هفتمین کنفرانس بین المللی انجمن رمز ایران

۲۴ و ۲۵ شهریورماه ۱۳۸۹، دانشگاه صنعتی خواجه نصیرالدین طوسی

!

انجمن رمزایران

دانشگاه صنعتی خواجه نصیرالدین طوسی

روشی مبتنی بر خوشه‌بندی برای تشخیص ناهنجاری پویا در شبکه‌های اقتضایی متحرک با پروتکل مسیریابی AODV

میثم علیخانی، مهدی آبادی

تهران، دانشگاه تربیت مدرس، دانشکده مهندسی برق و کامپیوتر، گروه مهندسی کامپیوتر

m.alikhany@modares.ac.ir, abadi@modares.ac.ir

چکیده

برای تشخیص ناهنجاری در شبکه‌های اقتضایی متحرک روش‌های مختلفی پیشنهاد شده است. در روش‌های تشخیص ناهنجاری ایستا، یک نما از رفتار عادی شبکه ایجاد شده و از آن در مرحله تشخیص استفاده می‌شود. با توجه به همبندی پویا در شبکه‌های اقتضایی متحرک، استفاده از یک نمای عادی از پیش تعریف شده نمی‌تواند رفتار شبکه را به خوبی توصیف کند. در این مقاله، روشی پویا برای تشخیص ناهنجاری در شبکه‌های اقتضایی متحرک با پروتکل مسیریابی AODV پیشنهاد می‌شود. در روش پیشنهادی، هر گره در هر پنجره زمانی مجموعه‌ای از بردارهای داده را استخراج می‌کند. به مجموعه بردارهای داده استخراج شده با توجه به نرخ تغییر همبندی شبکه و زمان استخراج وزن متفاوتی نسبت داده می‌شود. روش پیشنهادی از دو مرحله آموزش و تشخیص تشکیل می‌شود. در مرحله آموزش، الگوریتم خوشه‌بندی وزن‌دار WFWC بر روی مجموعه بردارهای داده اعمال شده و نمای عادی ایجاد می‌شود. در مرحله تشخیص، در پایان هر پنجره زمانی ابتدا بردارهای داده عادی به نمای ایجاد شده اضافه می‌شوند. سپس، با به‌روزرسانی وزن بردارهای داده و اعمال الگوریتم WFWC نمای عادی به‌روزرسانی می‌شود. نتایج آزمایش‌های انجام شده نشان می‌دهد که به‌روزرسانی نمای عادی باعث افزایش نرخ تشخیص و کاهش نرخ هشدار نادرست روش پیشنهادی می‌شود.

واژه های¡ کلیدی

شبکه اقتضایی متحرک، پروتکل AODV، تشخیص نفوذ، تشخیص ناهنجاری پویا، حمله سیاه‌چاله، خوشه‌بندی وزن‌دار شخیص.

۱- مقدمه !

هر شبکه اقتضایی متحرک (MANET[1]) شامل مجموعه‌ای از گره‌های بی‌سیم و متحرک است که در آن هیچ‌گونه زیرساخت یا نقطه دسترسی[2] متمرکزی از قبیل ایستگاه‌های پایه[3] وجود ندارد. امروزه از شبکه‌های اقتضایی متحرک به دلیل توسعه سریع و آسان در کاربردهای مختلفی از قبیل کنفرانس‌های ویدئویی، حوادث غیرمنتظره (سیل و زلزله) و محیط‌های نظامی و صنعتی استفاده می‌شود [۱]. شبکه‌های اقتضایی متحرک به دلیل داشتن ویژگی‌هایی ذاتی از قبیل همبندی[4] پویا و عدم وجود مدیریت امنیتی متمرکز، در مقایسه با سایر شبکه‌ها از آسیب‌پذیری بیشتری برخوردار هستند.

روش‌های تشخیص نفوذ به دو دسته کلی تشخیص مبتنی بر امضاء[5] و تشخیص ناهنجاری[6] تقسیم می‌شوند [۲]. در روش‌های تشخیص مبتنی بر امضاء، الگوهای نفوذهای شناخته شده با ترافیک ورودی مقایسه شده و در صورت تطبیق نفوذ تشخیص

[1] Mobile Ad hoc NETwork

[2] Access point

[3] Base stations

[4] Topology

[5] Signature-based detection

[6] Anomaly detection

داده می‌شود. مزیت این روش‌ها نرخ هشدار نادرست[1] پایین و عیب آن‌ها عدم تشخیص نفوذهای جدید است. در روش‌های تشخیص ناهنجاری، ابتدا یک نما[2] از رفتار عادی شبکه ایجاد می‌شود. سپس، هر ترافیکی که از نمای ایجاد شده انحراف داشته باشد به عنوان نفوذ تشخیص داده می‌شود. مزیت این روش‌ها تشخیص نفوذهای جدید و عیب آن‌ها نرخ هشدار نادرست بالا است. روش‌های تشخیص ناهنجاری به دو دسته شبه‌نظارتی[3] و بدون‌نظارت[4] تقسیم می‌شوند [۳]. روش‌های تشخیص ناهنجاری شبه‌نظارتی در مرحله آموزش، نیازمند مجموعه‌ای از داده‌های عادی برای ایجاد نما هستند. در حالی که روش‌های تشخیص ناهنجاری بدون‌نظارت در مرحله آموزش نیازی به داده‌های عادی ندارند. در این روش‌های تشخیص ناهنجاری فرض می‌شود که فراوانی داده‌های عادی از داده‌های ناهنجار خیلی بیشتر است و داده‌های عادی از لحاظ آماری متمایز از داده‌های ناهنجار هستند.

