نشریه علمی "پدافند الکترونیکی و سایبری"
سال هفتم، شماره ۴، زمستان ۱۳۹۸، ص ۸۹-۷۹

ارائه چارچوب مبتنی بر هستان‌شناسی برای ادغام داده‌های سخت و نرم در تحلیل امنیت سایبری

علی‌جبار رشیدی [1*]، سعداله سبحانی [2]، سید مجتبی حسینی [3]
۱- دانشیار، ۲- دانشجوی دکتری، ۳- استادیار، دانشگاه صنعتی مالک اشتر

(دریافت: ۹۷/۰۵/۲۳، پذیرش: ۹۷/۱۲/۱۴)

چکیده

در تحلیل امنیت سایبری، علاوه بر داده‌ها و اطلاعاتی که از حسگرهای ماشینی مانند سامانه‌های تشخیص نفوذ، دیواره‌های آتش و پویشگرهای آسیب‌پذیری به‌دست می‌آید (داده‌های سخت)، مشاهدات و برداشت‌های انسانی شامل گزارش‌های کاربران و مدیران شبکه از کارکرد عادی یا غیرعادی اجزای شبکه و تشخیص‌های صورت گرفته توسط تحلیلگرهای امنیتی از وضعیت امنیتی شبکه (داده‌های نرم) می‌تواند نقش مهمی در رسیدن به تخمین و تصمیم دقیق‌تر و مطمئن‌تر داشته باشد. ادغام داده‌های سخت و نرم در تحلیل امنیت سایبری دارای چالش‌هایی از قبیل طراحی چارچوب مدل‌سازی مسئله و نمایش انواع مختلف عدم قطعیت است. در این مقاله مدل جدیدی مبتنی بر هستان‌شناسی جهت ادغام داده‌های سخت و نرم به‌منظور به‌کارگیری در تحلیل امنیت سایبری ارائه می‌شود. در ابتدا مفاهیم و متغیرهای مسئله مدل می‌شوند و سپس با استفاده از مجموعه قواعد، استنتاج وضعیت امنیتی دارایی‌ها صورت می‌گیرد. همچنین مدل باور انتقال‌پذیر و قاعده ترکیب دمپستر- شفر برای مدل‌سازی یکپارچه عدم قطعیت و ادغام داده‌ها به کار گرفته شده است. نتایج به‌کارگیری مدل پیشنهادی در یک سناریوی نمونه از تحلیل امنیت سایبری، عملیاتی بودن آن را در ادغام داده‌های سخت و نرم سایبری نشان می‌دهد. انعطاف‌پذیری بالا و پویایی مدل با توجه به قابلیت توسعه هستان‌شناسی و پایگاه دانش، از ویژگی‌های مدل پیشنهادی است.

کلید واژه‌ها: ادغام داده‌های سخت و نرم، مدل‌سازی عدم قطعیت، تحلیل امنیت سایبری، هستان‌شناسی.

۱. مقدمه

این یک واقعیت پذیرفته شده است که انسان در بعضی از موارد دارای قدرت درک، پردازش و تشخیص بالاتری نسبت به ماشین است. از طرف دیگر برخی از حوزه‌ها و محیط‌ها به اندازه کافی در دسترس حسگرهای فیزیکی نیست، مثلاً تشخیص نیت یک فرد از روی دنباله حرکات و اقداماتی که انجام می‌دهد، بر اساس پردازش داده‌های حسگرهای ماشینی بسیار دشوار است، اما یک فرد خبره بر اساس دانش و تجربه خود، می‌تواند با اطمینان بالایی نیت فرد موردنظر را تشخیص دهد. موارد زیادی هم می‌تواند وجود داشته باشد که تشخیص ماشین یقینی و قطعی نیست و در این موارد نظر و تشخیص انسان می‌تواند برای به‌دست آوردن یقین و اطمینان بیشتر نسبت به‌درستی تشخیص صورت گرفته کمک کند و در نتیجه بهبود آگاهی وضعیتی را در پی داشته باشد. برای مثال، در تحلیل امنیت سایبری، بالا رفتن ترافیک شبکه در یک زمان خاص می‌تواند نشانگر یک وضعیت غیر نرمال باشد، و هم می‌تواند بر اساس شرایط زمانی خاص، یک وضعیت عادی باشد. نظر تحلیل‌گر یا مدیر شبکه با توجه به تجربه و خبرگی در این زمینه، می‌تواند در ارزیابی نهایی به سیستم کمک کرده و نتیجه را بهبود دهد.

