مجله علمی- پژوهشی "پدافند الکترونیکی و سایبری"
سال پنجم، شماره ۲، تابستان ۱۳۹۶، ص ۳۵- ۱۹

ارائه روش حل مسأله نقاط مرزی در نشان‌گذاری مبتنی بر فاصله در جریان شبکه گمنامی

احمد احمدی[۱]، مهدی دهقانی[۲*]، محمود صالح اصفهانی[۳]

۱- کارشناس ارشد مهندسی نرم‌فزار، ۲- دکتری کامپیوتر، ۳- استادیار، دانشگاه جامع امام حسین (ع)

(دریافت: ۹۴/۰۹/۰۹، پذیرش: ۹۵/۱۰/۱۴)

چکیده

توانمندی ردیابی نفوذ، نقش بازدارنده‌ای در ناامنی دارد. یکی از روش‌های ردیابی نفوذ، روش نشان‌گذاری ترافیک شبکه است. در این روش با تغییر در الگوی جریان شبکه گمنامی، ترافیک جریانی خاص نشان‌گذاری شده و در مرزهای خروجی شبکه آن جریان ردیابی می‌گردد. در این تحقیق، روش نشان‌گذاری مبتنی بر فاصله که تابه‌حال روی شبکه گمنامی مسیریابی پیازی ارزیابی نشده است، پیاده‌سازی شده و به‌صورت عملی در محیط واقعی بر روی شبکه مسیریابی پیازی مورد ارزیابی قرار گرفت. تحلیل نتایج نشان می‌دهد که این روش دارای نقطه ضعف نقاط مرزی است. راه‌کار پیشنهادی بهبود روش مبتنی بر فاصله، با ایجاد فضای خالی محافظ در انتهای هر فاصله برای رفع مسأله نقاط مرزی مورد استفاده قرار گرفت. بعد از پیاده‌سازی و ارزیابی عملی روش پیشنهادی بهبودیافته مبتنی بر فاصله، اندازه‌گیری دقت نرخ‌های کدگشایی، مثبت اشتباهی و منفی اشتباهی، نشان می‌دهد که روش پیشنهادی کارایی بهتری در مقایسه با روش اولیه نشان‌گذاری مبتنی بر فاصله دارد. هم‌چنین برای ارزیابی نامحسوسی با فرض سناریوی نفوذ، از روش‌های آماری K-S، آنتروپی و آنتروپی شرطی جهت تشخیص وجود نشان‌گذاری استفاده شد. نتایج نشان می‌دهد که روش پیشنهادی در آزمون‌های K-S و آنتروپی دارای سطح نامحسوسی قابل قبول است.

واژه‌های کلیدی: نشان‌گذاری، روش مبتنی بر فاصله، شبکه گمنامی

۱- مقدمه

نشان‌گذاری[۱]، کاربردهای متفاوتی در علوم مختلف ازجمله کامپیوتر دارد. از نشان‌گذاری برای مشخص‌کردن حق کپی محصولات چندرسانه‌ای استفاده می‌شود. در دهه اخیر از نشان‌گذاری برای نشانه‌دارکردن جریان شبکه به‌عنوان تحلیل ترافیک شبکه جهت شناسایی بهتر جریان‌ها، در مقایسه با تحلیل ترافیک غیرفعال و منفعل استفاده شده است. تاکنون طرح‌های نشان‌گذاری مختلفی طراحی شده که اساساً از ایده‌های نشان‌گذاری چندرسانه‌ای اقتباس شده‌اند ولی به‌طورکلی می‌توان تغییر در الگوی محتوای جریان شبکه را نشان‌گذاری ترافیک شبکه نامید [۱]، مانند ایجاد تغییر در زمان‌بندی بسته‌ها که باعث نشان‌گذاری در ترافیک شبکه می‌شود. نشان‌گذاری‌ها یک روش مستقل از محتوا را برای برچسب‌زدن به ترافیک ارائه می‌دهندکه جریان‌های همبسته بعد از گذشت از شبکه‌های مختلف قابل بازشناسی می‌باشند [۱].

کاربردهای نشان‌گذاری جریان شبکه شامل ردیابی مبدأ نفوذ، کشف گمنامی ایجادشده در شبکه‌های گمنامی و شناسایی شبکه بات می‌باشد. مهاجمین شبکه، معمولاً کوشش می‌کنند که مکان واقعی خودشان را با استفاده از بازتاب ترافیک[۲] خود از طریق تعدادی میزبان آلوده[۳]، پنهان کنند و بدین صورت ردپا یا هویت خود را پنهان می‌کنند. نشان‌گذاری ترافیک شبکه می‌تواند در شناسایی ترافیک‌های بازتاب شده استفاده شود، اگر جریان شبکه در نشان‌گذار نشان‌گذاری شود، در صورت عبور ترافیک نشان‌شده از کدگشا موجود در شبکه، آن جریان نشان شده، قابل

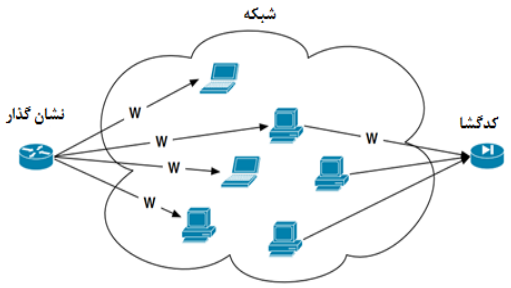
* رایانامه نویسنده مسئول: mdehghany@ihu.ac.ir

1- Watermark

2- Relaying

3- Zombi

تشخیص می‌باشد، شکل (۱) این موضوع را نمایش می‌دهد [۲].



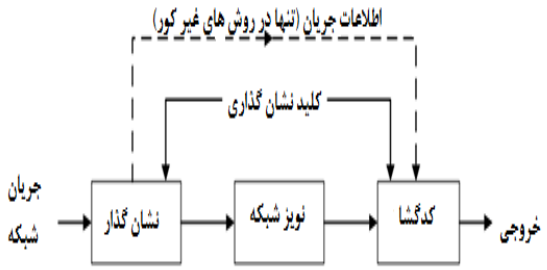
شکل (۱): ردیابی با استفاده از نشان‌گذاری [۲].

یکی از روش‌های نشان‌گذاری، روش مبتنی بر فاصله است. این روش به‌علت مسئله نقاط مرزی در اجرا باعث ایجاد تاخیرها و ناپایداری شده و در نتیجه در ردیابی نفوذ در شبکه گمنامی مسیریابی پیازی[1] دارای ضعف و ناکارآمدی می‌باشد. در این مقاله به این مسئله پرداخته شده و راه‌حلی برای بهبود آن پیشنهاد می‌شود و سپس با انجام آزمایش‌ها در محیط واقعی و ارزیابی روش پیشنهادی و مقایسه آن با روش مبتنی بر فاصله، ثابت خواهیم کرد که کارایی این روش بهبود یافته است.

در ساختار این مقاله ابتدا پیش زمینه در مورد روش نشان‌گذاری مبتنی بر فاصله در بخش ۲ مطرح می‌شود و در ادامه فعالیت‌های مرتبط در بخش ۳ مرور می‌شود سپس در بخش ۴ مسئله نقاط مرزی در روش مبتنی بر فاصله بررسی می‌شود و در بخش ۵ روش پیشنهادی حل مسئله نقاط مرزی ارائه شده و با ارزیابی روش پیشنهادی در بخش ۶ کارایی آن اثبات می‌گردد و در انتها نیز نتیجه‌گیری در بخش ۷ بیان می‌شود.

۲- پیش زمینه

نحوه کار نشان‌گذاری جریان شبکه بدین صورت است که وقتی یک جریان که از نشان‌گذار[2] عبور می‌کند، به‌وسیله تغییردادن اطلاعات زمان‌بندی بسته‌ها نشان‌گذاری می‌شود. برای مثال با به‌کاربردن تأخیر مشخصی بر روی بسته‌ها نشان‌گذاری انجام می‌شود. جریان از امتداد یک کانال نویزدار عبور می‌کند که ممکن است شامل شبکه‌های مختلف و سامانه‌های گمنامی باشد. این کانال تأخیرهای اضافی را نشان می‌دهد و ممکن است باعث از بین رفتن، جابه‌جایی و تکثیر بسته‌ها گردد. پس از گذشت از کانال، جریان به قسمت کدگشا[3] می‌رسد؛ که این قسمت جریان را برای یافتن الگوی نشان‌گذاری شده مندرج در آن، تفتیش می‌کند که الگوی کدگذاری باید مبتنی بر کلید محرمانه نشان‌گذاری باشد؛ که بین نشان‌گذار و کدگشا اشتراک گذاشته شده است [۳]. شکل (۲)، مدل عمومی نشان‌گذاری جریان شبکه را نمایش می‌دهد. کلید یا متغیرهای اصلی نشان‌گذاری از قبل بین کدگشا و نشان‌گذار به اشتراک درآمده است. به‌طورکلی، جریانی نشان‌گذاری می‌شود که احتیاج به شناسایی آن در خروجی‌های شبکه باشد. برای مثال، اگر جریانی در شبکه، به‌عنوان جریان مخرب تشخیص داده شد، آن‌گاه برای ردیابی مبدأ و منبع آن جریان به نشان‌گذار دستور داده می‌شود که جریان مخرب را نشان‌گذاری کند.



شکل (۲): مدل عمومی نشان‌گذاری جریان شبکه [۳].

حال اگر نفوذگر از یکی از میزبان‌های داخل شبکه برای نفوذ استفاده کند و به نوعی آدرس خود را پنهان کند، جریان پس از عبور از آن برای خروج از شبکه باید از درگاه‌های موجود استفاده کند که در این صورت هم‌چنان نشان‌گذاری انجام‌شده باقی می‌ماند که در صورت استقرار کدگشا در درگاه خروجی، جریان نشان‌گذاری‌شده تشخیص داده می‌شود و مبدأ اصلی و خارج از شبکه قابل تشخیص و شناسایی می‌باشد و می‌توان جریان تشخیص داده شده را در این‌جا ردیابی و مسدود کرد [۳]. یکی از روش‌های پنهان‌سازی مبدأ، درخواست استفاده از شبکه گمنامی است. شبکه‌های گمنامی، در مسیر رسیدن پیام از فرستنده به گیرنده پیام‌ها به‌طور مداوم و تکراری در نودهای شبکه که به آن‌ها مسیریاب پیازی می‌گویند، رمزنگاری و رمزگشایی می‌شوند. شبکه مسیریابی پیازی شبکه‌ای است که با استفاده از تونل‌های مجازی به افراد و گروه‌ها اجازه می‌دهد تا امنیت و حفاظت از حریم خصوصی خود را در اینترنت تقویت کنند. در هنگام دریافت پیام در شبکه مسیریابی پیازی، آخرین لایه از رمزنگاری پیازی را با کلید خصوصی خود رمزگشایی می‌کند این کار جهت آگاهی نسبت به نحوه ارسال پیام به مسیریاب پیازی بعدی صورت می‌گیرد. امکان این وجود ندارد که کسی با این رمزنگاری

1- The Onion Routing (TOR)
2- Watermarker
3- Decoder

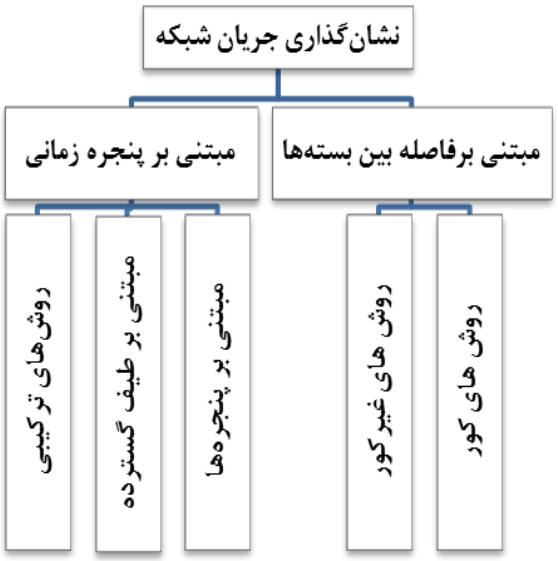
و رمزگشایی قادر به شناسایی مبدأ، مقصد و محتوای پیام باشد. یک سامانه گمنامی یک تعدادی جریان‌های ورودی را به تعدادی جریان‌های خروجی نگاشت می‌کند، درصورتی‌که رابطه بین آن‌ها پنهان می‌باشد و شبیه به یک شبکه ترکیبی، مسیریابی پیازی و یا یک پیشکار[1] ساده پیاده‌سازی شده است. هدف یک نفوذگر آن است که برای استتار خود، ترافیک خود را به یک جریان ورودی خروجی پیوند زند (و بالعکس). بافرض این‌که کدگشا و نشان‌گذار در خارج از دروازه‌های شبکه گمنامی می‌باشد، نشان‌گذاری جریان شبکه می‌تواند در از بین بردن حافظ گمنامی نفوذگر استفاده شود.

