International* Conference on Research in Networking (pp. 1263-1269). Springer, Berlin, Heidelberg.

[11] B.Al-Duwari and M. Govindarasu: *Novel Hybrid Schemes Employing Packet Marking and Logging for IP Traceback*, IEEE Trans. Parallel Distributed Syst., Vol. 17, No. 5, May 2006, pp. 403-418.

[12] M. H. yang, M. C. Yang: RIHT*: A Novel Hybrid IP Traceback Scheme*, IEEE Trans. Information Forensics and Security, Vol. 7, No. 2, April 2012, pp. 789-797.

[13] Paruchuri, V., Durresi, A., Kannan, R., & Iyengar, S. S. (2004). *Authenticated autonomous system traceback. In Advanced Information Networking and Applications*,

2004. AINA 2004. 18th International Conference on (Vol. 1, pp. 406-413). IEEE.
[14] C. Gong and K. Sarac, "*Toward a practical packet marking approach for IP traceback*," International Journal of Network Security, vol. 8, no. 3, pp. 71–84, 2009.
[15] V. Paruchuri, A. Durresi, and L. Barolli. *FAST: Fast Autonomous System Traceback.* 21st International Conference on Advanced Information Networking and Applications (AINA2007), pages 498–505, May 2007.
[16] *Regional Internet Registeries*. https://www.arin.net , November 15, 2017.
[17] Sangli, S., Tappan, D., & Rekhter, Y. (2006). *BGP extended communities attribute* (No. RFC 4360).
[18] H. Aljifri, "*IP traceback: a new denial-of-service deterrent*?" IEEE Security and Privacy, vol. 1, no. 3, pp. 24–31, 2003.
[19] Aghaei-Foroushani, V., & Zincir-Heywood, A. N. (2016, April). *Autonomous system based flow marking scheme for IP-Traceback*. In Network Operations and Management Symposium (NOMS), 2016 IEEE/IFIP (pp. 121-128). IEEE.
[20] Team Cymru - IP to ASN Mapping. [Online]. Available: http://www.team-cymru.org , August 29, 2016.
[21] *University of new Brunswick.NSL.KDD dataset*. http://www.unb.ca , August 10, 2017
[22] *Argus Network Monitoring and Analysing*. http://www.qosient.com , August 24, 2017
[23] Aghaei-Foroushani, V., & Zincir-Heywood, A. N. (2016, April). *Autonomous system based flow marking scheme for IP-Traceback.* In Network Operations and Management Symposium (NOMS), 2016 IEEE/IFIP (pp. 121-128). IEEE.
[24] V. Paruchuri, A. Durresi, and L. Barolli. *FAST: Fast Autonomous System Traceback.* 21st International Conference on Advanced Information Networking and Applications (AINA2007), pages 498–505, May 2007.
[25] C. Gong, T. Le, T. Korkmaz, and K. Sarac, "*Single Packet IP Traceback in AS-level Partial Deployment Scenario*," in Global Telecommunications Conference, 2005. GLOBECOM'05. IEEE, Nov. 2005,pp.1817–1821.
[26] *The CAIDA Skitter AS Links Dataset*. http://www.caida.org , November 13, 2017.
[27] *CIDR Report.* http://www.cidr-report.org , February 20, 2018

زیرنویس‌ها

۱ Distributed Denial of Service
۲ Internet Protocol
۳ Source Address IP
٤ Spoof
٥ IP Traceback
٦ Field
۷ Topology
۸ Path Signature
۹ Packet Marking
۱۰ Flow Marking
۱۱ Autonomous System
۱۲ Autonomous System Border Router
۱۳ Router
۱٤ Upstream Interface
۱٥ Probabilistic Marking
۱٦ Deterministic Marking
۱۷ Probabilistic Packet Marking
۱۸ Timestamp
۱۹ Deterministic Packet Marking
۲۰ Interface
۲۱ Bloom Filter
۲۲ False Positive
۲۳ Distributed Linked List Traceback
۲٤ Probabilistic Pipeline Packet Marking
۲٥ Core Router
۲٦ Time to Live
۲۷ Autonomous System Number
۲۸ Border Gateway Protocol
۲۹ Path Attribute
۳۰ Optional
۳۱ Transitive
۳۲ Partial Deployment
۳۳ Option
۳٤ Drop
۳٥ Identification Field
۳٦ Victim AS
۳۷ Dedicate Server
۳۸ Byte
۳۹ Type
٤۰ Value
٤۱ The Overlay Table
٤۲ Remote Neighbor
٤۳ Section
٤٤ Neighbors
٤٥ Upstream
٤٦ Downstream
٤۷ Path Signature Request
٤۸ Reply Attack
٤۹ Download
٥۰ Marking Rate
٥۱ IPv6