انواع داده‌ها و اطلاعات ورودی قابل استفاده در سامانه‌های ادغام [1] اطلاعات به دو دسته کلی داده‌های سخت و داده‌های نرم [2] تقسیم‌بندی می‌شود. داده‌های سخت به داده‌ها و اطلاعات به‌دست‌آمده از حسگرهای ماشینی و داده‌های نرم به داده‌ها و اطلاعات انسانی گفته می‌شود [۱].

در حوزه تحلیل امنیت سایبری، داده‌های سخت شامل هشدارهای تولیدشده توسط سامانه‌های تشخیص نفوذ (IDS [3])، وقایع ثبت‌شده دیواره‌های آتش [4] و خروجی‌های

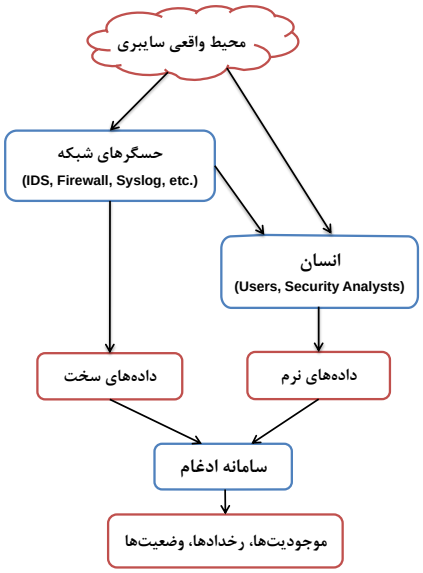
* رایانامه نویسنده مسئول: rashidi@mut.ac.ir

[1] Fusion
[2] Hard data and Soft data
[3] Intrusion Detection System
[4] Firewalls

پویشگرهای آسیب پذیری[1] هستند و داده های نرم شامل گـزارش کاربران یا مدیران شبکه‌ها از کارکرد عادی یا غیرعادی شـبکه یـا سیستم مورد استفاده‌شان و همچنـین اظهارنظرهـای تحلیلگـران امنیتی شبکه نسبت به رخدادها یا وضعیت‌هایی از محیط سایبری خواهد بود.

داده‌های نرم می‌تواند اطلاعات به‌دست آمده از داده‌های سخت را تأیید یا تکمیل کند. مثلاً گزارش مدیر شبکه مبنی بر کند شدن عملکرد سرور، احتمال وقوع یک حمله را که از روی داده‌های سخت به‌دست آمده، تقویت می‌کند. یا اینکه در یک سناریوی حمله، ممکن است برخی گام‌های حمله به خاطر استفاده مهاجم از یک مجوز تعریف شده، تشخیص داده نشود و گزارش کاربر مبنی بر فعالیت‌های غیرنرمال در صفحه کاربری‌اش، می‌تواند اطلاعات موردنیاز برای تشخیص حمله را تکمیل کند.

در این پژوهش، فرآیند کلی ورود داده‌های سخت و نرم به سامانه ادغام به این صورت در نظر گرفته شده است که گزارش‌هایی از محیط توسط حسگرهای ماشینی ارائه می‌شود که داده سخت را تشکیل می‌دهد. همچنین گزارش‌هایی هم توسط انسان تولید و ارائه می‌شود. این گزارش‌های انسانی که داده نرم نامیده می‌شود، بر اساس تحلیل‌هایی روی داده‌های حسگری و یا مبتنی بر مشاهدات مستقیم از محیط حاصل می‌شود. سپس گزارش‌ها به سامانه ادغام داده‌های سخت و نرم وارد می‌شوند تا با ادغام آن‌ها تخمین دقیق‌تر و با قابلیت اطمینان بیشتر از موجودیت‌ها، رخدادها و وضعیت‌های ممکن در محیط مورد بررسی به‌دست آید. این فرآیند برای محیط سایبری در شکل (۱) نشان داده شده است.