۲-۱- روش‌های نشان گذاری

روش‌های مختلفی برای نشان‌گذاری ترافیک شبکه، ابداع شده است. معمولاً این روش‌ها از روش‌های نشان‌گذاری دیجیتال الهام گرفته شده‌اند. این روش‌ها در بعضی از موارد با ترکیب دو یا چند روش، به‌صورت یک روش ترکیبی ارائه شده است. در بعضی از موارد با افزودن الگوریتم خاص به روش‌های قبلی، آن را اصلاح و بهینه کرده‌اند و در برابر بعضی از حملات، آن‌ها را مقاوم کرده‌اند. دسته‌بندی‌های متفاوتی از دیدگاه‌های مختلف برای طرح‌های نشان‌گذاری مطرح شده است. از منظر قابل تشخیص‌بودن جریان نشان‌گذاری‌شده در فاصله بین کدگشا و نشان‌گذار می‌توان طرح‌های نشان‌گذاری را به دو دسته محسوس و نامحسوس تقسیم نمود. از منظر هماهنگی بین کدگشا و نشان‌گذار این روش‌ها را می‌توان به دو دسته کور و غیرکور تقسیم‌بندی کرد. با مطالعه و بررسی تمامی روش‌های نشان‌گذاری ترافیک شبکه، تقسیم‌بندی جامعی از روش‌های نشان‌گذاری در شکل (۳) ارائه شده است [۴] که می‌توان تمامی آن‌ها را به دو قسمت مبتنی بر فاصله بین بسته‌ها و مبتنی بر پنجره زمانی تقسیم کرد. در روش‌های مبتنی بر فاصله بین بسته‌ها [۵-۷] منحصراً بر روی تأخیرهای بین بسته‌ها عمل می‌کند. این روش‌ها در برابر تلفات بسته‌ها، جابه‌جایی و جا اندازی[2] مقاوم نیستند. ولی در روش‌های مبتنی بر پنجره زمانی [۲ و ۸-۱۶] یک عمل را در یک فاصله زمانی کامل انجام می‌دهد که این نوع نشان‌گذاری‌ها مستعد حمله چندجریانی[3] می‌باشند که البته در بعضی از آن‌ها، با اعمال تغییراتی در برابر حملات مقاوم می‌شوند. در روش‌های مبتنی بر فاصله بین بسته‌ها سه روش تاکنون ارائه شده است که این بخش را می‌توان در دو قسمت کور و غیرکور تقسیم‌بندی کرد.

روش‌های مبتنی بر پنجره زمانی تنوع بیشتری را به خود اختصاص داده‌اند، به‌طوری‌که می‌توان آن را به سه قسمت مختلف مبتنی بر پنجره‌ها، مبتنی بر طیف گسترده و روش‌های ترکیبی دسته‌بندی کرد.

نتایجی که در بررسی تحلیلی روش‌ها به‌دست آمده است، این بود که روش‌های مبتنی بر پنجره زمانی نسبت به روش‌های مبتنی بر فاصله بین بسته‌ها از خود مقاومت بیشتری در برابر حملات دارند و تمامی روش‌ها به‌جز روش‌های نسل اول دارای نامحسوسی می‌باشند [۱۷].



شکل (۳): دسته‌بندی روش‌های نشان‌گذاری [۴].

۲-۲- روش نشان‌گذاری مبتنی بر فاصله[4]

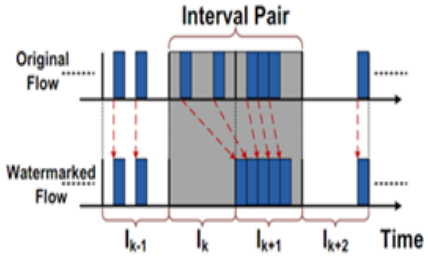
روش مبتنی بر فاصله یکی از روش‌های مبتنی بر پنجره زمانی می‌باشد که توسط پیون و همکارانش ارائه شده است و با روش مبتنی بر مرکز فاصله[5] شباهت دارد [۱۸].

در کل رویکرد و طرح مبتنی بر فاصله در برابر بسته‌بندی‌کردن مجدد جریان‌ها استحکام دارد. این روش جزء روش‌هایی است که نرخ بسیار پایین خطای مثبت اشتباهی و منفی اشتباهی دارد. زمان‌های ورود بسته‌ها براساس یک مجموعه از فاصله‌های زمانی از قبل انتخاب شده دست‌کاری می‌شود. درج نشان‌گذاری با استفاده از تغییرات نرخ و آهنگ ترافیک در فاصله‌های متوالی حاصل می‌شود. این طرح در ابتدا ارائه از استحکام مناسبی

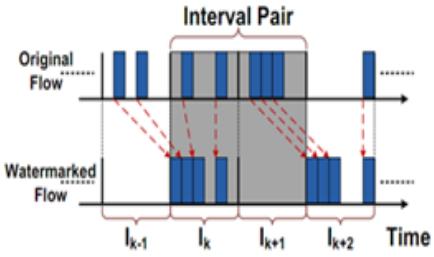
1- Proxy
2- Insertions
3- Multi Flow Attack
4- Interval Based Watermark
5- Interval Centroid Based Watermarking

برخوردار بوده است ولی با ارائه حملات چندجریانی استحکامی در برابر این حمله نداشته و نامحسوسی خود را از دست داده است.

شکل (۴) نحوه نشان‌گذاری مبتنی بر فاصله را نمایش می‌دهد. در این طرح دو زوج پنجره برای درج نشان‌گذاری درنظر گرفته شده است که برای ارسال بیت‌های نشان‌گذاری استفاده می‌شوند، همان‌طورکه در شکل (۴) مشخص است، نحوه قراردادن بسته‌ها در این دو زوج فاصله زمانی براساس بیت نشان‌گذاری متفاوت است.



الف)نشان گذاری مبتنی بر فاصله درج مقدار یک



ب) نشان گذاری مبتنی بر فاصله - درج مقدار صفر

شکل (۴): نشان‌گذاری به روش مبتنی بر فاصله [۱۴].

قبل از اجرای این روش، قسمت‌های نشان‌گذار و کدگشا متغیرهای اصلی را به اشتراک می‌گذارند. متغیرهای اصلی و اشتراکی شامل T>۰ طول فاصله نشان‌گذاری، W مقدار بیت‌های نشان‌گذاری با L>۰ بیت و همچنین O>۰ آفست شروع زمانی نشان‌گذاری در سمت نشان‌گذار و r>۰ نیز تعداد تکرار یا افزونگی در مقدار نشان‌گذاری می‌باشد. همچنین S به‌عنوان تابع انتخاب فاصله به‌عنوان متغیرهای اشتراکی می‌باشند؛ که این تابع باید بتواند افزونگی r را برای تعداد معین بسته به بیش‌ترین مقدار برساند و همچنین این تابع باید بتواند مجموع تعداد بسته‌ها را بدون تأثیر در دیگر فاکتورها به کم‌ترین برساند و بتواند در انتخاب درج نشان‌گذاری در جریان به‌طور تصادفی عمل کند. جدول (۱) متغیرهای مربوط به روش مبتنی بر فاصله را توضیح می‌دهد.

جدول (۱): شرح متغیرهای اصلی و اشتراکی در روش مبتنی بر فاصله

متغیر	توضیح متغیر
T	طول فاصله نشان‌گذاری
W	مقدار بیت‌های نشان‌گذاری
L	تعداد بیت‌های مقدار نشان‌گذاری
R	تعداد تکرار در مقدار نشان‌گذاری
O	آفست شروع زمانی نشان‌گذاری در سمت نشان‌گذار

برای ارسال یک جریانی که با L بیت نشان‌گذاری شده است و دارای r بار افزونگی است نیاز به 3*L*r فاصله می‌باشد که هر بیت نشان‌گذاری با دیگر بیت‌ها تداخلی ندارد.

برای آن‌که مراحل کار در نشان‌گذاری مبتنی بر فاصله بهتر درک شود ابتدا متغیرهای را که برای شرح مراحل نیاز است، معرفی می‌کنیم. فرض می‌شود i=1…r نشان‌دهنده شماره تکرار درج نشان‌گذاری می‌باشد و j=1…L شماره مکان بیت نشان‌گذاری از ۱ تا L می‌باشد. برای نمایش هر بیت از نشان‌گذاری یک پنجره شامل سه فاصله با طول T درنظر گرفته می‌شود. فاصله اول را با $I_{i,j,1}$ و فاصله دوم و سوم را به ترتیب با $I_{i,j,2}$ و $I_{i,j,3}$ نمایش می‌دهیم که از این زوج فاصله جهت درج نشان‌گذاری استفاده می‌شود. به‌طور مثال، $I_{2,4,3}$ نمایان‌گر فاصله سوم مربوط به بیت چهارم نشان‌گذاری و در دومین تکرار نشان‌گذاری می‌باشد. برای درک بهتر فرآیند موجود در روش مبتنی بر فاصله بررسی می‌شود:

فرآیند کار در نشان‌گذار:

وقتی جریان بالادستی برای نشان‌گذاری وارد نشان‌گذار می‌شود، اعمال زیر برای هر بسته ورودی Pk انجام می‌شود:

۱- ابتدا با استفاده از تابع انتخاب فاصله S شاخص فاصله آن شناسایی می‌شود؛ که در این‌جا طول فاصله T و آدرس و نقطه شروع نیز O است.

۲- تأخیر بسته Pk به‌اندازه T به‌صورت زیر محاسبه می‌شود:

اگر Pk در فاصله Ii,j,1 باشد آن‌گاه استنتاج می‌کنیم که wj=0 است.

اگر Pk در فاصله Ii,j,2 باشد آن‌گاه استنتاج می‌کنیم که wj=1 است.

اگر Pk در فاصله Ii,j,3 باشد آن‌گاه استنتاج می‌کنیم که wj=0 است.

اگر بسته Pk به‌اندازه‌ای به تأخیر بیفتد تا در فاصله بعدی قرار بگیرد، مقدار تأخیر می‌تواند کم‌ترین مقدار ممکن یا تصادفی باشد، به‌شرطی‌که بسته Pk به‌طور صحیح در فاصله مربوط به خود قرار بگیرد. البته تصادفی‌کردن مقدار تأخیر در برابر مداخلات

تحلیل زمانی نفوذگران جهت تخریب و حمله به نشان‌گذاری مطلوب‌تر می‌باشد.

مورد بعدی که باید در نشان‌گذاری مبتنی بر فاصله رعایت شود، آن است که اگر بسته Pk-1 به‌گونه‌ای تأخیر داشته باشد که ترتیب خود را با بسته Pk از بین برود برای حفظ ترتیب انتقال بسته‌ها و جلوگیری از بسته‌بندی مجدد در جریان خروجی، باید بسته Pk تأخیر داشته باشد که در این موارد از تأخیر حداقلی بین بسته‌های Pk و Pk-1 استفاده می‌شود [۱۹].

فرآیند کار در کدگشا:

کدگشا با مقادیر و پارامترهای یکسان T,S,w که در نشان‌گذار استفاده شده است هم اکنون می‌تواند جهت کدگشایی نشان‌گذاری w' از یک جریان خروجی و تطبیق آن با w استفاده کند. برای بررسی نحوه مراحل کدگشایی ابتدا مفاهیم و متغیرهای بررسی می‌شود. مقادیر xi,j,2 و xi,j,3 تعداد بسته‌های دریافت شده به‌ترتیب در فاصله‌های Ii,j,2 و Ii,j,3 می‌باشند. O' آفست شروع زمانی در سمت کدگشا و w'i را مقدار نشان‌گذاری کدگشایی شده در i ام تکرار از تعداد افزونگی r قرار می‌دهیم. W'I,j را نیز بیت j ام مقدار نشان‌گذاری کدگشایی‌شده از تکرار i ام نمایش می‌دهیم. فرآیند کدگشایی به‌صورت زیر می‌باشد:

۱- تنظیم اولیه آفست کدگشایی O' به O-|δ| که |δ| بیانگر بیش‌ترین یا بالاترین اختلاف زمانی بین نشان‌گذار و کدگشا می‌باشد.

۲- برای هر بسته ورودی Pk، با استفاده از تابع انتخاب فاصله S، شاخص یا اندیکس فاصله آن‌را با طول فاصله T و آفست O' شناسایی می‌کنیم.

۳- برای هر زوج فاصله Ii,j,2 و Ii,j,3 به‌ترتیب مقادیر xi,j,2 و xi,j,3 محاسبه و ذخیره می‌شود.

۴- بعد از کامل‌شدن هر تکرار i از r، برای هر بیت نشان‌گذاری، Xi,j,2 و Xi,j,3 به‌صورت زیر محاسبه می‌شوند:

$$X_{i,j,3} = i^{-1} \sum_{u=1}^{i} x_{u,j,3} \quad \text{و} \quad X_{i,j,2} = i^{-1} \sum_{u=1}^{i} x_{u,j,2}$$

آن‌گاه برای محاسبه نشان‌گذاری i امین افزونگی به‌صورت زیر عمل می‌کنیم:

هرگاه که Xi,j,3 < Xi,j,2 باشد، آن‌گاه w'i,j = 0 است.

هرگاه که Xi,j,3 > Xi,j,2 باشد، آن‌گاه w'i,j = 1 است.

در غیر این صورت، برای i>1 داریم w'i,j= w'i-1,j است. (برای i=1 داریم w'1,j = 0)

با مقایسه w'i با w، اگر فاصله همینگ کوچک‌تر مساوی حد آستانه کدگشایی شد، H(w, w'i)≤ h آن‌گاه نشان‌گذاری به‌درستی کدگشایی و شناخته شده است و نتیجه مثبت است.