در این مقاله، یک روش تشخیص ناهنجاری بدون‌نظارت پیشنهاد شده است. روش پیشنهادی، از دو مرحله آموزش و تشخیص تشکیل می‌شود. در مرحله آموزش، با استفاده از یک الگوریتم خوشه‌بندی وزن‌دار با شعاع ثابت نمای عادی ایجاد می‌شود. در مرحله تشخیص، با استفاده از نمای عادی بردارهای داده ناهنجار تشخیص داده می‌شوند. سپس، با به‌روزرسانی وزن بردارهای داده عادی و اعمال الگوریتم خوشه‌بندی وزن‌دار با شعاع ثابت نمای عادی به‌روزرسانی می‌شود.

در ادامه، در بخش ۲ به کارهای مرتبط اشاره می‌شود. در بخش ۳ پروتکل مسیریابی AODV و در بخش ۴ انواع حملات در برابر این پروتکل مسیریابی به صورت اجمالی شرح داده می‌شوند. در بخش ۵ روشی برای تشخیص ناهنجاری پویا در شبکه‌های اقتضایی متحرک پیشنهاد می‌شود. در بخش ۶ نتایج آزمایش‌های انجام شده برای ارزیابی روش پیشنهادی ارائه می‌شود و در نهایت در بخش ۷ نتیجه‌گیری به عمل می‌آید.

۲- کارهای مرتبط

برای تشخیص نفوذ به شبکه‌های اقتضایی متحرک روش‌های مختلفی پیشنهاد شده است که در ادامه به برخی از آن‌ها اشاره می‌شود.

لی و همکارانش [۴] روشی پیشنهاد کرده‌اند که از بسته‌های اضافی CREQ و CREP در فرآیند مسیریابی استفاده می‌کند. در این روش هر گره میانی که مسیری به گره مقصد دارد، علاوه بر ارسال بسته RREP به سمت گره مبدا یک بسته CREQ به گره بعدی خود در سمت گره مقصد ارسال می‌کند. گره بعدی با دریافت بسته CREQ و با جستجو در حافظه نهان[5] خود وجود یک مسیر به گره مقصد را بررسی کرده و در صورت یافتن مسیر با بسته CREP به گره مبدا پاسخ می‌دهد. گره مبدا با مقایسه مسیرهای موجود در بسته‌های RREP و CREP از صحت مسیر در بسته RREP دریافتی اطمینان حاصل می‌کند. در صورتی که بسته RREP توسط گره مقصد ارسال شده باشد، نیازی به ارسال بسته‌های CREQ و CREP نمی‌باشد. در این روش، عملیات اضافی ارسال بسته‌های CREQ و CREP باعث به وجود آمدن سربار[6] اضافی در فرآیند مسیریابی می‌شود و در نتیجه کارآیی شبکه را با مصرف پهنای باند کاهش می‌دهد.

هوآنگ و همکارانش [۵] روشی پیشنهاد کرده‌اند که در آن هر گره بر جریان بسته‌ها نظارت می‌کند. در این روش، ویژگی‌های متعددی وابسته به ترافیک و همبندی شبکه تعریف شده و با استفاده از وابستگی بین این ویژگی‌ها ناهنجاری تشخیص داده می‌شود. هوآنگ و همکارانش [۶] رفتار عادی پروتکل مسیریابی AODV را با استفاده از یک آتوماتای حالت متناهی توسعه‌یافته (EFSA۷) مدل کرده‌اند و برای تشخیص حملات از هر دو روش تشخیص ناهنجاری و تشخیص مبتنی بر توصیف استفاده کرده‌اند.

سان و همکارانش [۷] یک روش تشخیص ناهنجاری با توجه به وجود تحرک در شبکه‌های اقتضایی متحرک پیشنهاد کرده‌اند. در این روش، ابتدا در مرحله آموزش با استفاده از مدل‌های مختلف تحرک فعالیت‌های مسیریابی جمع‌آوری شده و برای هر سطح تحرک متوسط نرخ تغییر پیوند (LCR8) محاسبه می‌شود. از فعالیت‌های مسیریابی جمع‌آوری شده برای ایجاد نماهای عادی استفاده می‌شود. سپس در مرحله تشخیص، هر سیستم تشخیص نفوذ محلی به طور متناوب نرخ تغییر پیوند (LCRrecent) را برای فعالیت‌های مسیریابی اخیر گره خود محاسبه کرده و از بین نماهای عادی نمایی را انتخاب می‌کند که LCR آن کمترین فاصله اقلیدسی را با LCRrecent داشته باشد. در هر بازه زمانی، هر گره LCR را با توجه به همسایگان جدید و قدیمی خود محاسبه می‌کند. بنابراین، برای محاسبه LCR تعداد کل گره‌ها در شبکه در نظر گرفته نمی‌شود. اما باید توجه داشت که دلیل تغییر حالت شبکه ممکن است به واسطه ظاهرشدن و ناپدیدشدن ناگهانی

[1] False Alarm Rate (FAR)
[2] Profile
[3] Semi-supervised
[4] Unsupervised
[5] Cache
[6] Overhead
[7] Extended Finite State Automaton
[8] Link Change Rate

سایر گره‌ها در شبکه باشد. هنگامی که رفتار گره‌ها در مرحله تشخیص متفاوت از مرحله آموزش باشد، استفاده از یک نمای عادی از پیش تعریف شده نمی‌تواند رفتار شبکه را به خوبی توصیف کند.

کوروساوا و همکارانش [۸] روشی پویا برای تشخیص حمله سیاه‌چاله[1] پیشنهاد کرده‌اند. در این روش از تشخیص ناهنجاری شبه‌نظارتی استفاده می‌شود که در مرحله آموزش نیازمند یک مجموعه داده‌های فاقد حمله است.