شکل (۱): فرآیند کلی ورود داده‌های سخت و نرم به سامانه ادغام

هر کدام از داده‌های سخت و داده‌های نرم و منابع آن‌ها دارای ویژگی‌های خاص خود است و انواع مختلفی از عدم قطعیت و ناکاملی اطلاعات می‌تواند در آن‌ها وجود داشته باشد. برای مثال داده‌های سخت بیشتر دارای مقادیر کمّی هستند و داده‌های نرم دارای مقادیر کیفی هستند و یا در گزارش‌های انسانی از کلمات فازی بیشتر استفاده می‌شود. با توجه به این تفاوت‌ها ادغام داده‌های نرم با داده‌های سخت دارای مشکلات و چالش‌های خاصی هم در سطح معماری و هم در سطح روش‌ها و الگوریتم‌ها است.

از مسائل و چالش‌هایی که در ادغام داده‌های سخت و نرم وجود دارد، موارد زیر را می‌توان برشمرد:

- طراحی یک معماری و چارچوب کلی مشخص‌کننده فرآیند پردازش و ادغام داده‌ها
- نحوه مدل‌سازی مسئله و روش استنتاج و استفاده از دانش زمینه
- نحوه مدل کردن عدم‌قطعیت و ابهام موجود در داده‌های نرم و مخصوصاً میزان باور و اطمینان تصریح‌شده در کنار داده‌ها
- نحوه اعتبارسنجی منابع داده‌های انسانی و نحوه دخالت دادن اعتبار داده‌ها در فرآیند ادغام
- وجود ناسازگاری در داده‌ها و شواهد و نیاز به حل و رفع و اداره کردن آن‌ها

یک نیازمندی مهم در تحلیل امنیت سایبری با استفاده از منابع داده‌ای مختلف، طراحی یک مدل مفهومی یکپارچه از مفاهیم و موجودیت‌های فضای مسئله است. با توجه به قابلیت‌ها و مزایایی که هستان‌شناسی در مدل کردن مفاهیم مسئله و نمایش دانش حوزه دارد، توسعه یک هستان‌شناسی امنیت، یک راه‌حل ممکن برای این نیازمندی است.

در حوزه تحلیل امنیت سایبری، در تحقیقات مختلفی به استنتاج مبتنی بر هستان‌شناسی پرداخته شده است، اما در این تحقیقات داده نرم به‌عنوان یک نوع داده ورودی قابل استفاده در استنتاج وضعیت امنیتی شبکه به کار گرفته نشده است. همچنین در استنتاج‌های مبتنی بر هستان‌شناسی، به موضوع ادغام شواهد مختلف به‌دست‌آمده برای یک فرضیه و به‌کارگیری عدم قطعیت در فرآیند استنتاج، خوب پرداخته نشده است. از طرف دیگر در کارهای انجام‌شده در حوزه ادغام داده‌های سخت و نرم، از هستان‌شناسی به‌عنوان ابزاری برای مدل‌سازی مفهومی مسئله و مبنای مرتبط‌سازی و ادغام داده‌ها استفاده نشده است.

بر این اساس، در این مقاله چارچوب جدیدی جهت ادغام

[1] Vulnerability Scanners

داده‌های سخت و نرم به‌منظور به‌کارگیری در تحلیل امنیت سایبری ارائه می‌شود. در چارچوب پیشنهادی، از هستان‌شناسی برای تعریف نحو و معنای مورد استفاده در بیان داده‌ها و اطلاعات قابل استنتاج و مرتبط‌سازی آن‌ها با هم استفاده می‌شود. همچنین مدل باور انتقال‌پذیر و قاعده ترکیب دمپستر برای مدل‌سازی عدم قطعیت و ادغام داده‌های سخت و نرم به کار گرفته می‌شود.

در ادامه این مقاله در بخش دوم به سابقه تحقیقات در حوزه ادغام داده‌های سخت و نرم و همچنین در زمینه تحلیل امنیت سایبری مبتنی بر هستان‌شناسی پرداخته شده است. در بخش سوم چارچوب و روش پیشنهادی تشریح می‌گردد. در بخش چهارم نحوه به‌کارگیری مدل پیشنهادی در یک سناریوی نمونه از تحلیل امنیت سایبری نشان داده می‌شود و در بخش پنجم نیز جمع‌بندی و نتیجه‌گیری مقاله ارائه می‌شود.