(که در این‌جا h حد آستانه می‌باشد و مقدار آن بین ۰ تا L است.)

۵- گام‌های فوق را تا زمانی ادامه می‌دهیم که یا با r امین تکرار نشان‌گذاری به جواب مثبت برسیم یا جریان به انتها برسد.

۶- اگر در انتهای مرحله چهارم نتیجه منفی باشد باید مراحل از گام دوم مجدداً تکرار شود و به O' مقدار v (مقدار v بین صفر تا T است) را اضافه شود. البته می‌توان در محاسبات واقعی و زیر بار هر آزمایش را با آفست‌های مختلف و به‌صورت محاسبات موازی انجام داد.

یکی از موارد استفاده حد آستانه فاصله همینگ برای خارج‌کردن نویز در فرآیند کدگشایی شامل تغییر شکل‌های فعال[1] می‌باشد. گرچه حد آستانه بالاتر، باعث افزایش نرخ شناسایی می‌شود و ممکن است که یک جریان ارتباطی نشان‌گذاری شده را به‌درستی تشخیص دهد ولی باعث افزایش نرخ مثبت اشتباهی[2] خواهد شد؛ که در این صورت احتمال دارد که یک جریان ارتباطی سالم را به اشتباه، به‌عنوان یک جریان نشان‌گذاری‌شده کدگشایی کند [۱۹].

البته محاسبه نرخ کدگشایی منتظره و نرخ مثبت اشتباهی منتظره برای L بیت نشان‌گذاری با استفاده از حد آستانه فاصله همینگ در تحقیقات سابق [۷ و ۲۰] انجام شده است و پیشنهاد شده است که h به‌گونه‌ای انتخاب گردد که یک تراز و موازنه بین نرخ کدگشایی و نرخ مثبت اشتباهی برقرار شود.

۲-۳- معیارهای ارزیابی نشان‌گذاری

در مجموع می‌توان معیارهای ارزیابی روش‌های نشان‌گذاری را در دو مورد استحکام[3] و نامحسوسی[4] خلاصه کرد. برای به‌دست‌آوردن و مقایسه کارایی روش‌های نشان‌گذاری می‌توان براساس این معیارها آن‌ها را ارزیابی و بررسی کرد. معیار استحکام، میزان سختی حذف نشان‌گذاری یا محدودکردن نرخ کدگشایی صحیح را در کدگشا نشان می‌دهد. معیار نامحسوسی، میزان سختی تشخیص نشان‌گذاری را از طریق مقایسه مشخصات ترافیک با ترافیک مجاز نشان می‌دهد. یا به‌عبارتی دیگر، کاربرد نامحسوسی توانایی پنهان‌بودن نشان‌گذاری از تشخیص و شناسایی در مسیر جریان می‌باشد؛ زیرا درصورت محسوس‌بودن نشان‌گذاری، شناسایی آن توسط نفوذگرها و نفوذگران حرفه‌ای قابل انجام خواهد بود لذا آن‌ها می‌توانند با تشخیص وجود نشان‌گذاری، طرحی برای حمله به نشان‌گذاری را پیاده‌سازی کنند و آن‌را از بین ببرند. برای تشخیص نامحسوسی روش نشان‌گذاری مبتنی بر

1- Active
2- False Positive Rate
3 -Robustness
4- Stealthness

فاصله از آزمون‌های کولموگروف اسمیرنوف[1]، آنتروپی تصحیح شده[2] و آنتروپی شرطی تصحیح‌شده[3] استفاده شده است. آزمون کولموگروف اسمیرنوف یا آزمون K-S یک آزمون غیرپارامتریک مفید و عمومی برای مقایسه این‌که آیا دو نمونه متعلق به توزیع مشابه هستند یا نه، استفاده می‌شود. چون آزمون K-S غیرپارامتریک است، متکی بر هیچ فرضی در رابطه با توزیع نیست و به‌همین دلیل، می‌تواند برای هر توزیع مورد استفاده قرار گیرد [۲۱]؛ اما محدود به توزیع‌های پیوسته است.

پنگ و همکارانش نشان دادند که چه‌طور با آزمون کولموگروف- اسمیرنوف می‌توان در مدولاسیون شاخص میزان[4]، نشان‌گذاری‌ها که به‌صورت تأخیر بین بسته‌ای درج شده است کارایی را شناسایی کرد [۲۲]. در تشخیص مبتنی بر آنتروپی[5] سعی در اندازه‌گیری شباهت نسبی بین ترافیک مجاز و ترافیک مشکوک به جریان نشان‌گذاری است. از آن‌جایی‌که در جریان نشان‌گذاری‌شده اطلاعات با تغییر در زمان‌بندی تاخیرهای بین بسته‌ای کدگذاری می‌شود، توزیع زمان‌های بین بسته‌ها را تحت تاثیر قرار می‌دهد. برای مشاهده چنین تغییری در توزیع از میزان آنتروپی استفاده می‌شود. محاسبه آنتروپی براساس توزیع نمونه‌های تاخیرهای بین بسته‌ای ترافیک مجاز شناخته شده انجام می‌شود. هرگونه انحراف از این توزیع آموزشی باعث افت آنتروپی محاسبه شده می‌شود [۲۳]. روش تشخیص مبتنی بر آنتروپی شرطی تصحیح‌شده را برای تمایز بین قاعده‌مندی ترافیک جریان نشان‌گذاری از ترافیک سالم ارائه کرده است [۲۳].

۳- فعالیت های مرتبط

یکی از اولین روش‌های نشان‌گذاری، روش ونگ و همکارانش می‌باشد. آن‌ها نشان‌گذاری مبتنی بر IPD[6] را معرفی می‌کنند که قبل از اجرا نیازمند آن است که بین نشان‌گذار و کدگشا هماهنگی مستحکمی انجام شود [۷]. ونگ و همکارانش، ICBW[7] را ارائه کردند. این طرح مبتنی بر تقسیم‌بندی جریان در فاصله‌های با طول مساوی می‌باشد [۱۵]. پیون و همکارانش طرح IBW[8] را ارائه کرده‌اند. در این طرح، نشان‌گذاری‌کننده و شناسایی‌کننده بر روی پارامترهای محرمانه هماهنگ می‌شوند [۱۴]. این طرح در برابر بسته‌بندی مجدد[9] استحکام دارد. البته این روش دارای نرخ‌های بسیار پایین در خطای مثبت اشتباهی و خطای منفی اشتباهی می‌باشد ولی در برابر حملات چندجریانی مقاومت ندارد [۲۴]. ژنگ و همکارانش، طرح IBW را در برابر حملات چندجریانی بهینه کرده‌اند [۲۵]. هومن صدر و همکارانش روشی ارائه می‌دهند که روشICBW را در برابر حملات چندجریانی مقاوم می‌کند و روش جدید را MAR-ICBW[10] نامیده‌اند [۱۳]. ونگ و همکارانش در تحقیقاتی دیگر روش جدیدی را با استفاده از روش ICBW و ترکیب آن با تکنیک کدگذاری طیف گسترده، ارائه می‌دهند که نام آن را با توجه به شیوه استفاده‌شده در آن DICBW[11] قرار می‌دهند که کارایی و محرمانگی بالاتری نسبت به ICBW دارد و قادر به ردیابی جریان‌های چندگانه می‌باشد [۱۰].

یو و همکارانش طرح DSSS[12] را ارائه کرده‌اند که این روش یک نشان‌گذاری دودویی می‌باشد که در جریان تعبیه شده و نامحسوس است ولی در برابر حملات چندجریانی آسیب‌پذیر است [۱۶]. با توجه به این‌که روش DSSS برای ردیابی جریان‌ها نیاز به نرخ ثابتی دارد، لو ژنگ و همکارانش روشی را برای بهبود این محدودیت ارائه دادند [۱۲]. هم‌چنین لیو و همکارانش روش ترکیبی جدیدی را با استفاده از روش‌های طیف گسترده و ICBW به نام ICBSSW ارائه کرده‌اند [۸]. هومن صدر و همکارانش روش RAINBOW را ارائه می‌دهند که مبتنی بر بسته می‌باشد، این طرح به‌علت استفاده از یک سامانه هماهنگ‌کننده که توسط نشان‌گذار و کدگشا قابل دسترس است، که این روش اولین روش نشان‌گذاری جریان به‌صورت غیرکور است [۶]. آن‌ها برای بالابردن کارایی شناسایی در RAINBOW از الگوریتم رمز تجمع تکراری[13] استفاده و روش C-RAINBOW را ایجاد می‌کنند [۵]. لیانگچن ژنگ و همکارانش روشی به نام MMAR-SSW را ارائه می‌کنند که در برابر حملات MSAC و MFA مقاوم می‌باشد. این روش دارای محرمانگی بیش‌تری است [۱۱]. هوآنگ و همکارانش روشی را برای بهبود DSSS ارائه می‌دهند. در این روش با طولانی‌ترکردن شبه نویز در روش سابق، نامحسوسی نشان‌گذاری را افزایش می‌دهند [۹]. هومن‌صدر و همکارانش با ارائه طرح جدید SWIRL اولین رویکرد عملی را برای تحلیل ترافیک مقیاس بزرگ آماده کرده‌اند [۲].

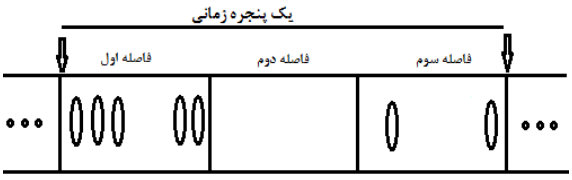
۴- مسئله نقاط مرزی در روش مبتنی بر فاصله

یکی از مشکلات مهم در روش مبتنی بر فاصله این است که این

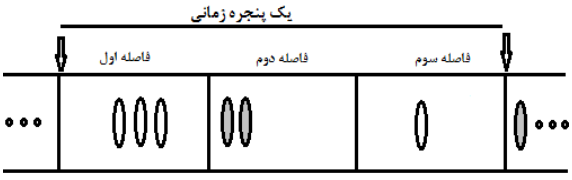
1- Kolmogorov-Smirnov test
2- Corrected Entropy
3- Corrected Conditional Entropy(CCE)
4- QIM
5- Entropy Based Detection
6- Inter Packet Delay
7- Interval Centroid-based Watermarking
8- Interval-Based Watermarking
9- Repacketzation
10- Multi-flow Attack Resistant–Interval Centroid Based Watermark.
11- Double Interval Centroid Based Watermark.
12- Direct Sequence Spread Spectrum.
13- Repeat-Accumulate Code

روش در برابر ردیابی جریان در شبکه گمنامی دارای ضعف می‌باشد. از این‌رو، با بررسی گام‌های روش و اجرای آن در شبکه گمنامی مسیریابی پیازی راه‌حلی را برای بهبود آن پیشنهاد می‌شود که بتواند کارایی بهتر در شبکه گمنامی داشته باشد. شبکه گمنامی مسیریابی پیازی یا تور برای امنیت ارتباطات، هر ارتباط را چندبار در ایستگاه‌های کاری شبکه به‌صورت رمزشده در داخل تونل‌های ارتباطی قرار می‌دهد؛ که این خود باعث می‌گردد که زمانی خاص برای تونل‌سازی ارتباط و رمزنگاری و هم‌چنین بسته‌بندی مجدد علاوه‌بر تأخیرهای رایج مانند لغزش زمانی و زمان مورد نیاز برای ارسال، به سرجمع زمان اضافه می‌گردد که این خود باعث به‌هم‌زدن الگوی ارسال روش مبتنی بر فاصله خواهد شد.

البته درصورتی‌که بسته‌های جریان مورد نظر دارای حجم کم باشند، مشکل بسته‌بندی مجدد برطرف خواهد شد ولی تأخیرهایی که ناشی از تونل‌سازی می‌باشد، باعث می‌گردد که بسته‌هایی که در مرزهای انتهای فاصله‌های یک پنجره زمانی روش مبتنی بر فاصله می‌باشند، با اعمال این‌گونه تأخیرها، وارد فاصله بعدی یا حتی وارد پنجره زمانی بعدی گردند. شکل (۵) این مشکل را نمایش می‌دهد. همان‌طوری‌که در شکل (۵- الف) ملاحظه می‌شود نشان‌گذار می‌خواهد بیت صفر را در نشان‌گذاری ارسال کند لذا آرایش بسته‌ها را به‌صورت شکل (الف) می‌چینند.



الف) نحوه قرار گرفتن بسته ها در یک پنجره در نشان گذار



ب) نحوه دریافت بسته ها در یک پنجره در کدگشا

شکل (۵): مشکل محاسبه نشان‌گذاری در روش مبتنی بر فاصله.