۳- پروتکل مسیریابی AODV

پروتکل AODV [۹] یک پروتکل مسیریابی واکنشی[2] است. این پروتکل از مفهوم شماره توالی مقصد[3] در پروتکل مسیریابی DSDV برای نگهداری آخرین اطلاعات مسیریابی استفاده می‌کند. فرض کنید گره مبدا S قصد دارد تا ارتباطی را با گره مقصد D برقرار کند. در صورت نداشتن اطلاعات مسیریابی، گره S فرآیند کشف مسیر را با همه‌پخشی[4] یک بسته RREQ به گره‌های همسایه خود آغاز می‌کند. هر گره همسایه N با دریافت بسته RREQ و در صورت داشتن اطلاعات مسیریابی و تازه بودن این اطلاعات با یک بسته RREP به گره S پاسخ می‌دهد. در غیر این صورت، فیلد شمارنده گام[5] بسته RREQ را به اندازه یک واحد افزایش داده و این بسته را مجدداً به همسایه‌های خود همه‌پخشی می‌کند. همچنین، اطلاعات مسیریابی را برای ایجاد مسیر معکوس نگه‌داری می‌کند. گره N برای حصول اطمینان از تازه بودن اطلاعات مسیریابی، شماره توالی مقصد در بسته RREQ دریافتی را با شماره توالی گره D در جدول مسیریابی خود مقایسه می‌کند. در صورت کوچکتر بودن شماره توالی گره D در جدول مسیریابی، این شماره توالی را با شماره توالی مقصد در بسته RREQ به‌روزرسانی می‌کند. در صورتی که گره N چندین بسته RREP را دریافت کند، بسته‌ای را انتخاب می‌کند که دارای شماره توالی مقصد بزرگتری است. در صورت یکسان بودن شماره توالی مقصد بسته‌های RREP دریافتی، بسته‌ای را انتخاب می‌کند که دارای شمارنده گام کوچکتری است.

۴- انواع حملات در برابر پروتکل AODV

انواع حملات در برابر پروتکل AODV را می‌توان به دو دسته اختلال در مسیریابی و مصرف منابع تقسیم کرد [۱۰]. در حملات اختلال در مسیریابی، گره مهاجم از ویژگی‌های پروتکل AODV سوءاستفاده می‌کند. یکی از انواع حملات اختلال در فرآیند مسیریابی حمله سیاه‌چاله [۱۱] است. در این حمله، گره مهاجم با جعل فیلدهای شماره توالی مقصد، شمارنده گام و غیره در بسته‌های RREQ و RREP همه بسته‌های گره(های) قربانی را به سمت خود جذب کرده و از رسیدن آن‌ها به مقصد جلوگیری به عمل می‌آورد. در ادامه، دو نوع از این حمله معرفی می‌شود.

۴-۱. حمله سیاه‌چاله با بسته‌های RREP جعلی

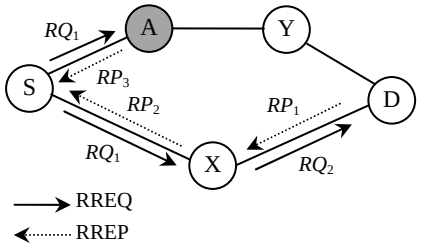
در این حمله، گره مهاجم پس از دریافت یک بسته RREQ خود را به عنوان یک گره میانی که مسیری تازه به مقصد دارد معرفی می‌کند. بدین منظور، گره مهاجم یک بسته RREP جعلی به شکل زیر تولید کرده و به سمت گره مبدا ارسال می‌کند:

مقدار فیلد شمارنده گام به عدد یک کاهش داده می‌شود.

آدرس مبدا اولیه و آدرس مقصد نهایی بدون تغییر در فیلدهای آدرس گره مبدا و آدرس گره مقصد بسته قرار داده می‌شوند.

شماره توالی مقصد حداقل یک واحد افزایش می‌یابد.

آدرس گره مبدا در سرآیند بسته با یک آدرس نامعتبر در شبکه جایگزین می‌شود.



شکل ۱: حمله سیاه‌چاله با بسته‌های RREP جعلی

شکل ۱ مثالی از حمله سیاه‌چاله با بسته‌های RREP جعلی را نمایش می‌دهد. فرض کنید گره مبدا S قصد دارد تا ارتباطی را با گره مقصد D برقرار کند. همچنین، فرض کنید مقدار شماره توالی گره D در جدول مسیریابی گره S برابر با ۲۰ باشد. گره X با دریافت بسته RQ1 آن را به سمت گره D هدایت[6] می‌کند. گره مهاجم A با دریافت بسته RQ1 با بسته RP3 به گره S پاسخ می‌دهد. گره S براساس فیلد شماره توالی مقصد مسیر معرفی شده توسط مهاجم را انتخاب کرده و داده‌های خود را به سمت گره نامعتبر Z در شبکه ارسال می‌کند. جزئیات بسته‌های فوق در جدول ۱ نمایش داده شده است.

[1] Blackhole
[2] Reactive
[3] Destination sequence number
[4] Broadcasting
[5] Hop count
[6] Forward

جدول ۱: مقدار بسته‌های RREQ و RREP

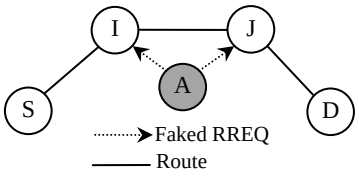
	RQ_1	RQ_2	RP_1	RP_2	RP_3
آدرس IP مبدا	S	X	D	X	Z
شماره توالی مقصد	۲۰	۲۰	۲۱	۲۱	۳۰
آدرس مبدا اولیه	S	S	S	S	S
آدرس مقصد نهایی	D	D	D	D	D
شمارنده گام	۱	۲	۱	۲	۱

۲-۴. حمله سیاه‌چاله با جعل بسته‌های RREQ

در این حمله، گره مهاجم پس از دریافت یک بسته RREQ تغییرات زیر را در آن اعمال کرده و بسته جعلی را در شبکه همه‌پخشی می‌کند:

آدرس مبدا اولیه و آدرس مقصد نهایی تغییری نمی‌کنند.
آدرس گره مبدا در سرآیند بسته با یک آدرس نامعتبر در شبکه جایگزین می‌شود.
فیلد شماره توالی مبدا حداقل یک واحد افزایش می‌یابد و یا فیلد شمارنده گام به عدد یک کاهش می‌یابد.
شناسه همه‌پخشی حداقل یک واحد افزایش می‌یابد.