۲. سابقه تحقیقات

در زمینه ادغام داده‌های سخت و نرم، تحقیقات مختلفی در قالب یک پروژه پنج ساله بین دانشگاهی تحت مدیریت و راهنمایی هال، لیناس و دیگر همکارانشان، از جمله [۲] و [۳] به جنبه‌های مختلف این مسئله پرداخته است. این پروژه با هدف استفاده از گزارش‌های میدانی در کشف اقدامات آشوبگرانه و تروریستی تعریف شده است. یک نکته مهم در این تحقیقات، مربوط به چارچوب و معماری کلی ادغام است، مثلاً این‌که قبل از شروع ادغام داده‌ها چه پردازش‌ها و تبدیل‌هایی لازم است صورت گیرد، ادغام داده‌های سخت و نرم در چه مرحله از دنباله پردازش‌های لازم بر روی داده‌ها صورت گیرد، و این‌که ادغام داده‌ها و اطلاعات دارای چه مراحل کلی است.

در معماری ارائه‌شده در [۲] برای ادغام داده‌های سخت و نرم، سه مرحله کلی ارجاع‌دهی مشترک[1]، مرتبط‌سازی داده‌ها[2] و تخمین وضعیت[3] در نظر گرفته شده است. ارجاع‌دهی مشترک شامل عملیاتی برای قالب‌بندی مشترک داده‌ها، هم‌تراز کردن مکانی و زمانی داده‌ها، و نرمال‌سازی ضرایب اطمینان و نمایش‌های مختلف عدم قطعیت است. مرتبط‌سازی داده‌ها شامل عملیاتی برای تولید، ارزیابی و انتخاب فرضیه‌ها بر اساس ارتباط بین داده‌های ورودی است. تخمین وضعیت شامل عملیاتی برای استخراج و مشخص‌سازی ویژگی‌های شیء و ویژگی‌های وضعیت است.

در [۴] نظریه امکان بر اساس توانایی‌اش در نمایش مناسب عدم قطعیت‌های موجود در گزارش‌های انسانی و به خاطر سادگی محاسباتی آن، به‌عنوان مدل پایه برای نمایش و مدل‌سازی عدم قطعیت در داده‌ها، انتخاب شده است و برای داده‌های با نمایش احتمالاتی، یک تبدیل به نمایش امکانی انجام می‌شود.

در [۵] برای ادغام داده‌های سخت و نرم در کاربرد ردگیری اهداف، از نظریه مجموعه تصادفی[4] استفاده شده است. طبق این تحقیق، نظریه مجموعه تصادفی، امکان نمایش و مدل‌سازی انواع مختلف عدم قطعیت موجود در داده‌های نرم و داده‌های سخت را فراهم می‌کند. برای ادغام داده‌های نرم، توسعه‌ای از روش فیلتر کالمن به نام فیلتر شهودی کالمن (KEF[5]) به کار گرفته شده است. KEF با به‌کار گرفتن نظریه مجموعه تصادفی در داخل چارچوب تخمین بیزین مشتق شده است و قادر به پردازش داده‌های غیردقیق و همچنین غیرشفاف است. در این تحقیق، فرض شده است که هر کدام از منابع ورودی نرم و سخت، داده‌ها و اطلاعاتی را در مورد نحوه حرکت اهداف مورد بررسی در یک محیط، بر اساس قالب و نحو از پیش تعیین‌شده، ارائه می‌کنند. مطابق این نحو، گزارش دهنده می‌تواند میزان اطمینان خود را در مورد اطلاعات ارائه شده در قالب کلماتی مانند "قطعاً" و "شاید" بیان کند.

در [۶] یک چارچوب تحلیلی مبتنی بر نظریه باور دمپستر- شفر برای ادغام داده‌های سخت و نرم ارائه شده و از یک رویکرد شرطی به نام مفهوم شرطی Fagin-Halpern برای به‌روزرسانی شواهد و باور استفاده می‌شود. در این تحقیق، از یک راهبرد مرحله‌ای و گام‌به‌گام برای به‌روزرسانی شواهد استفاده می‌شود.