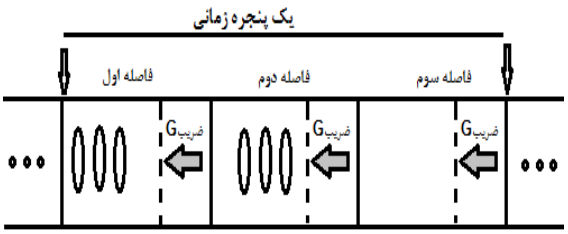
بعضی از بسته‌ها به‌طور اتفاقی در حاشیه و مرز فاصله‌ها قرار می‌گیرند. با گذشت از شبکه گمنامی و اعمال تأخیرهای تونل‌سازی به بسته‌ها، مجموع تأخیرها در بعضی از بسته‌ها بالا می‌رود. بعد از رسیدن به کدگشا بسته‌ها به‌علت تأخیرها جابه‌جا می‌شوند که در این مواقع محاسبات قسمت کدگشا دچار اشتباه شده و نرخ خطای مثبت اشتباهی و منفی اشتباهی در کدگشایی بالا می‌رود و در نتیجه با کاهش نرخ کدگشایی روبرو شده و لذا کارایی نشان‌گذاری پایین خواهد آمد.

به‌طورمثال، شکل (۵- ب) در مقایسه تعداد بسته‌های موجود در فاصله دوم و فاصله سوم، تعداد بسته‌ها در فاصله دوم بیش‌تر خواهد بود درحالی‌که در حقیقت تعداد بسته‌ها در فاصله دوم باید کم‌تر باشند؛ که این خود باعث مشکل محاسبه نشان‌گذاری می‌شود.

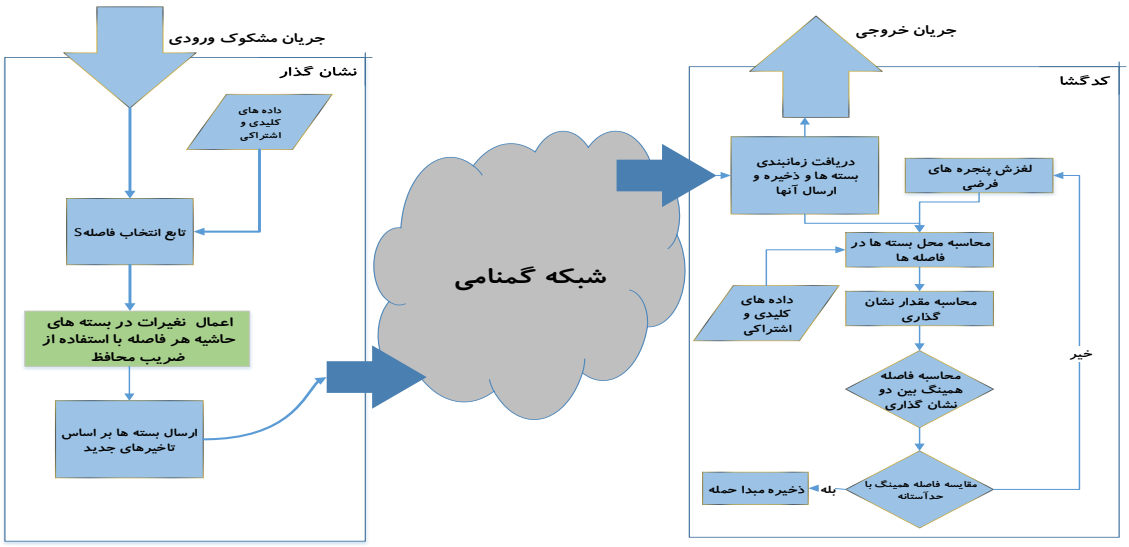
۵- روش پیشنهادی حل مسئله نقاط مرزی

با بررسی یکی از مشکلات اساسی روش مبتنی بر فاصله در شبکه گمنامی مسیریابی پیازی متوجه خواهیم شد که وجود بسته در کران‌های انتهایی فاصله‌ها باعث کاهش کارایی نشان‌گذاری مبتنی بر فاصله می‌گردد. با توجه به این‌که مکان هربسته توسط تابع S در نشان‌گذار محاسبه و در مکان خود قرار داده می‌شود، لذا برای اصلاح روش باید تغییراتی در تابع S اعمال شود. جهت اصلاح روش مبتنی بر فاصله یک ضریب محافظ برای تهی و خالی‌کردن کران و مرزهای انتهای فاصله درنظر گرفته می‌شود.

این ضریب براساس مقدار آن اندازه‌ای را به‌عنوان محافظ، خالی از بسته نگه می‌دارد. این اندازه از زمان که به‌نوعی سکوت ارسال بسته نیز می‌باشد. همین‌طور که در شکل (۶) دیده می‌شود فاصله اول و دوم در محدوده مجاز قرارگرفته‌اند و وارد قسمت بسته‌ها در محافظ نشده‌اند. نحوه اعمال ضریب محافظ G به این صورت است که $\frac{1}{G}$ از کل فاصله زمانی مورد نظر از بسته‌ها تهی می‌گردد که این تغییرات توسط تابع S در نشان‌گذار اعمال می‌شود. در این تابع زمان ارسال هر بسته براساس مقدار بیت نشان‌گذاری تعیین می‌گردد. در هنگام مقررشدن این‌که هر بسته در چه فاصله زمانی ارسال گردد زمان مقررشده با بازه محافظ نیز کنترل می‌شود تا بسته در کران انتهای فاصله قرار نداشته باشد.



شکل (۶): اعمال محافظ در روش پیشنهادی برای فاصله‌ها.



شکل (۷): فرآیند روش پیشنهادی.

اگر بسته در کران انتهایی قرار داشته باشد با کم‌کردن زمان مقررشده، ارسال بسته را در بازه مجاز انتقال می‌دهد. با این تغییرات روش مبتنی بر فاصله در برابر شبکه گمنامی، کارایی بهتری خواهد داشت و نرخ خطای آن به مراتب کم‌تر خواهد شد. با ارائه روش پیشنهادی جهت بهبود روش مبتنی بر فاصله در کل می‌توان فرآیند روش بهبودیافته را به‌صورت شکل (۷) نمایش داد.

همان‌طوری‌که ملاحظه می‌کنید کلیت فرآیند در دو قسمت نشان‌گذار و کدگشا بیان شده است؛ که برای بهبود روش در شبکه مسیریابی پیازی تنها در قسمت نشان‌گذار تغییراتی انجام می‌شود. جریانی که برای نشان‌گذاری وارد می‌شود ابتدا بسته‌های جریان توسط تابع انتخاب فاصله S با توجه به مقادیر متغیرهای اشتراکی در پنجره‌های فرضی قرار داده می‌شوند و در ادامه بسته‌های که در حاشیه انتهایی فاصله‌ها قرار گرفته‌اند به حاشیه امن انتقال داده می‌شوند. سپس با ترتیب درنظر گرفته‌شده بسته‌ها در شبکه گمنامی به سمت مقصد ارسال می‌شوند.

در قسمت کدگشا بسته‌ها دریافت‌شده و پس از استخراج تأخیر بین بسته‌ی بسته‌ها، آن‌ها به سمت مقصد اصلی هدایت می‌شوند. زمان‌بندی دریافت‌شده با توجه به مقادیر متغیرهای اشتراکی در پنجره‌های فرضی قرار گرفته می‌شوند سپس مقدار نشان‌گذاری محاسبه‌شده و با مقدار نشان‌گذاری اصلی مقایسه می‌شوند در این مقایسه مقدار بیت اختلافی به‌عنوان عدد همینگ به‌دست می‌آید. بعد عدد همینگ با حد آستانه نشان‌گذاری مقایسه شده اگر کمتر باشد جریان دریافتی نشان‌گذاری‌شده می‌باشد و در غیر این صورت، لغزشی اندک به پنجره‌های فرضی داده می‌شود و دوباره محاسبه انجام می‌گردد که این کار بنا به‌ضرورت می‌تواند تا چند بار تکرار شود.

۶- روش ارزیابی و آزمایش

جهت ارزیابی استحکام و نامحسوسی روش پیشنهادی در محیط واقعی و در بستر اینترنت آزمایش‌ها انجام می‌گردد. برای شناخت روش ارزیابی ابتدا به بیان شرایط محیط و سپس به بررسی ابزار و روش انجام آزمایش پرداخته می‌شود.

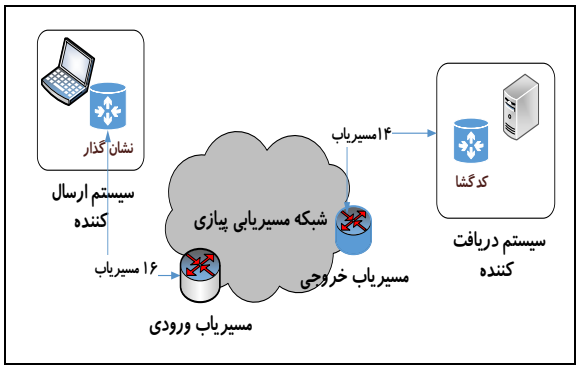
۶-۱- محیط و مفروضات انجام آزمایش

در این تحقیق بر روی نشان‌گذاری ترافیک شبکه گمنامی تمرکز می‌نمایم و در فضای آزمایشگاهی جریان ورودی به یک شبکه گمنامی مسیریابی پیازی را با روش نشان‌گذاری مبتنی بر فاصله نشان‌گذاری کرده و براساس معیارهای استحکام و نامحسوسی به‌صورت آزمایشگاهی آن را بررسی و ارزیابی می‌کنیم. فرض بر این است که جریان مورد نظر از شبکه گمنامی استفاده می‌کند و جریان ورودی به شبکه گمنامی در دسترس می‌باشد؛ و نفوذگر هیچ دسترسی به ایستگاه‌های مسیریابی داخل شبکه گمنامی ندارد. همچنین فرض می‌کنیم که در هنگام آزمایش، مشکلات خاص[۱]، آزمایش را تحت تأثیر خود قرار ندهد و در یک محیط عادی در بستر اینترنت آزمایش‌ها انجام می‌شود. به‌طورکلی در آزمون‌ها بسته‌های جریان ورودی به شبکه گمنامی، بر روی رایانه

۱- هرگونه مشکل خواسته یا ناخواسته که موجب اختلال در آزمایش‌ها گردد. مانند تغیر ناگهانی مسیر در شبکه مسیریابی پیازی که به‌علت نحوه آزمایش و پیاده‌سازی آزمایشگاهی باعث از بین رفتن اتصالات شده که منجر به این می‌شود که کنترل روال برنامه آزمایشگاهی خارج شود.

همراه متصل به شبکه مسیریابی پیازی به‌صورت تصادفی تولید می‌شوند که آن را مربوط به یک مهاجم فرض می‌کنیم و سپس همان رایانه همراه جریان مفروض را نشان‌گذاری کرده و از طریق بستر شبکه مسیریابی پیازی به سمت یک سرور در اینترنت ارسال می‌کند.

جریان با ورود به شبکه گمنامی و عبور از آن وارد سرور مجازی قابل دسترس در اینترنت شده و در آن سرور، برنامه مربوط به کدگشایی جریان را دریافت و درصورت داشتن الگوی نشان‌گذاری آن را شناسایی می‌کند. ما در این تحقیق تنها بخش‌های اصلی سامانه نشان‌گذاری را پیاده‌سازی کرده‌ایم و لذا برای تحلیل ساده‌تر و ملموس‌تر، از مبدأ و مقصد واقعی جریان صرف‌نظر خواهیم کرد. شکل (۸) حالت کلی محیط آزمایشگاهی را نمایش می‌دهد.



شکل(۸): ترسیم محیط و شرایط آزمایشگاه.

برای آماده‌شدن بستر آزمایش‌ها، یک سرور مجازی با سیستم‌عامل Kali (از توزیع‌های امنیتی و رایج لینوکس) در میزبانی[1] واقع در کشور آمریکا اجاره شد و با استفاده از رایانه همراه، اتصال از طریق شبکه مسیریابی پیازی به سرور مجازی برقرار شد. ارتباطاتی که از رایانه همراه به سرور مجازی وجود داشت یکی اتصال putty به پورت SSH سرور برای انتقال دستور به سرور می‌باشد و ارتباط بعدی برنامه نشان‌گذار که اطلاعات را به سمت شبکه مسیریابی پیازی هدایت می‌کرد تا به سرور مجازی یا همان برنامه کدگشا برسد.

معمولاً در هر بار اتصال و اجرای آزمایش‌ها، به‌دلیل پویایی بستر، مسیر و تعداد ماشین‌های موجود در مسیر متفاوت بوده است، ولی به‌طورکل می‌توان عنوان کرد که تعداد ماشین‌های بین رایانه همراه تا ایستگاه ورودی شبکه مسیریابی پیازی حدود ۱۶ ماشین می‌باشد و بعد از خروج از ایستگاه خروجی شبکه گمنامی تعداد ماشین‌ها موجود تا سرور مجازی حدود ۱۴ ماشین است.

زمان‌های آزمون‌ها معمولاً در زمان‌های متفاوت در شبانه‌روز اجرا شده است. به‌طوری‌که به علت تعداد زیاد آزمون‌های برای اجرا در حالت‌های مختلف و اجرای هر آزمون به‌طور متوسط ۱۵ بار برای هر حالت، زمان زیادی را خواهد داشت. لذا در ساعات مختلف شبانه‌روز حتی نیمه‌شب‌ها آزمون‌ها انجام شده است و نتایج آن ثبت گردیده است.