سایر گره‌ها با دریافت این بسته، مسیرهای خود به سمت مقصد را براساس گره نامعتبر تنظیم می‌کنند. در نتیجه مسیرهای واقعی با مسیرهای جعلی جایگزین می‌شوند. شکل ۲ مثالی از این حمله را نمایش می‌دهد.



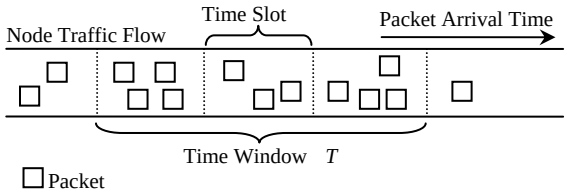
شکل ۲: حمله سیاه چاله با جعل بسته‌های RREQ

۵- روش پیشنهادی

در روش پیشنهادی، هر گره در هر بازه زمانی[1] با نظارت بر جریان ترافیک حالت شبکه را با بردار داده $x_i = [x_i^1, x_i^2, ..., x_i^p]^T$ نمایش می‌دهد که در آن هر x_i^j یک ویژگی قابل اندازه‌گیری است. چندین بازه زمانی یک پنجره زمانی[2] را تشکیل می‌دهند. بنابراین، هر گره در هر پنجره زمانی مجموعه‌ای از بردارهای داده را استخراج می‌کند. در شکل ۳ چگونگی استخراج یک بردار داده از جریان ترافیک گره نمایش داده شده است.

[1] Time slot

[2] Time window



شکل ۳: استخراج بردارهای داده در یک پنجره زمانی

در این مقاله، برای توصیف رفتار شبکه‌های اقتضایی متحرک با پروتکل مسیریابی AODV از چهارده ویژگی ارائه شده در [۱۲] استفاده می‌شود. تحرک گره‌ها در شبکه‌های اقتضایی متحرک موجب تغییر همبندی و در نتیجه باعث کاهش کارآیی شبکه می‌شود [۱۳]. هر مجموعه از بردارهای داده متناسب با حالت شبکه در زمان استخراج خود می‌باشد که با توجه به تغییر رفتار شبکه در زمان‌های بعدی این مجموعه از بردارهای داده نمی‌تواند رفتار شبکه را به خوبی توصیف کند. به عبارت دیگر، از اهمیت این مجموعه از بردارهای داده در زمان‌های بعدی کاسته می‌شود. به این دلیل وزن‌دهی به هر بردار داده می‌تواند در تشخیص ناهنجاری به صورت پویا و منطبق با تغییر رفتار شبکه مفید باشد. این وزن‌دهی را می‌توان بر اساس نرخ تغییر همبندی شبکه و زمان استخراج آن بردار داده انجام داد. برای وزن‌دهی به بردارهای داده استخراج شده در هر پنجره زمانی t از رابطه (۱) [۱۲] استفاده می‌شود:

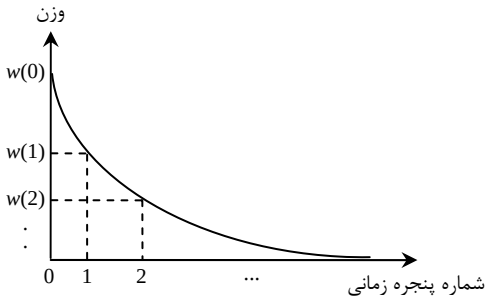
$$\begin{cases} w(t) = w(0)e^{-(NMR)(\Delta T/100)t} & t \le m \\ w(t) = 0 & t > m \end{cases} \quad (۱)$$

که $w(0)$ وزن اولیه برای مجموعه بردارهای داده، NMR[3] نرخ تحرک گره، T اندازه پنجره زمانی و t شماره پنجره زمانی است. هر گره تنها بردارهای داده استخراج شده در m پنجره زمانی قبلی را نگه‌داری می‌کند. وزن‌های $w(t), t = 1,..., m$ متناظر با هر پنجره زمانی به گونه‌ای محاسبه می‌شوند که

$$\sum_{t=1}^{m} w(t) = 1 \quad (۲)$$

با گذشت زمان و افزایش شماره پنجره زمانی مقدار وزن مجموعه بردارهای داده کاهش می‌یابد. شکل ۴ این وضعیت را نمایش می‌دهد.

[3] Node Mobility Ratio



شکل ۴: تغییر وزن هر مجموعه از بردارهای داده با گذشت زمان

یک معیار تحرک در شبکه‌های اقتضایی متحرک تعداد گره‌های همسایه هر گره است [۱۳]. پارامتر *NMR* با استفاده از این معیار محاسبه می‌شود. اگر *FNL*[1] نشان دهنده گره‌های همسایه هنگام استخراج مجموعه بردارهای داده در اولین پنجره زمانی و *CNL*[2] نشان دهنده گره‌های همسایه در پنجره زمانی جاری باشد، آنگاه پارامتر *NMR* با استفاده از رابطه (۳) محاسبه می‌شود:

$$NMR = \frac{|FNL - CNL| + |CNL - FNL|}{n} \qquad (۳)$$

که n تعداد کل گره‌ها در شبکه است.

روش پیشنهادی از دو مرحله آموزش و تشخیص تشکیل می‌شود که در ادامه هر یک از این دو مرحله شرح داده می‌شوند.