در [۷] یک رویکرد مبتنی بر شبکه‌های بیزین چند- موجودیتی فازی (Fuzzy MEBN[6]) برای ادغام داده‌های سخت و نرم ارائه شده است. Fuzzy MEBN از قدرت بیانگری منطق مرتبه اول برای روابط معنایی بهره می‌گیرد، از توانایی شبکه‌های بیزین در مدیریت عدم قطعیت استفاده می‌کند، و منطق فازی را هم برای مدل کردن ابهام موجود در واژه‌های زبانی به کار می‌گیرد. در رویکرد ارائه شده، ابتدا موجودیت‌هایی که تأثیر بالایی در ایمنی سیستم دارند، به همراه روابط علّی و روابط معنایی بین موجودیت‌ها تعیین می‌شوند. سپس یک مدل بر اساس Active Fuzzy-MEBN ارائه می‌شود که اجازه می‌دهد موجودیت‌ها و روابط آن‌ها در یک حوزه خاص مدل شود، ارزیابی وضعیت جاری سیستم انجام شود، حالت‌های آتی آن‌ها پیش‌بینی شود و درنهایت تصمیم بهینه پیشنهاد شود.

[1] Common Referencing
[2] Data Association
[3] State estimation
[4] Random Set Theory
[5] Kalman Evidential Filter
[6] Fuzzy Multi-Entity Bayesian Networks

جدول (۱) بر اساس معیارهایی از قبیل معماری کلان و روش ادغام مورد استفاده، به دسته‌بندی و مقایسه تحقیقات قبلی ادغام داده‌های سخت و نرم پرداخته است. لازم به ذکر است که با توجه به متفاوت بودن مسئله مورد بررسی این تحقیقات و یکسان نبودن نوع داده‌ها و نتایج ادغام، مقایسه از جهت قوت و ضعف روش‌ها امکان‌پذیر نیست و دسته‌بندی و مقایسه صورت گرفته با هدف کنار هم قرار دادن ویژگی‌های تحقیقات قبلی از چند جنبه مختلف است.

جدول (۱): دسته‌بندی و مقایسه تحقیقات قبلی ادغام داده‌های سخت و نرم

مرجع پژوهش	مسئله مورد بررسی	معماری کلان و فرآیند پردازش و ادغام اطلاعات	چارچوب نظری مدل‌سازی عدم قطعیت و ادغام	نحوه نمایش و مدل‌سازی داده‌ها
[۲-۴]	کشف اقدامات آشوبگرانه و تروریستی	معماری مبتنی بر مدل JDL	نظریه امکان به‌عنوان چارچوب هدف نمایش عدم قطعیت و تبدیل نمایش احتمالی به آن	گراف داده و گراف الگو نشان‌دهنده روابط بین موجودیت‌ها
[۵]	ردگیری اهداف و طبقه‌بندی اهداف	چارچوب پردازشی مبتنی بر فیلتر کالمن	نظریه مجموعه تصادفی	قالب و نحو از پیش تعیین‌شده
[۷]	تشخیص وضعیت خطر در شبکه‌های موردی خودرویی	معماری مبتنی بر Fuzzy MEBN	شبکه‌های بیزین چند موجودیتی فازی	مقادیر فازی یا احتمالی قابل انتساب به گره‌های شبکه بیزین
[۶]	بررسی شواهد جرم‌یابی	بدون معماری کلان و صرفاً دنباله‌ای از به‌روزرسانی شواهد	نظریه دمپستر-شفر و رویکرد شرطی برای به‌روزرسانی شواهد	شواهد و توابع جرم مبتنی بر چارچوب تشخیص نظریه باور

با توجه به بررسی صورت گرفته، محدودیت‌ها و کاستی‌های زیر در تحقیقات پیشین ادغام داده‌های سخت و نرم وجود دارد:

- ضعف در ارائه یک چارچوب یا ساختار مفهومی برای مدل‌سازی متغیرها یا مفاهیم مسئله و روابط بین آن‌ها
- مناسب بودن چارچوب ادغام ارائه‌شده برای مسائل محدود با ویژگی‌های خاص: مثلاً استفاده از شبکه‌های بیزین فازی که در تحقیقات پیشین وجود دارد، برای مسائلی که دارای متغیرهای محدود با روابط احتمالاتی بین آن‌ها بوده و دانش پیشین در مورد احتمالات وجود داشته باشد، مناسب است.