با توجه به عدم وجود مجموعه داده‌های معتبر، به‌دلیل بهانه تحریم‌های غیرقانونی علیه ایرانیان، برای زمان‌بندی بسته‌های ورودی به نشان‌گذار از توزیع پارتو[2] به‌صورت تصادفی استفاده شد که این بسته‌های ورودی دارای تأخیر زمانی میانگین ۲۳۳ میلی‌ثانیه و انحراف معیار ۴۰ میلی‌ثانیه می‌باشند که به‌طور متوسط نرخ ارسال آن‌ها به‌طور میانگین برابر ۴/۲۹ بسته در ثانیه می‌باشد. انتخاب توزیع پارتو به‌این‌علت بود که این توزیع با ترافیک‌های SSH و Telnet در شبکه منطبق می‌باشد که این خود تا حدودی افزایش دقت آزمون‌ها را به همراه خواهد داشت.

برای محاسبه کارایی و استحکام روش پیشنهادی با اجراکردن روش پیشنهادی و روش سابق بر روی محیط واقعی و اعمال تغییرات و اندازه‌گیری و محاسبه دقت نرخ کدگشایی، نرخ‌های مثبت اشتباهی و منفی اشتباهی در بین حدود ۲۰ جریانی که شامل بیش از ۱۰۰۰ بسته می‌باشد در هر نوع آزمون از آزمون‌ها، مقادیری آماری به‌دست آمد.

۶-۲- روش و ابزار انجام آزمایش‌ها

در روش آزمایش به این صورت عمل شده که در سمت سرویس‌گیرنده که فرض شده بود جریان‌ها از آن عبور می‌کنند و سپس وارد شبکه مسیریابی پیازی می‌شوند، ابتدا زمان‌بندی بسته‌ها به‌صورت تصادفی تولیدشده و به تصادف بعضی از جریان‌ها پس از نشان‌گذاری وارد تونل رمز شبکه مسیریابی پیازی می‌شوند و بعضی نیز بدون نشان‌گذاری وارد شبکه مسیریابی پیازی می‌شوند. بعد از گذر از شبکه گمنامی، جریان‌ها وارد سرویس‌دهنده در بستر اینترنت شده که این سرویس‌دهنده همان کدگشا می‌باشد، هر جریان بعد از ورود بسته‌ها الگوی آن استخراج‌شده و درصورت داشتن الگوی نشان‌گذاری، جریان به‌عنوان جریان نشان‌گذاری معرفی می‌شود.

این استخراج الگو به این‌گونه صورت می‌پذیرد که تأخیر بین بسته‌ی بسته‌ها محاسبه‌شده و ذخیره می‌شود و زمان‌بندی را در فایلی مجزا جهت بررسی‌های آماری ذخیره می‌کند. هم‌چنین برای اتصال به شبکه مسیریابی پیازی از نرم‌افزار TOR Browser نسخه ۲/۳/۲۵ استفاده گردید. قسمت‌های نشان‌گذار و کدگشا با

1- Hosting

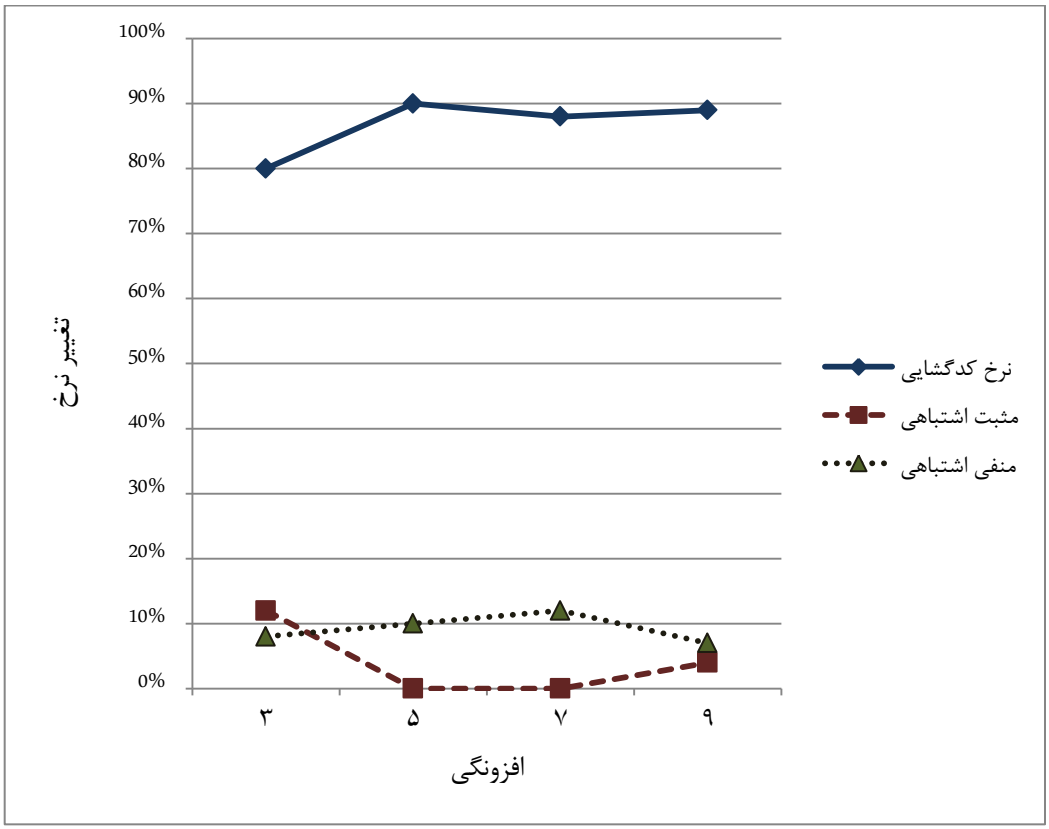
2- Pareto distribution

زبان پایتون پیاده سازی گردید که از زبان پایتون نسخه ۲/۷/۲ هم در دو سمت ارتباط استفاده شد.

فایل‌های متنی ذخیره‌شده توسط نرم‌افزار کدگشایی، شامل تأخیرهای بین بسته‌ای و نمایش محاسبات کدگشا جهت به دست‌آوردن مقدار نشان‌گذاری می‌باشد، از این اطلاعات برای بررسی نامحسوسی از روش آماری مرتبه دوم به روش KS از نرم‌افزار Matlab جهت محاسبات این روش استفاده گردید. البته بعضی محاسبات نیز با نرم‌افزار Excel محاسبه گردید. هم‌چنین برای بررسی نامحسوسی با روش آنتروپی و آنتروپی شرطی از نرم‌افزارهای آزمایش‌شده در [۲۳] استفاده شد، این نرم‌افزارها با زبان C++، الگوریتم‌های دو روش آماری مرتبه دوم را پیاده‌سازی کرده است.

۷- ارزیابی استحکام روش پیشنهادی

برای ارزیابی استحکام روش پیشنهادی ابتدا مطابق مطالب قبلی بسته‌ها تولید و نشان‌گذاری‌شده در بستر شبکه مسیریابی پیازی به‌سمت کدگشا ارسال شده‌اند. حجم بسته‌های ارسالی کم‌تر از ۴۹۸ بایت می‌باشد و فرض شده است در شرایط عادی و بدون هیچ حمله‌ای این ارسال‌ها صورت می‌پذیرد. برای ارزیابی نرخ کدگشایی روش پیشنهادی در شبکه مسیریابی پیازی، متغیر T برابر ۹۰۰ میلی‌ثانیه، تعداد بیت نشان‌گذاری ۱۵ بیت، حد آستانه ۵، ضریب محافظ ۴ و مقدار آفست ۱۰۰ میلی‌ثانیه تنظیم شده است و افزونگی نشان‌گذاری در مقادیر مختلف ۳، ۵، ۷ و ۹ تغییر می‌کند.



شکل(۹): نرخ کدگشایی نشان‌گذاری روش پیشنهادی براساس تغییرات افزونگی.

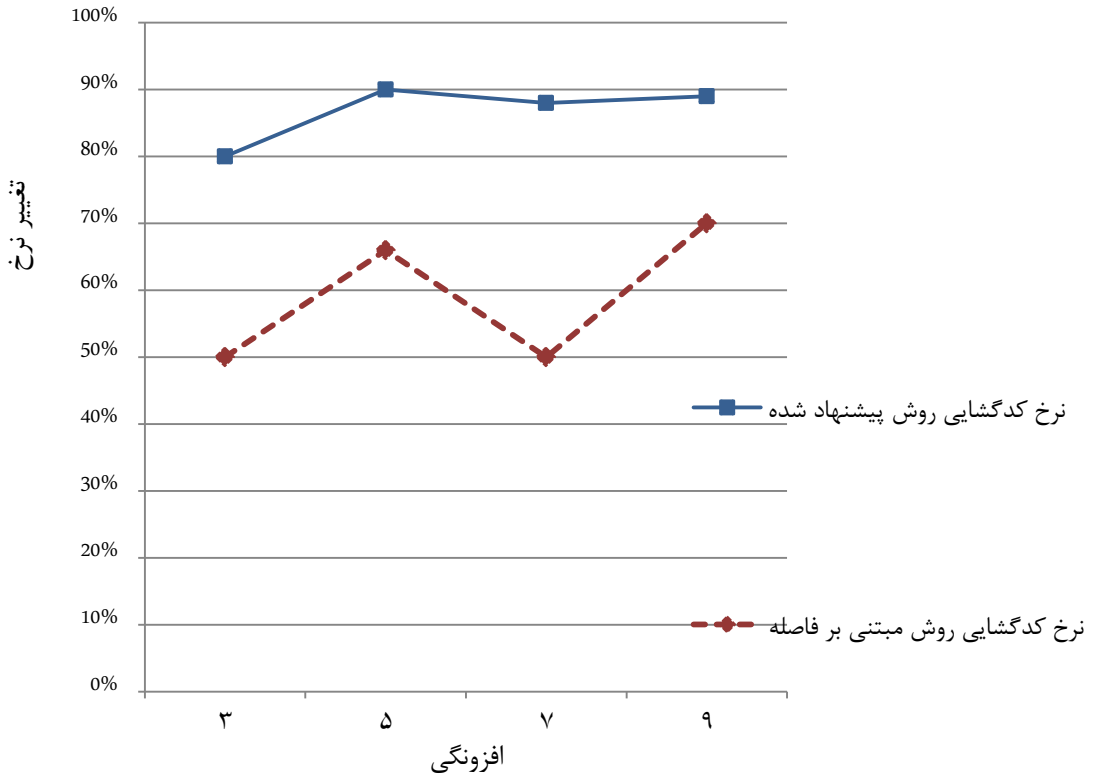
برای محاسبه نرخ کدگشایی ۲۰ جریانی که به‌صورت تصادفی نشان‌گذاری شده‌اند را مورد ارزیابی در کدگشا قرار داده شدند که براساس درصد درست و صحیح کدگشایی جریان‌های نشان‌گذاری شده و نشده، نرخ کدگشایی استخراج می‌شود و براساس درصد نادرست کدگشایی جریان‌های نشان‌گذاری شده و نشده، نرخ‌های مثبت اشتباهی و منفی اشتباهی استخراج می‌شوند.

نمودار شکل (۹)، این تغییرات نرخ را براساس تغییر متغیر افزونگی نمایش می‌دهد که مطابق آن، با افزایش افزونگی نشان‌گذاری جریان‌ها، نرخ کدگشایی افزایش نسبی دارد و به همین نسبت نرخ‌های مثبت اشتباهی و منفی اشتباهی نیز کاهش نسبی دارد و در کل کارایی روش پیشنهادی در شبکه مسیریابی پیازی مناسب بهره‌برداری می‌باشد.

۷-۱- مقایسه کارایی روش اصلی و روش پیشنهادی

برای آن‌که میزان بهبود روش پیشنهادی مشاهده شود، آن‌را با روش اصلی یا همان روش مبتنی بر فاصله در یک محیط آزمایشی مساوی آزمایش و مقایسه می‌کنیم. در آزمون ارزیابی نرخ کدگشایی روش پیشنهادی و روش اصلی مبتنی بر فاصله در شبکه مسیریابی پیازی، متغیر T برابر ۹۰۰ میلی‌ثانیه، تعداد بیت نشان‌گذاری ۱۵ بیت، حد آستانه ۵، ضریب محافظ برای روش پیشنهادی ۴ و مقدار آفست ۱۰۰ میلی‌ثانیه تنظیم شده است و افزونگی نشان‌گذاری در مقادیر مختلف ۳، ۵، ۷ و ۹ تغییر می‌کند. هر آزمون برای ۲۰ جریان مختلف تکرار گردید و بر اساس مقایسه درستی کدگشایی با حالت جریان واقعی نرخ کدگشایی استخراج گردید.

شکل (۱۰)، نمودار مقایسه نرخ کدگشایی روش پیشنهادی و روش مبتنی بر فاصله را نمایش می‌دهد. همان‌طورکه ملاحظه می‌شود با افزایش افزونگی روند نرخ کدگشایی روش مبتنی بر فاصله نامنظم است ولی در کل تفاوت قابل ملاحظه‌ی با روند نرخ کدگشایی روش پیشنهادی دارد.