۵-۱. مرحله آموزش

در این مرحله، هر گره یک مجموعه از بردارهای داده را استخراج کرده و به آن وزن اولیه $w(0)$ را نسبت می‌دهد. سپس، این مجموعه از بردارهای داده را نرمال‌سازی کرده و الگوریتم خوشه‌بندی وزن‌دار WFWC[3] را بر روی آن اعمال می‌کند. در نهایت، از بین خوشه‌های تولید شده خوشه‌های عادی را تشخیص داده و از آن‌ها برای ایجاد نمای عادی استفاده می‌کند.

۵-۱-۱. نرمال‌سازی

مقادیر ویژگی‌های هر بردار داده می‌توانند اختلاف قابل توجهی با یکدیگر داشته باشند. بنابراین، هنگامی که فاصله بین دو بردار داده محاسبه می‌شود، ویژگی‌هایی که مقادیر بزرگتری دارند بر ویژگی‌ها با مقادیر کوچکتر غلبه می‌کنند. برای اطمینان از این که همه ویژگی‌ها تاثیر یکسانی بر محاسبه فاصله دارند، هر ویژگی $x_i^j, j = 1,..., p$ باید قبل از اجرای الگوریتم خوشه‌بندی به بازه [۱و۰] مقیاس داده شود:

$$\hat{x}_i^j = \frac{x_i^j - \min(x^j)}{\max(x^j) - \min(x^j)} \qquad (۴)$$

که $\min(x^j)$ و $\max(x^j)$ به ترتیب کوچکترین و بزرگترین مقادیر j امین ویژگی هستند.

۵-۱-۲. الگوریتم خوشه‌بندی وزن‌دار WFWC

الگوریتم خوشه‌بندی وزن‌دار WFWC گونه‌ای از الگوریتم K-Means [۱۴] با شعاع ثابت است. در این الگوریتم مجموعه‌ای از بردارهای داده با وزن‌های متفاوت به عنوان ورودی دریافت شده و مجموعه‌ای از خوشه‌ها با شعاع ثابت تولید می‌شود. فرآیند خوشه‌بندی در سه مرحله انجام می‌شود: در مرحله اول، یک مجموعه از خوشه‌ها با شعاع ثابت تولید می‌شود. بدین منظور، ابتدا بردار داده x_i از مجموعه بردارهای داده X انتخاب می‌شود. در صورت تهی بودن مجموعه خوشه‌های C بردار داده x_i به عنوان مرکز اولین خوشه در نظر گرفته می‌شود. در غیر این صورت، اگر فاصله نزدیکترین خوشه از x_i کمتر از باشد، x_i به این خوشه اضافه می‌شود و در صورت برآورده نشدن این شرط یک خوشه جدید با مرکز x_i تولید می‌شود. در مرحله دوم، مرکز هر خوشه با توجه به وزن بردارهای داده موجود در آن خوشه و با استفاده از رابطه (۵) به‌روزرسانی می‌شود:

$$\mu_j = \frac{\sum_{x_i \in c_j} w_i x_i}{\sum_{x_i \in c_j} w_i} \qquad (۵)$$

procedure WFWC

input:
$X = \{(x_1, w_1), (x_2, w_2), ..., (x_n, w_n)\}$ // A set of data vectors
c // The cluster radius
output:
C // A set of clusters
begin
$C := \phi$
repeat
for each data vector $x_i \in X$ **do**
if $C = \phi$ **then**
Make a new cluster c_1 with centroid μ_1
$\mu_1 := \{x_i\}, C := \{c_1\}$
else
Find the nearest cluster $c_j \in C$ to x_i
if $d(x_i, c_j) < c$ **then**
$c_j := c_j \cup \{x_i\}$

[1] First Neighborhood List
[2] Current Neighborhood List
[3] Weighted Fixed Width Clustering

```
        else
            Make a new cluster c_k with centroid μ_k
            μ_k := {x_i}, C := C ∪ {c_k}
        end if
      end if
    end for
    for each cluster c_j ∈ C do
      Update centroid μ_j using equation (5)
    end for
    for each two clusters c_j, c_k ∈ C do
      if d(c_j, c_k) < c then
        Merge clusters c_j and c_k into a new cluster c*
        with centroid μ*
        Update centroid μ* using equation (6)
      end if
    end for
  until C does not change
end procedure
```

شکل ۵: شبه کد الگوریتم خوشه‌بندی وزن‌دار WFWC

در مرحله سوم، خوشه‌هایی که فاصله مراکز آن‌ها کمتر از شعاع ثابت باشد، با یکدیگر ادغام شده و مرکز خوشه جدید با استفاده از رابطه (۶) محاسبه می‌شود:

$$\mu^* = (w_j\mu_j + w_k\mu_k)/(w_j + w_k) \qquad (۶)$$

که w_j و w_k به ترتیب متوسط وزن‌های بردارهای داده موجود در خوشه‌های c_j و c_k می‌باشند. مراحل فوق آنقدر تکرار می‌شوند تا در مراکز خوشه‌های تولید شده هیچ گونه تغییری ایجاد نشود. در شکل ۵ شبه کد الگوریتم خوشه‌بندی وزن‌دار WFWC نمایش داده شده است.

۵-۱-۳. ایجاد نمای عادی

فرض کنید C مجموعه خوشه‌های حاصل از اعمال الگوریتم خوشه‌بندی وزن‌دار WFWC روی مجموعه بردارهای داده X باشد. برای ایجاد نمای عادی، ابتدا خوشه‌های عادی تشخیص داده می‌شوند. بدین منظور، از میانگین فاصله بین خوشه‌ای (ICD[1]) با l خوشه نزدیک استفاده می‌شود [۱۵]. برای هر خوشه c_j میانگین فاصله بین خوشه‌ای ICD_j با استفاده از رابطه (۷) محاسبه می‌شود:

$$ICD_j = \begin{cases} \frac{1}{l}\sum_{k}^{l} d(c_j, c_k) & l \le |C| - 1 \\ \frac{1}{|C|-1}\sum_{k}^{|C|-1} d(c_j, c_k) & l > |C| - 1 \end{cases} \qquad (7)$$

که $|C|$ تعداد کل خوشه‌ها و $d(c_j, c_k)$ فاصله اقلیدسی بین خوشه‌های c_j و c_k است. پارامتر l که تعداد نزدیکترین خوشه‌ها را نمایش می‌دهد توسط کاربر تعیین می‌شود. در ادامه، میانگین و انحراف معیار برای ICD همه خوشه‌ها محاسبه شده و خوشه‌ای که ICD آن از میانگین به‌علاوه یک انحراف معیار کمتر باشد به عنوان خوشه عادی تشخیص داده می‌شود.