بنابراین، از بررسی تحقیقات پیشین، نیاز به ارائه یک چارچوب و روش ادغام داده‌های سخت و نرم مناسب برای مسئله تحلیل امنیت سایبری احساس می‌شود.

در زمینه تحلیل امنیت سایبری و آگاهی وضعیتی سایبری تحقیقات خیلی گسترده‌ای انجام شده است که بر اساس هدف‌گذاری و روش‌های مورد استفاده در این مقاله، به برخی از آن‌ها اشاره می‌شود.

در [۱۳] به‌عنوان محصول یک تلاش مشارکتی برای تعریف و توسعه یک زبان استاندارد برای نمایش ساخت‌یافته اطلاعات تهدیدهای سایبری، معماری و زبان STIX[1] ارائه شده است و در سال ۲۰۱۷ نیز نسخه ۲ آن عرضه شده است. این معماری و زبان، امکان توصیف و مشخص‌سازی اطلاعات موردنیاز برای کاربردهایی از قبیل تجزیه و تحلیل تهدیدهای سایبری، مدیریت فعالیت‌های واکنشی و مقابله‌ای و به اشتراک‌گذاری اطلاعات را فراهم می‌کند.

زبان STIX دارای مؤلفه‌های مختلفی برای توصیف تهدید از جنبه‌های گوناگون است. در این زبان، یک حمله یا تهدید سایبری با عناصر و ویژگی‌هایی شامل کنشگر تهدید، روش‌ها و ابزار مورد استفاده برای حمله، مقصد بهره‌کشی (آسیب‌پذیری‌ها و ضعف‌های مورد استفاده در حمله)، مشاهده‌پذیرها و نشانگرهای حمله، و اقدامات دفاعی ممکن توصیف و بیان می‌شود.

در کارهای تحقیقاتی مختلفی از هستان‌شناسی برای

[1] Structured Threat Information eXpression

رده‌بندی خودکار حملات، آسیب‌پذیری‌ها و هشدارها، تعیین خط‌مشی‌های امنیتی، تشخیص نفوذ، و استدلال در مورد آگاهی وضعیتی استفاده شده است. در [۸] از هستان‌شناسی و قواعد تعریف‌شده توسط افراد خبره، در آگاهی وضعیتی امنیت شبکه استفاده شده است. همچنین در [۹] یک رویکرد مبتنی بر هستان‌شناسی و زبان قواعد وب معنایی برای مدل کردن اطلاعات و عملیات سامانه‌های مدیریت اطلاعات و رخدادهای امنیتی ارائه شده است.

در [۱۱] ابتدا یک طبقه‌بندی از مفاهیم مرتبط با حملات شبکه‌ای ارائه شده است که شامل پنج بعد اثر حمله، بردار حمله، مقصد حمله، آسیب‌پذیری و دفاع است. سپس بر اساس طبقه‌بندی ارائه‌شده، یک هستان‌شناسی ساخته شده و با استفاده از پایگاه داده ملی آسیب‌پذیری (NVD) با اطلاعاتی در زمینه آسیب‌پذیری‌ها، ضعف‌ها و امتیازدهی آسیب‌پذیری‌ها پر شده است. در نهایت یک چارچوب مبتنی بر هستان‌شناسی برای ارزیابی امنیتی شبکه و سامانه‌های کامپیوتری و ارزیابی تأثیر حمله روی سامانه ارائه شده است.

در [۱۲] با استفاده از توانایی‌های استنتاجی مدل مبتنی بر هستان‌شناسی، یک چارچوب کارا برای تولید گراف‌های حمله و ارزیابی امنیت شبکه ارائه شده است. در این چارچوب، یک هستان‌شناسی برای تعریف مفاهیم امنیت شامل دستگاه‌ها و مؤلفه‌های شبکه، آسیب‌پذیری و حمله و مشخص کردن روابط بین آن‌ها طراحی شده است و از قواعد SWRL برای استنتاج رفتارهای حمله بالقوه و دسترسی‌ها یا امتیازات گوناگونی که مهاجمین از طریق این حملات می‌توانند به‌دست آورند، استفاده می‌کند. چارچوب ارائه‌شده برای ارزیابی امنیت شبکه، مراحل مختلفی شامل ساخت دانش حمله، به‌دست آوردن اطلاعات شبکه، نمونه‌سازی از هستان‌شناسی امنیت، و تولید گراف حمله را پیاده‌سازی می‌کند.