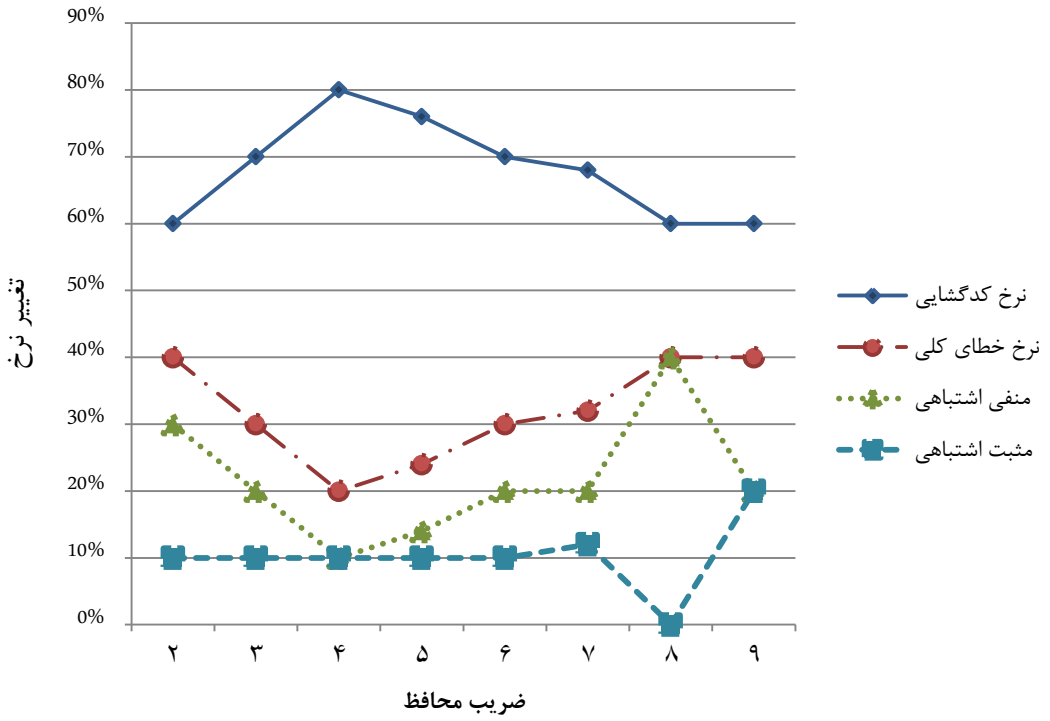
شکل (۱۰): نمودار مقایسه استحکام روش پیشنهادی و روش مبتنی بر فاصله

تفاوت نرخ کدگشایی دو روش با مقادیر مختلف افزونگی از ۲۰٪ تا ۴۰٪ می‌باشد. لذا با توجه به این آزمایش می‌توان بهبودی نرخ کدگشایی روش پیشنهادی را در شبکه مسیریابی پیازی اثبات کرد.

۷-۲- کارایی بر اساس میزان حاشیه امن

اساس تغییرات در روش پیشنهادی ایجاد حاشیه امن در هر فاصله برای قرارگیری بسته‌ها می‌باشد. برای ایجاد و اندازه حاشیه امن در فاصله‌ها یک ضریب با نام ضریب محافظ G درنظر گرفته شد که مقدار و میزان حاشیه امن براساس این ضریب محاسبه می‌شود. سؤالی که مطرح می‌شود آن است که چه مقداری برای این ضریب تنظیم گردد که نرخ کدگشایی به بیشنه حالت خود برسد. برای محاسبه بهترین ضریب ممکن در شبکه مسیریابی پیازی روش پیشنهادی را به حالت بهینه برساند، آزمونی طراحی شد که در آن برای هر ضریب محافظ حدود ۱۰ الی ۲۰ جریان تولید شدند و این جریان‌ها به‌صورت تصادفی با مشخصاتی یکسان نشان‌گذاری شده و از طریق شبکه مسیریابی پیازی به سمت کدگشا ارسال شدند. در قسمت کدگشا درصد درستی کدگشایی از این جریان‌ها به‌عنوان نرخ کدگشایی معرفی شد و براساس خطاها و کدگشایی اشتباه نیز نرخ‌های مثبت اشتباهی و منفی اشتباهی محاسبه شدند. در این آزمون، متغیر T برابر ۹۰۰ میلی‌ثانیه، تعداد بیت نشان‌گذاری ۱۵ بیت، حد آستانه ۵، تعداد افزونگی برای نشان‌گذاری ۳ و مقدار آفست ۱۰۰ میلی‌ثانیه

تنظیم شده است و ضریب محافظ نشان‌گذاری از ۲ الی ۹ تغییر می‌کند. شکل (۱۱)، نمودار کارایی روش پیشنهادی را براساس تغییرات حاشیه امن با استفاده از ضریب محافظ را نمایش می‌دهد. نکته‌ای که باید در این آزمایش و آزمایش‌های بعدی به آن اشاره کرد، این است که بالابردن تعداد افزونگی، حجم ارتباطات و زمان اجرای آزمایش را بالا می‌برد و از طرفی تغییرهای ناگهانی مسیر در شبکه مسیریابی پیازی با توجه به مفروضات و شرایط آزمایش، منجر به خرابی در آزمایش می‌شود، در این آزمایش‌ها از افزونگی ۳ با نرخ کارایی متوسط استفاده می‌شود که تعداد بسته‌های کم‌تر با زمان کم‌تری برای نشان‌گذاری احتیاج دارد که این خود باعث بالابردن دقت آزمایش‌ها می‌شود.



شکل (۱۱): نمودار کارایی روش پیشنهادی براساس تغییرات حاشیه امن.

بعد از انجام این آزمایش و محاسبه نتایج آن و با ملاحظه نمودار مشاهده می‌شود که نرخ کدگشایی در شرایط و متغیرهای مساوی در ضریب محافظ چهار به بهینه‌ترین حالت خود رسیده است و نمودار بعد و قبل از این نقطه نرخ‌های کدگشایی پایین‌تری را نمایش می‌دهد. لذا بهترین ضریب محافظ برای ایجاد حاشیه امن در روش پیشنهادی نشان‌گذاری در شبکه گمنامی ضریب چهار می‌باشد و این بدین معنی است که به میزان یک چهارم انتهایی هر فاصله زمانی نشان‌گذاری در نشان‌گذار خالی از هر بسته می‌باشد که به این صورت در انتهای هر فاصله، حاشیه امنی ایجاد می‌شود.

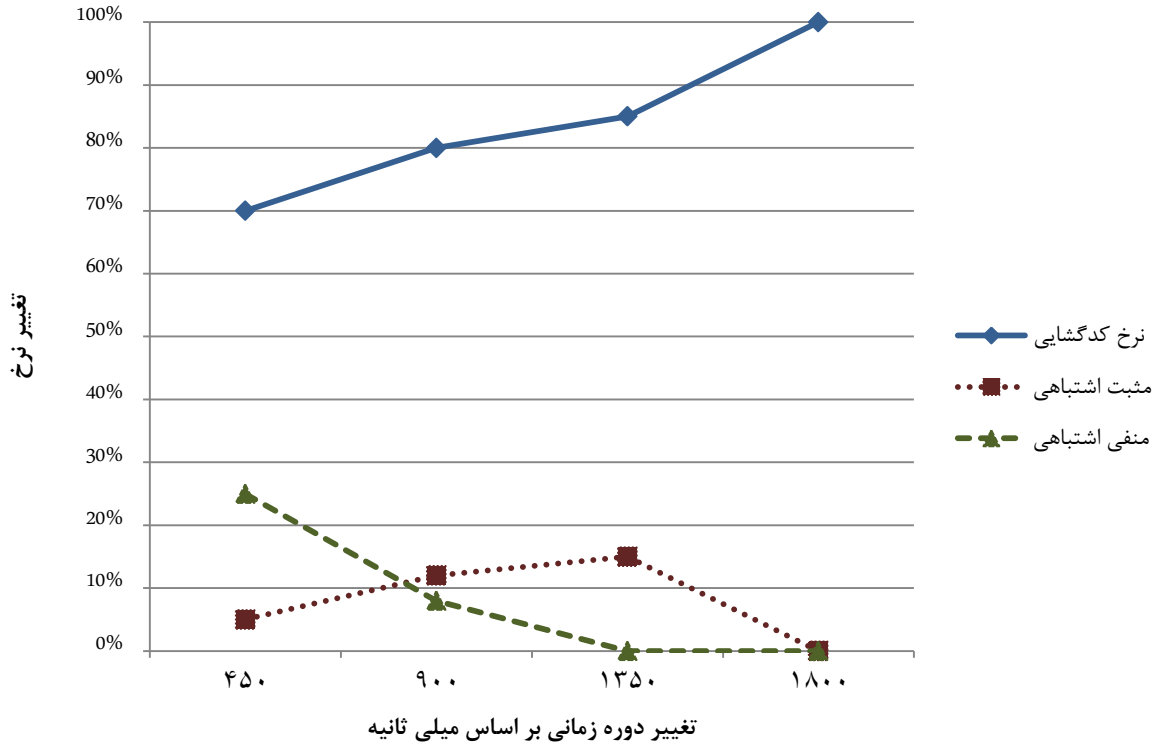
از طرفی دیگر نرخ‌های مثبت اشتباهی و منفی اشتباهی در ضریب چهار به کم‌ترین حالت خود می‌رسند ولی در ادامه در ضریب هشت دچار تغییراتی می‌شوند که روند شیب نمودار را تغییر می‌دهند این تغییر در آزمایش استحکام ما اهمیت خاصی ندارد. زیرا در نمودار، خطای کلی دارای شیب یکسان و بدون تغییر در ضریب هشت است و همچنین در آزمایش انجام‌شده مثبت اشتباهی به‌علت نمایش ترافیک سالم به‌جای ترافیک نشان‌گذاری‌شده دارای اهمیت بالایی نیست به این علت که اگر سامانه ردیابی یک شبکه حساس، ترافیکی را به اشتباه نشان‌گذاری‌شده تشخیص دهد می‌توان گفت که این خطا قابل چشم‌پوشی است ولی اگر همین سامانه ترافیک نشان‌گذاری را به اشتباه سالم تشخیص دهد (منفی اشتباهی) دیگر ردیابی آن ترافیک قابل انجام نخواهد بود. در نمودار شکل (۱۱) نرخ منفی اشتباهی روند یکسان دارد و در ضریب هشت به صفر می‌رسد و در این آزمون بنا به‌دلایل عنوان‌شده و همچنین به‌علت توجه و تمرکز به نرخ کدگشایی، تغییر نرخ مثبت اشتباهی را بی‌اهمیت فرض می‌کنیم.

۷-۳- کارایی براساس تغییرات فاصله زمانی

در آزمون بعدی نرخ کدگشایی روش پیشنهادی براساس تغییرات فاصله زمانی T آزمایش می‌شود، برای این آزمون نیز شبیه آزمون‌های قبل، به ازای تغییر هر فاصله زمانی به‌طور متوسط ۱۵ جریان مختلف ایجاد و نشان‌گذاری می‌کنیم و در شبکه گمنامی مسیریابی پیازی ارسال می‌کنیم و نرخ‌های کدگشایی و مثبت‌اشتباهی و منفی اشتباهی در قسمت کدگشا با توجه به نوع جریان ارسال شده، محاسبه می‌شوند.

در این آزمون، تعداد بیت نشان‌گذاری ۱۵ بیت، حد آستانه ۵، تعداد افزونگی برای نشان‌گذاری ۳، مقدار آفست ۱۰۰ میلی‌ثانیه و ضریب محافظ نشان‌گذاری نیز در بهینه حالت خود یعنی ۴ تنظیم شده است و فاصله زمانی یا متغیر T در مقادیر مختلف ۴۵۰، ۹۰۰، ۱۳۵۰ و ۱۸۰۰ میلی‌ثانیه تغییر می‌کند. در این شرایط کارایی و توان نرخ کدگشایی روش پیشنهادی کنترل و بررسی می‌شود. شکل (۱۲) نمودار کارایی روش پیشنهادی را براساس تغییرات فاصله زمانی را نمایش می‌دهد.

نتایج حاصله از این آزمایش نشان داد که با افزایش فاصله زمانی نرخ کدگشایی به‌طور صعودی بالا می‌رود که به همین ترتیب، نرخ‌های مثبت اشتباهی و منفی اشتباهی نیز کاهش پیدا می‌کند. این روند افزایش را در شبکه گمنامی مسیریابی پیازی ادامه دادیم تا با تنظیم فاصله زمانی ۱۸۰۰ میلی‌ثانیه نرخ کدگشایی نشان‌گذاری ۱۰۰٪ تمامی جریان‌های ارسالی بود؛ زیرا با تنظیم این فاصله زمانی، لغزش زمانی و تأخیرهای موجود در شبکه مسیریابی پیازی باعث خرابی در الگوی نشان‌گذاری نخواهند شد.



شکل (۱۲): نمودار کارایی روش پیشنهادی براساس تغییرات فاصله زمانی.

همین‌طورکه در شکل (۱۲) نمایش داده شده است، نرخ مثبت اشتباهی در ابتدا به‌صورت صعودی بالا می‌رود و نرخ منفی اشتباهی نیز به‌صورت نزولی پایین می‌آید.