۵-۲. مرحله تشخیص

در این مرحله، هر گره در هر پنجره زمانی یک مجموعه از بردارهای داده را استخراج کرده و به آن وزن اولیه $w(0)$ را نسبت می‌دهد. سپس، این مجموعه از بردارهای داده را نرمال‌سازی کرده و با نمای عادی مقایسه می‌کند تا وجود هر گونه ناهنجاری در شبکه را تشخیص دهد. در نهایت، با استفاده از بردارهای داده عادی، نمای عادی را به‌روزرسانی می‌کند.

۵-۲-۱. تشخیص ناهنجاری

نمای عادی شامل تعدادی خوشه با شعاع ثابت است. فرض کنید C_n مجموعه خوشه‌های عادی و X مجموعه بردارهای داده استخراج شده در پنجره زمانی جاری باشد. برای تشخیص ناهنجاری در شبکه، ابتدا فاصله نزدیکترین خوشه $c_j \in C_n$ از هر بردار داده $x_i \in X$ محاسبه می‌شود. سپس در صورتی که فاصله به دست آمده بزرگتر از باشد، آن بردار داده به عنوان ناهنجار تشخیص داده می‌شود. در نهایت، در صورتی که تعداد بردارهای داده ناهنجار در هر پنجره زمانی از یک حد آستانه بیشتر باشد، این رفتار به عنوان حمله گزارش می‌شود. در غیر این صورت، بردارهای داده عادی به نمای عادی اضافه می‌شوند.

۵-۲-۲. به‌روزرسانی نمای عادی

برای به‌روزرسانی نمای عادی، ابتدا وزن هر یک از بردارهای داده در نمای عادی با استفاده از رابطه (۱) و با توجه به زمان استخراج آن بردار داده به‌روزرسانی می‌شود. سپس بردارهایی که متعلق به بیش از m پنجره زمانی قبلی بوده‌اند، از نمای عادی حذف می‌شوند. در نهایت، الگوریتم خوشه‌بندی وزن‌دار WFWC روی بردارهای داده باقی‌مانده اعمال می‌شود تا نمای عادی جدید به دست آید.

[1] Inter Cluster Distance

۶- نتایج آزمایش‌ها

در این بخش، نتایج آزمایش‌های انجام شده برای ارزیابی کارآیی روش پیشنهادی شرح داده می‌شود.

۶-۱. محیط شبیه‌سازی

برای شبیه‌سازی شبکه‌های اقتضایی متحرک از شبیه‌ساز NS2 [۱۶] استفاده شد. در این شبیه‌سازی، از مدل ترافیکی CBR با طول بسته داده ۵۱۲ بایت و مدل تحرک RWP[1] [۱۷] در یک ناحیه به ابعاد ۱۰۰۰m ۱۰۰۰m استفاده شد. مدل ترافیکی CBR با استفاده از برنامه cbrgen.tcl و مدل تحرک RWP با استفاده از برنامه اجرایی setdest تولید شد. تعداد کل گره‌های شبکه ۳۰ گره در نظر گرفته شد.

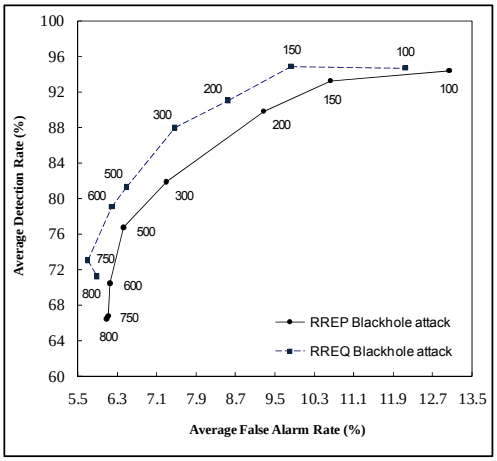
جدول ۲: پارامترهای شبیه‌سازی

Parameter	Value
Simulation time	10000(s)
No. of nodes	30
Routing protocol	AODV
Traffic model	CBR
Mobility model	RWP
Pause time	10(s)
Maximum mobility	20(m/s)
Maximum connection	30
Simulation area	1000(m) × 1000(m)
Transmission rate	250(m)
Maximum bandwidth	2(Mbps)
No. of malicious nodes	1

چنین فرض شد که هر گره به صورت مستقل یک مقصد تصادفی را انتخاب کرده و با سرعت یکنواختی بین صفر تا ۲۰m/s به سمت آن مقصد حرکت می‌کند. سپس، با رسیدن به مقصد به اندازه ۱۰ ثانیه توقف کرده و با انتخاب یک مقصد تصادفی دیگر به حرکت خود ادامه می‌دهد. در جدول ۲ خلاصه‌ای از پارامترهای شبیه‌سازی نمایش داده شده است [۶].

۶-۲. نتایج شبیه‌سازی

برای ایجاد نمای عادی، در مرحله آموزش، ابتدا مجموعه‌ای از بردارهای داده توسط هر گره استخراج شد. مدت زمان مورد استفاده برای جمع‌آوری این مجموعه از بردارهای داده، برابر ۱۰۰۰ ثانیه (یک دهم زمان کل شبیه‌سازی) و اندازه بازه زمانی برای استخراج هر بردار داده برابر ۵ ثانیه [۵، ۶] در نظر گرفته شد. به همه بردارهای داده وزن یکسان $w(0) = 1$ نسبت داده شد. سپس، با اجرای الگوریتم خوشه‌بندی وزن‌دار WFWC بر روی مجموعه بردارهای داده فوق، نمای عادی ایجاد شد. در آزمایش‌های انجام شده، یکی از گره‌ها به عنوان گره مهاجم در نظر گرفته شد. این گره در فاصله زمانی ۳۵۰۰ تا ۶۰۰۰ ثانیه حملات سیاه‌چاله را انجام داد. در مرحله تشخیص، هر گره در پایان هر پنجره زمانی و با توجه به بردارهای داده عادی در m پنجره زمانی قبلی نمای عادی را به‌روزرسانی کرد. با توجه به نتایج آزمایش‌های انجام شده، پارامتر $m = 3$ در نظر گرفته شد. برای ارزیابی روش پیشنهادی، از دو معیار نرخ تشخیص و نرخ هشدار نادرست استفاده شد.

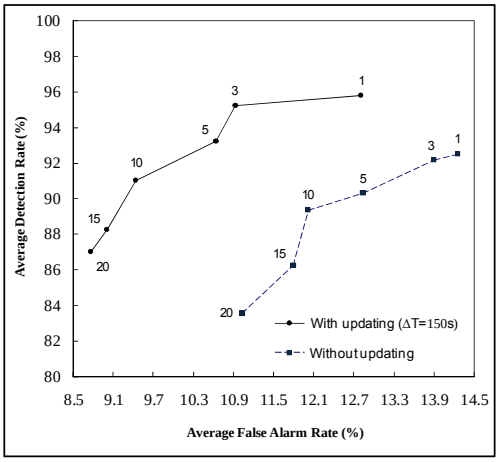


شکل ۶: متوسط نرخ تشخیص در برابر نرخ هشدار نادرست برای حمله سیاه‌چاله با بسته‌های RREP جعلی و حمله سیاه‌چاله با جعل بسته‌های RREQ به ازای اندازه‌های متفاوت پنجره زمانی

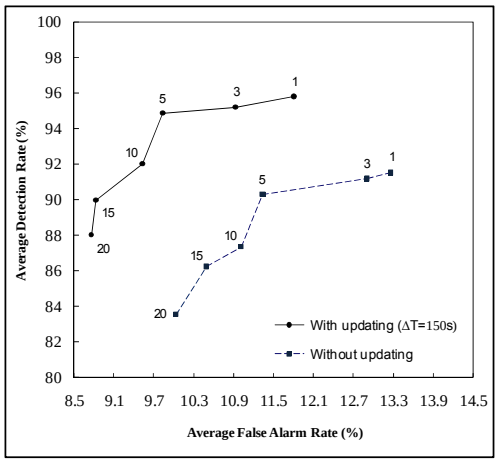
در شکل ۶ متوسط نرخ تشخیص در برابر نرخ هشدار نادرست برای حمله سیاه‌چاله با بسته‌های RREP جعلی و حمله سیاه‌چاله با جعل بسته‌های RREQ به ازای اندازه‌های متفاوت پنجره زمانی (T) نمایش داده شده است. در این آزمایش، نرخ تحرک ۵ m/s و شعاع ثابت خوشه‌بندی 0.35 = در نظر گرفته شد. با توجه به شکل فوق مشخص می‌شود که با کاهش اندازه پنجره زمانی یا به عبارت دیگر با به‌روزرسانی سریع نمای عادی نرخ تشخیص افزایش پیدا می‌کند. البته باید توجه داشت که با کاهش اندازه پنجره زمانی تعداد دفعات به‌روزرسانی نمای عادی افزایش می‌یابد که این مساله باعث افزایش سربار محاسباتی و در نتیجه افزایش مصرف انرژی در هر گره می‌شود.

در شکل‌های ۷ و ۸ متوسط نرخ تشخیص در برابر نرخ هشدار نادرست برای دو نوع حمله سیاه‌چاله به ازای نرخ‌های متفاوت تحرک نمایش داده شده است. در این آزمایش‌ها، روش پیشنهادی در دو حالت به‌روزرسانی نمای عادی و عدم به‌روزرسانی نمای عادی مورد ارزیابی قرار گرفت. اندازه پنجره زمانی $T = 150s$ و شعاع ثابت خوشه‌بندی 0.35 = در نظر گرفته شد. با توجه به شکل‌های فوق مشخص می‌شود که به‌روزرسانی نمای عادی، تاثیر قابل ملاحظه‌ای در افزایش نرخ تشخیص و کاهش نرخ هشدار نادرست دارد.

[1] Random WayPoint Model



شکل ۷: متوسط نرخ تشخیص در برابر نرخ هشدار نادرست برای حمله سیاه‌چاله با بسته‌های RREP جعلی به ازای نرخ‌های تحرک



شکل ۸: متوسط نرخ تشخیص در برابر نرخ هشدار نادرست برای حمله سیاه‌چاله با جعل بسته‌های RREQ به ازای نرخ‌های تحرک