۸- ارزیابی نامحسوسی روش پیشنهادی

به‌منظور تشخیص و شناسایی نشان‌گذاری مبتنی بر فاصله، ما آزمون‌های آماری تشخیص را بر روی نمونه‌های ترافیک نشان‌گذاری شده و ترافیک سالم اجرا می‌کنیم. در حقیقت نامحسوسی روش توانایی پنهان‌بودن نشان‌گذاری از تشخیص و شناسایی در مسیر جریان می‌باشد؛ زیرا درصورت محسوس‌بودن نشان‌گذاری، شناسایی آن توسط نفوذگر حرفه قابل انجام خواهد بود لذا آن‌ها می‌توانند با حمله به نشان‌گذاری، آن‌را از بین ببرند. ولی در فرض تحقیق آمده است که نشان‌گذاری در بستر شبکه مسیریابی پیازی صورت می‌پذیرد و نفوذگران که از این بستر برای نفوذ و گم‌کردن ردپا استفاده می‌کنند، هیچ‌گونه دسترسی و

اطلاعی از مسیریاب‌های و ایستگاه‌های داخل شبکه مسیریابی پیازی ندارند. لذا توانمندی شناسایی و تشخیص نشان‌گذاری‌ها با فرض تحقیق برای نفوذگر میسر نخواهد بود. ولی می‌توان با ایجاد سناریوهای متفاوت، مدل‌هایی را برای شناسایی نشان‌گذاری توسط نفوذگران طرح‌ریزی کرد. در شکل (۱۳)، مکان‌های که یک نفوذگر می‌تواند با ایجاد دسترسی به تشخیص نشان‌گذاری و در پی آن حمله به نشان‌گذاری استفاده نماید، مشخص شده است. حتی نفوذگر می‌تواند دسترسی خود را بر روی مسیریاب‌های داخل شبکه مسیریابی پیازی ایجاد نماید و از این طریق حمله به نشان‌گذاری را اجرا کند. در حقیقت نفوذگر با داشتن یک پراکسی در بین مسیر حمله خود که محل آن در قبل از شبکه مسیریابی پیازی (فضای مشخص‌شده الف) یا بعد از شبکه مسیریابی پیازی (فضای مشخص‌شده ب) و یا در داخل شبکه مسیریابی پیازی است، می‌تواند سناریوی تشخیص و حمله به نشان‌گذاری را انجام دهد. او می‌تواند جهت تشخیص نشان‌گذاری از آزمون‌های آماری نامحسوسی استفاده نماید. در آزمون نامحسوسی فرض می‌کنیم که برای آزمایش در یکی از محیط‌های الف یا ب دسترسی وجود دارد.

برای محاسبه نامحسوسی، با توجه به رمزنگاری و بسته‌بندی مجدد موجود در شبکه گمنامی، نمرات آستانه به‌صورت ۱۰ امین و ۹۰ امین صدک، یعنی کم‌ترین و بیش‌ترین نمرات در آزمایش‌های مختلف بر روی نمونه‌های سالم به‌دست می‌آید.

برای تعیین نمرات آستانه در آزمون‌های آنتروپی شرطی تصحیح‌شده و کولموگروف- اسمیرنوف ما از صدک ۹۰ ام و برای آزمون‌های آنتروپی تصحیح‌شده از صدک ۱۰ ام نمرات به‌دست‌آمده از نمونه‌های ترافیک سالم، استفاده می‌کنیم.

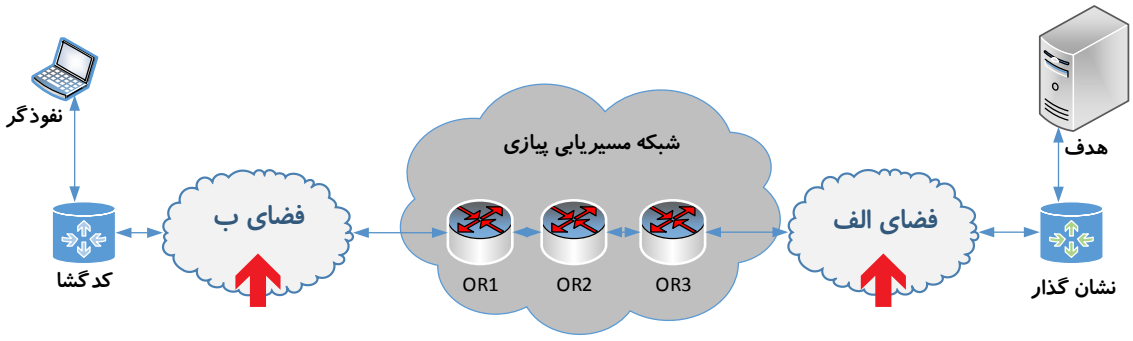
۸-۱- محاسبه نمرات آستانه

در آزمون کولموگروف- اسمیرنوف درصورتی‌که نمره آزمون کم‌تر از آستانه باشد، حاکی از آن است که نمونه نزدیک به رفتار عادی است. با این‌حال، اگر نمونه منطبق بر رفتار مناسب نباشد، نمره آزمون بزرگ‌تر از آستانه خواهد شد و احتمال وجود جریان نشان‌گذاری‌شده را نشان می‌دهد [۲۳].

در آزمون آنتروپی تصحیح‌شده اگر نمره آزمون کم‌تر از نمره آستانه باشد، بیانگر این است که نمونه منطبق بر توزیع مناسب نبوده و به احتمال زیاد از نوع ترافیک نشان‌گذاری شده است. در آزمون آنتروپی شرطی تصحیح‌شده اگر نمره آزمون بالاتر یا خیلی پایین‌تر از نمرات آستانه باشد، نشان‌دهنده احتمال وجود جریان نشان‌گذاری شده است.

هنگامی‌که نمره آزمون آنتروپی شرطی تصحیح‌شده بسیار پایین باشد، نمونه بسیار قاعده‌مند بوده و وقتی‌که نمره آزمون آنتروپی شرطی تصحیح‌شده بالاتر از نمره آستانه و یا خیلی نزدیک به آنتروپی مرتبه اول باشد، نمونه عدم همبستگی را نشان می‌دهد [۲۳]. نمرات آستانه برای آزمون‌های آنتروپی تصحیح‌شده، آنتروپی شرطی تصحیح‌شده و کولموگروف- اسمیرنوف در جدول (۲) نشان داده شده است.

جدول(۳) مقادیر به‌دست‌آمده از آزمون‌ها را بر روی فاصله زمانی‌های مختلف را نشان می‌دهد و جدول(۴) خلاصه‌ای از آزمون‌ها را به همراه نتایج نهایی آزمون نمایش می‌دهد.



شکل (۱۳): مدل تهدید فرضی برای نشان‌گذاری در شبکه مسیریابی پیازی.

جدول (۲): نمرات آستانه برای آزمون‌های مختلف نامحسوسی

نوع آزمون	نمرات آستانه	نرخ مثبت اشتباهی
کولموگروف- اسمیرنوف (KS)	KS ≤ ۰/۲۸۱۸	%۱۰
آنتروپی تصحیح شده (CEN)	۲۱/۶۳۸۱ ≤ CEN	%۱۰
آنتروپی شرطی تصحیح شده (CCE)	CCE ≤ ۲/۰۸۸۴	%۱۰

جدول (۳): مقادیر به‌دست‌آمده از آزمون K-S براساس فاصله زمانی.

فاصله‌زمانی (میلی‌ثانیه)	آزمون K-S			نتیجه آزمون
	میانگین	انحراف معیار	صدک ۹۰ ام	
۴۵۰	۰/۲۰۶۷	۰/۰۶۵۱	۰/۲۵۲۷	نامحسوس
۹۰۰	۰/۲۰۱۵	۰/۰۳۷۷	۰/۲۳۸	نامحسوس
۱۳۵۰	۰/۲۳۷	۰/۰۳۹۸	۰/۲۷۸	نامحسوس
۱۸۰۰	۰/۱۷۵	۰/۰۵۲۵	۰/۲۳۳۶	نامحسوس

جدول (۴): بررسی نتایج آزمون آنتروپی و آنتروپی شرطی.

نوع آزمون	فاصله زمانی میلی‌ثانیه	نتیجه مقایسه با حد آستانه	میانگین مقادیر آزمون‌ها	نتیجه آزمون
آنتروپی	۴۵۰	کوچک‌تر	۱۵/۶۶۰۸۲۱	محسوس
آنتروپی	۹۰۰	کوچک‌تر	۱۵/۸۱۹۹۶۹	محسوس
آنتروپی	۱۳۵۰	کوچک‌تر	۱۵/۸۱۹۷۷۱	محسوس
آنتروپی	۱۸۰۰	کوچک‌تر	۱۶/۴۰۲۷۶۷	محسوس
آنتروپی شرطی	۴۵۰	کوچک‌تر	۱/۹۰۸۰۲۳	نامحسوس
آنتروپی شرطی	۹۰۰	کوچک‌تر	۱/۹۴۸۳۴۸	نامحسوس
آنتروپی شرطی	۱۳۵۰	کوچک‌تر	۱/۹۵۱۵۴۲۵	نامحسوس
آنتروپی شرطی	۱۸۰۰	کوچک‌تر	۱/۸۶۲۹۲۲۸	نامحسوس

هنگامی‌که نمره آزمون آنتروپی شرطی تصحیح‌شده بسیار پایین باشد، نمونه بسیار قاعده‌مند بوده و وقتی‌که نمره آزمون آنتروپی شرطی تصحیح‌شده بالاتر از نمره آستانه و یا خیلی نزدیک به آنتروپی مرتبه اول باشد، نمونه عدم همبستگی را نشان می‌دهد [۲۳]. نمرات آستانه برای آزمون‌های آنتروپی تصحیح‌شده، آنتروپی شرطی تصحیح‌شده و کولموگروف- اسمیرنوف در جدول (۲) نشان داده شده است.

برای محاسبه حد آستانه در هر آزمون از دو نمونه سالم که به توزیع یکسان تعلق دارند، استفاده می‌شود. هر نمونه شامل تأخیرهای بین بسته‌ای جریان‌های سالم می‌باشند. برای محاسبه حد آستانه در آزمون آنتروپی تصحیح‌شده، یک نمونه سالم به‌عنوان مجموعه آموزشی با تعداد ۲۹ هزار عدد تأخیر بین بسته‌ای انتخاب شد و با نمونه سالمی دیگر با تعداد ۱۰ هزار عدد تأخیر بین بسته‌ای به‌صورت ۲۰۰۰ تایی به برنامه محاسبه آنتروپی تصحیح‌شده داده شد. مقادیر به‌دست‌آمده با میانگین ۲۰/۶۶۳ و اختلاف معیار ۱/۰۶۳۴۶۷ می‌باشد که با احتساب خطای مثبت اشتباهی ۱۰٪ مقدار حد آستانه برابر ۲۱/۴۵۷۹۲۷ شد.

هم‌چنین برای محاسبه حد آستانه در آزمون آنتروپی شرطی تصحیح‌شده، یک نمونه سالم به‌عنوان مجموعه آموزشی با تعداد ۲۹ هزار عدد تأخیر بین بسته‌ای انتخاب شد و با همان نمونه سالم به‌صورت ۲۰۰۰ تایی به برنامه محاسبه آنتروپی شرطی تصحیح‌شده داده شد. مقادیر به‌دست‌آمده با میانگین ۱/۸۹۲۹۴۲۵ و اختلاف معیار ۰/۱۶۷۰۱۳ می‌باشد که با احتساب خطای مثبت اشتباهی ۱۰٪ مقدار حد آستانه برابر ۲/۰۸۸۴ به‌دست آمد. در محاسبه حد آستانه آزمون K-S نیز یک نمونه سالم ۲۹ هزارتایی یک بار به‌عنوان نمونه آزمایش و یک بار به‌عنوان نمونه آموزشی انتخاب شدند و با استفاده از تابع محاسبه K-S در نرم‌افزار Matlab مقدار آن چندین بار محاسبه شد که میانگین آن ۰/۲۱۱۵ شد و مقدار حد آستانه با احتساب خطای مثبت اشتباهی ۱۰٪، ۰/۲۸۱۸ محاسبه شد.

۸-۲- آزمون کولموگروف- اسمیرنوف

برای محاسبه K-S هم از نمونه تست و هم از نمونه سالم که به توزیع یکسان تعلق دارند، استفاده می‌شود که هر نمونه شامل تأخیرهای بین بسته‌ای جریان سالم و جریان تست می‌باشند. اگر حداکثر فاصله کم‌تر از حد آستانه باشد، نمونه مورد بررسی به مجموعه سالم تعلق خواهد داشت. با این‌حال، اگر حداکثر فاصله فراتر از حد آستانه باشد به یک توزیع متفاوت تعلق دارد و به احتمال زیاد ترافیک از نوع نشان‌گذاری خواهد بود. برای اجرای این آزمون نمونه سالم شامل ۱۰۰۰ عدد تأخیر بین بسته‌ای

ترافیک سالم انتخاب شد و با نمونه‌های جریان‌های نشان‌گذاری شده با فاصله زمانی‌های متفاوت و تعداد تأخیر بین بسته‌ای مساوی ترافیک سالم با هم در آزمون K-S آزمایش شدند.