۷- نتیجه‌گیری

با توجه به این که همبندی پویا در شبکه‌های اقتضایی متحرک موجب تغییر در رفتار شبکه می‌شود، استفاده از یک نمای عادی از پیش تعریف شده نمی‌تواند رفتار شبکه را به خوبی توصیف کند. در این مقاله، روشی پویا برای تشخیص ناهنجاری در این شبکه‌ها پیشنهاد شده است. در روش پیشنهادی، هر گره در هر پنجره زمانی مجموعه‌ای از بردارهای داده را استخراج می‌کند. به مجموعه بردارهای داده استخراج شده با توجه به نرخ تغییر همبندی شبکه و زمان استخراج وزن متفاوتی نسبت داده می‌شود. روش پیشنهادی از دو مرحله آموزش و تشخیص تشکیل می‌شود. در مرحله آموزش، الگوریتم خوشه‌بندی وزن‌دار WFWC بر روی مجموعه بردارهای داده اعمال شده و نمای عادی ایجاد می‌شود. در مرحله تشخیص، در پایان هر پنجره زمانی ابتدا بردارهای داده عادی به نمای عادی اضافه می‌شوند. سپس، با به‌روزرسانی وزن بردارهای داده در نمای عادی و اعمال الگوریتم WFWC این نما به‌روزرسانی می‌شود. با انجام آزمایش‌های مختلف عملکرد روش پیشنهادی برای تشخیص حمله سیاه‌چاله با بسته‌های RREP جعلی و حمله سیاه‌چاله با جعل بسته‌های RREQ مورد ارزیابی قرار گرفت. در [۷] روشی ارائه شده است که ناهنجاری را منطبق با مدل‌های تحرک تشخیص می‌دهد. اما این روش از یک نمای عادی از پیش تعریف شده استفاده می‌کند و در آن امکان به‌روزرسانی نمای عادی وجود ندارد. روش‌های ارائه شده در [۸، ۱۲] از تشخیص ناهنجاری شبه‌نظارتی استفاده می‌کنند که در صورت وجود هر گونه ناهنجاری در مجموعه داده‌های آموزشی، این روش‌ها قادر به تشخیص آن‌ها نمی‌باشند. در حالی که روش پیشنهادی در این مقاله یک روش تشخیص ناهنجاری بدون‌نظارت است که در آن نیازی به داده‌های فاقد حمله در مرحله آموزش وجود ندارد. همچنین، در روش پیشنهادی نمای عادی در هر پنجره زمانی به‌روزرسانی می‌شود. با توجه به نتایج آزمایش‌های انجام شده مشخص شد که این به‌روزرسانی تاثیر قابل ملاحظه‌ای در افزایش نرخ تشخیص و کاهش نرخ هشدار نادرست دارد.

سپاسگزاری

این تحقیق با حمایت مالی مرکز تحقیقات مخابرات ایران و تحت قرارداد با کد شناسایی ۱۲۸-۱۲-۸۸ انجام شده است.

مراجع

[1] I. Chlamtac, M. Conti, and J. N. Liu, “Mobile ad hoc networking: imperatives and challenges,” Ad Hoc Networks, vol. 1, no. 1, pp. 13–64, July 2003

[2] H. Debar, M. Dacier, and A. Wespi, “A revised taxonomy for intrusion detection systems,” Annals of Telecommunications, vol. 55, no. 7–8, pp. 361–78, July 2000.

[3] E. Eskin, A. Arnold, M. Prerau, L. Portnoy, and S. Stolfo, “A geometric framework for unsupervised anomaly detection: detecting intrusions in unlabeled data,” *Applications of Data Mining in Computer Security*, Kluwer, 2002.

[4] S. Lee, B. Han, and M. Shin, “Robust routing in wireless ad hoc networks,” in Proceedings of the 2002 International Conference on Parallel Processing Workshops, pp. 73, Vancouver, Canada, August 2002.

[5] Y. A. Huang, W. Fan, W. Lee, and P. S. Yu, “Cross-feature analysis for detecting ad-hoc routing anomalies,” in *Proceedings of the 23rd International Conference on Distributed Computing Systems*, pp. 478–487, Washington DC, USA, May 2003.

[6] Y. A. Huang and W. Lee, “Attack analysis and detection for ad hoc routing protocols,” in *Proceedings of the 7th International Symposium on Recent Advances in Intrusion Detection (RAID’04)*, pp. 125–145, Riviera, French, September 2004.

[7] B. Sun, K. Wu, and U. Pooch, “Towards adaptive

intrusion detection in mobile ad hoc networks," in *Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM'04)*, vol. 6, pp. 3551–3555, Dallas, TX, USA, November 2004.
[8] S. Kurosawa, H. Nakayama, N. Kato, A. Jamalipour, and Y. Nemoto, "Detecting Blackhole attack on AODV-based mobile ad hoc networks by dynamic learning method," *International Journal of Network Security*, vol. 5, no. 3, pp. 338–346, November 2007.
[9] C. E. Perkins, E. M. B. Royer, and S. R. Das, Ad hoc On-Demand Distance Vector (AODV) Routing, RFC 3561, July 2003.
[10] P. Ning and K. Sun, "How to misuse AODV: A case study of insider attacks against mobile ad-hoc routing protocols," in *Proceedings of the 4th Annual IEEE Information Assuranca Workshop*, pp. 60–67, West Point, NY, USA, June 2003.
[11] C. Hongsong, J. Zhenzhou, and H. Mingzeng, "A novel security agent scheme for AODV routing protocol based on thread state transition," *Asia Journal of Information Technology*, vol. 5, no. 1, pp. 54–60, 2006.
[12] H. Nakayama, S. Kurosawa, A. Jamalipour, Y. Nemoto, and N. Kato, "A dynamic anomaly detection scheme for AODV-based mobile ad hoc networks," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 5, pp. 2471–2481, June 2009.
[13] J. Tsumochi, K. Masayama, H. Uehara, and M. Yokoyama, "Impact of mobility metric on routing protocols for mobile ad hoc networks," in *Proceedings of the IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM'03)*, vol. 1, pp. 322–325, August 2003.
[14] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, , pp. 281–297, Berkeley, University of California Press, January 1967.
[15] S. Rajasegarar, C. Leckie, M. Palaniswami, and J. C. Bezdek, "Distributed anomaly detection in wireless sensor networks," in *Proceedings of 10th IEEE International Conference on Communications Systems (ICCS'03)*, Singapore, October 2006.
[16] NS2: The Network Simulator

http://www.isi.edu/nsnam/ns/.
[17] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless Communications and Mobile Computing*, vol. 2, no. 5, pp. 483–502, September 2002.