مطابق آزمایش‌ها با آزمون K-S تمامی جریان‌ها با فاصله زمانی متفاوت دارای مقدار پایین‌تر از حد آستانه هستند، لذا با استفاده از آزمون K-S نمی‌توان نشان‌گذاری روش پیشنهادی را تشخیص داد و این روش با آزمون K-S نامحسوس خواهد بود.

۸-۳- آزمون‌های آنتروپی و آنتروپی شرطی

در تشخیص مبتنی بر آنتروپی سعی در اندازه‌گیری شباهت نسبی بین ترافیک مجاز و ترافیک مشکوک به نشان‌گذاری است. از آن‌جایی‌که در نشان‌گذاری اطلاعات با تغییر در زمان‌بندی تأخیرهای بین بسته‌ای کدگذاری می‌شود، توزیع زمان‌های بین بسته‌ها را تحت تأثیر قرار می‌دهد. برای مشاهده چنین تغییری در توزیع از میزان آنتروپی استفاده می‌شود. محاسبه آنتروپی براساس توزیع نمونه‌های تأخیرهای بین بسته‌ای ترافیک مجاز شناخته‌شده انجام می‌شود. هرگونه انحراف از این توزیع آموزشی باعث افت آنتروپی محاسبه‌شده می‌شود و می‌توان از آن به‌عنوان یک معیار برای تشخیص نشان‌گذاری استفاده کرد. به‌طورکلی برای تشخیص نشان‌گذاری از دو برنامه آنتروپی و آنتروپی شرطی که در [۲۳] استفاده شده است، استفاده می‌کنیم. برای محاسبه مقدارهای هر دو نوع آزمون آنتروپی هم از نمونه تست و هم از نمونه سالم که به توزیع یکسان تعلق دارند، استفاده می‌شود که هر نمونه شامل تأخیرهای بین بسته‌ی جریان سالم و جریان تست می‌باشند. نمونه سالم که همان نمونه آموزشی ۲۹ هزارتایی می‌باشد با نمونه‌های تستی مختلف که با فاصله‌های زمانی مختلف نشان‌گذاری شده‌اند، در آزمون بررسی می‌شوند.

نتیجه‌ای که بعد از انجام آزمون‌ها به‌دست آمد، آن بود که آزمون آنتروپی شرطی در هیچ یک از انواع نشان‌گذاری روش پیشنهادی، توانایی تشخیص را ندارد ولی آزمون آنتروپی تصحیح‌شده تمامی جریان‌های نشان‌گذاری را تشخیص داده است. لذا به‌طورکلی می‌توان گفت که با فرض سناریوی در شکل (۱۳)، روش پیشنهادی دارای نامحسوسی لازم در برابر آزمون آنتروپی تصحیح‌شده نیست ولی با فرض تحقیق روش پیشنهادی قابلیت تشخیص نخواهد داشت.

آزمون آنتروپی شرطی از دسته آزمون‌های قاعده‌مندی است. به نظر می‌رسد عدم تشخیص نشان‌گذاری توسط آزمون آنتروپی شرطی به این جهت باشد که عملاً دست‌کاری تأخیرهای بین بسته‌ی در منحنی توزیع آن‌ها قاعده‌مندی ایجاد نمی‌کند و همان توزیع از یک T به T بعدی شیفت داده می‌شود. به نظر می‌رسد یکی دیگر از دلایل عدم تشخیص، تعداد کم بسته‌ها در هر پنجره باشد. از این‌رو، جابه‌جایی دو یا سه بسته موجود در یک پنجره روی قاعده‌مندی تأثیر زیادی ندارد و شاید هم تشخیص آزمون KS نیز به‌همین دلیل باشد. ولی اگر تابع تولید بسته‌ها تنظیم شود که تعداد زیادی بسته در یک پنجره (3T) وجود داشته باشد؛ آن‌گاه دست‌کاری آن‌ها هم روی شکل و هم روی قاعده‌مندی تأثیر قابل مشاهده و محسوسی خواهد داشت؛ که این موضوع می‌تواند در تحقیقات آتی و پیشنهادی لحاظ گردد.

۹- نتیجه‌گیری

در روش مبتنی بر فاصله جهت نشان‌گذاری ترافیک شبکه، جابه‌جایی بسته‌های انتهایی هر فاصله به فاصله بعد، به‌خصوص بعد از گذشت از شبکه گمنامی مسیریابی پیازی، سبب کاهش استحکام نشان‌گذاری می‌شود. جهت حل این مسأله، یک ضریب محافظ برای خالی‌کردن مرزهای انتهای هر فاصله درنظر گرفته شد. براساس مقدار این ضریب بخشی از هر فاصله از یک پنجره به‌عنوان محافظ، خالی از بسته نگه داشته می‌شود. برای ارزیابی میزان ارتقای استحکام روش پیشنهادی، روش قبلی و پیشنهادی در شرایط یکسان آزموده شدند که با مقایسه نتایج، نرخ کدگشایی روش پیشنهادی بهبود قابل توجهی را نشان می‌دهد. در گام بعد به‌منظور استخراج مقدار بهینه ضریب محافظ جهت کنترل میزان حاشیه امن در فاصله‌ها، آزمون دیگری انجام شد که بهترین ضریب مقدار چهار به‌دست آمد. هم‌چنین استحکام روش پیشنهادی براساس تغییرات فاصله زمانی بررسی شد، نتیجه این آزمایش نشان می‌دهد که با افزایش فاصله زمانی استحکام نشان‌گذاری افزایش می‌یابد.

در آخرین آزمایش نامحسوسی روش پیشنهادی با استفاده از روش‌های آماری مرتبه دوم مانند K-S، آنتروپی و آنتروپی شرطی مورد بررسی قرار گرفت که روش پیشنهادی توسط روش آنتروپی کاملاً محسوس بود ولی با روش‌های آنتروپی شرطی و K-S نامحسوس بود. بدین‌صورت روش پیشنهادی با یک روش از سه روش تشخیص نشان‌گذاری ذکرشده قابل تشخیص است.

لذا جهت کار در آینده می‌توان با استفاده از الگوریتم‌های نامحسوس کانال‌های پوششی زمان‌بندی‌دار، به روش‌های نامحسوس جدید نشان‌گذاری دست یافت. هم‌چنین برای ادامه تحقیق می‌توان بر روی نشان‌گذاری با ترافیک‌های کاربردی متفاوت مانند HTTP، SSH، RDP و غیره پژوهش را ادامه داد. ارائه روش‌های حمله و فریب در ردیابی توسط نشان‌گذاری جریان شبکه به‌منظور پنهان‌کردن مبدأ نفوذ را می‌توان از دیگر پیشنهادات پژوهشی آینده معرفی کرد.

۱۰- مراجع

[1] A. Houmansadr, T. Coleman, N. Kiyavash, and N. Borisov, "On the channel capacity of network flow watermarking," 2009.

[2] A. Houmansadr and N. Borisov, "SWIRL: A Scalable Watermark to Detect Correlated Network Flows," in Network and Distributed System Security Symposium. Internet Society, Feb 2011.

[3] A. Houmansadr, "Design, Analysis, And implementation of effective network flow watermarking schemes," Doctor of Philosophy in Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, University of Illinois, Urbana-Champaign, 2012.

[4] A. Ahmadi, "Intruder Tracing in Anonymous Networks Using the Network Flow Watermarking," M.S. Thesis, I.T.C. Department, Imam Hossein Comprehensive University, Tehran, 2014, (in Persian).

[5] A. Houmansadr and N. Borisov, "Towards Improving Network Flow Watermarks using the Repeat-accumulate Codes," in 36th International Conference on Acoustics, Speech and Signal Processing, 2011.

[6] A. Houmansadr, N. Kiyavash, and N. Borisov, "Rainbow A Robust and Invisible Non-Blind Watermark for Network flow," in Network and Distributed System Security Symposium, Feb. 2009.

[7] X. Wang and D. S. Reeves, "Robust correlation of encrypted attack traffic through stepping stones by anipulation of interpacket delays," in ACM Conference on Computer and Communications Security, New York, NY, USA, pp. 20-29, 2003.

[8] J. Luo, X. Wang, and M. Yang, "An interval centroid based spread spectrum watermarking scheme for multi-flow traceback," Journal of Network and Computer Applications Elsevier, 2011.

[9] J. Huang, X. Pan, X. Fu, and J. Wang, "Long PN Code Based DSSS Watermarking," presented at the Annual IEEE International Conference on Computer Communications (INFOCOM), 2011.

[10] X.Wang, J. Luo, and M. Yang, "A Double Interval Centroid-Based Watermark for Network Flow Traceback," in 14th International Conference on Computer Supported Cooperative Work in Design. (CSCWD'2010), Shanghai, China, pp. 146–151, April 2010.

[11] L. Zhang, Z. Wang, Q. Wang, and F. Miao, "MSAC and Multi-flow Attacks Resistant Spread Spectrum Watermarks for network flows," in Information and Financial Engineering (ICIFE), 2010 2nd IEEE International Conference on, pp. 438-441, 2010.

[12] L. Zhang, J. Luo, and M. Yang, "An Improved DSSS-Based Flow Marking Technique for Anonymous Communication Traceback," in Multidisciplinary Autonomous Networks and Systems (MANS 09), Brisbane, Australia, pp. 563 - 567, May 2009.

[13] A. Houmansadr, N. Kiyavash, and N. Borisov, "Multi-flow attack resistant watermarks for network flows," in IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1497-1500, 2009.

[14] Y. J. Pyun, Y. H. Park, X. Wangy, D. S. Reeves, and P. Ning, "Tracing Traffic through Intermediate Hosts that Repacketize Flows," IEEE Conference on Computer Communications (IN-FOCOM), pp. 634–642, May 2007.

[15] X. Wang, S. Chen, and S. Jajodia, "Network Flow Watermarking Attack on Low-Latency Anonymous Communication Systems," IEEE Symposiumon Security and Privacy, pp. 116–130, 2007.

[16] W. Yu, XinwenFu, S. Graham, D. Xuan, and W. Zhao, "DSSS-Based Flow Marking Technique for Invisible Traceback," presented at the EEE Security and Privacy Symposium (S&P), 2007.

[17] A. Ahmadi, M. Dehghani, and M. S. Esfehani, "Survey of Intruder Tracing Methods in Anonymous Networks Using The Network Flow Watermarking," Passive Defence Quarterly, vol. 6, pp. 27-36, Summer 2015. (in Persian)

[18] N. Kiyavash, A. Houmansadr, and N. Borisov, "Multi-flow attacks against network flow watermarking schemes," in USENIX Security Symposium, Berkeley, CA, USA:USENIX Association, 2008.

[19] Y. J. Pyun, Y. Park, D. S. Reeves, X. Wang, and P. Ning, "Interval-based flow watermarking for tracing interactive traffic," Computer Networks 56, Elsevier, pp. 1646-1665, 2012.

[20] X. Wang, D. S. Reeves, P. Ning, and F. Feng, "Robust Network-Based Attack Attribution through Probabilistic Watermarking of Packet Flows," Department of Computer Science, NC State University, 2005.

[21] R. Duda, P. Hart, and D. Stork, "Pattern Classification(2ndEdition)," John Wiley & Sons, 2001.

[22] P. Peng, P. Ning, and D. S. Reeves, "On the secrecy of timing-based active watermarking trace-back techniques," in IEEE Symposium on Security and Privacy, pp. 334-349, May 2006.

[23] B. Birami, "Covert Timing Channel Detection With Use Statistic Methods," M.S. Thesis, I.T.C. College, Imam Hossein Comprehensive University, Tehran, 2012. (in Persian).

[24] N. Kiyavash, A. Houmansadr, and N. Borisov, "Multi-flow attacks against network flow watermarks analysis and countermeasures," arXiv preprint arXiv:1203.1390, 2012.

[25] L. Zhang, Z. Wang, J. Xu, and Q. Wang, "Multi-flow Attack Resistant Interval-Based Watermarks for Tracing Multiple Network Flows," in Computing and Intelligent Systems, ed: Springer, pp. 166-173, 2011.

Journal Of Electronical & Cyber Defence
Vol. 5, No. 2, 2017, Serial No. 18

A Problem Solving Method to Boundary of Interval Based Watermark in Anonymous Network Flows

A. Ahmadi, M. Dehghani*, M. Saleh Esfehani

*Imam Hossein University

(Received: 30/11/2015, Accepted: 03/01/2017)

ABSTRACT

Ability of intruder tracing, is deterrence to insecurity. One of the methods of intruder tracing is network flow watermark. In this technique, the pattern of network flow is changed, to watermark the special flow of traffic and we can be tracing it in the output boundaries of the network. In this research interval based watermark (IBW) method that ever on TOR anonymous network has not been evaluated, is implement to be practical in the real environment was evaluated on TOR. The analysis results show that this method has the weakness of the border. The proposed method to improve IBW method, with creating a blank space as guard at boundary intervals was used to solve problem in boundary. After implement and evaluate the proposed method, Measurement accuracy based on decoding rates, false positive and false negative. Show proposed method has better performance compared to old method. Also for evaluate stealthness with assume intruder scenario, was used statistic methods: K-S test, entropy test and conditional entropy test to detect watermark. The results show that proposed method in the K-S test and entropy test has acceptable stealthness level.

Keywords: Watermark, Interval Based Watermark, Anonymous Network

* Corresponding Author Email: mdehghany@ihu.ac.ir