در [۱۴] چارچوبی برای تعیین وضعیت خدمت‌رسانی یک شبکه بعد از حملات منع خدمت توزیع‌شده روی آن، با تلفیق (ادغام) اطلاعات حسگرهای فنی و بشری ارائه شده است. در این تحقیق فرض شده است که حسگرهای فنی، تخمینی از وضعیت خدمت‌رسانی شبکه را در قالب پارامتر تأخیر زمانی در پاسخ به خدمات درخواستی فراهم می‌کنند، کاربران انسانی میزان رضایت خود از کیفیت خدمت‌دهی را اعلام می‌کنند، و شبکه‌ها و رسانه‌های خبری در مورد وضعیت اهداف مورد حمله اظهارنظرهایی با عناوین قطع خدمت، تحت تعمیر، و رد یا تکذیب ارائه می‌کنند. چارچوب ارائه‌شده در این مقاله با در نظر گرفتن خصیصه‌ها و پارامترهایی برای هرکدام از اطلاعات فراهم شده، ابتدا بر اساس جدول قوانین و وضعیت‌های مربوط به ترکیب خصیصه‌های هر کدام از حسگرها به تلفیق داده‌های آن‌ها به‌صورت مستقل می‌پردازد و سپس بر اساس یک وزن دهی به حسگرها، به تلفیق نتایج قبلی پرداخته و تخمین نهایی از وضعیت خدمت‌رسانی شبکه هدف حمله را به‌دست می‌آورد. با اینکه این تحقیق شروع خوبی برای استفاده از نظرات انسانی در ارزیابی وضعیت یک شبکه است، ولی پرداختن به یک پارامتر محدود از وضعیت شبکه برای یک حمله مشخص از پیچیدگی مسئله کاسته است و بنابراین، نیاز به استفاده از هستان‌شناسی برای مدل کردن مفاهیم مسئله و مرتبط‌سازی آن‌ها باهم وجود نداشته است.

کاستی اصلی مرتبط با موضوع این پژوهش، در تحقیقات پیشین تحلیل امنیت سایبری، به کار نگرفتن داده نرم در انواع داده‌های ورودی و پرداخت ناکافی به موضوع ادغام شواهد با عدم قطعیت‌های مختلف در فرآیند استنتاج و تخمین وضعیت سایبری است.

۳. مدل پیشنهادی

با توجه به ویژگی‌های خاص مسئله‌ی آگاهی وضعیتی سایبری از جمله وجود روابط معنایی گوناگون بین مفاهیم مختلف موجود در مسئله و همچنین پویایی مسئله، نیازمند ارائه یک چارچوب برای مدل‌سازی مسئله با امکان مرتبط‌سازی و ادغام داده‌های سخت و نرم هستیم. در چارچوب ارائه‌شده در این مقاله، نیازمندی‌های زیر در مسئله ادغام داده‌های سخت و نرم لحاظ شده است:

- مفاهیم مختلف موجود در فضای مسئله و روابط بین این مفاهیم در قالب یک ساختار انعطاف‌پذیر با امکان استدلال و استنتاج بر روی آن، مدل شود.
- نقش داده‌های نرم یعنی مشاهدات و استنتاج‌های انسانی با ویژگی‌های متفاوتی از داده‌های سخت از جمله فازی بودن یا کلی بودن، در فرآیند حل مسئله مشخص باشد.
- امکان نمایش و مدیریت انواع مختلف عدم قطعیت و نقص و ناکاملی اطلاعات را فراهم کند.
- امکان مرتبط‌سازی و ادغام داده‌های سخت و نرم را در فرآیند استنتاج فراهم کند.
- اطلاعاتی در مورد اعتبار و قابلیت اعتماد منابع داده‌ها و همچنین میزان باور تصریح‌شده توسط افراد نسبت به داده‌های ارائه‌شده خود، در فرآیند استنتاج و ادغام استفاده شود.

دانش دامنه موردنیاز برای آگاهی وضعیتی شامل دو دسته است:

۱) دانش در مورد اینکه چه کلاس‌ها یا اشیای با چه صفاتی در حوزه مسئله وجود دارد و چه روابطی ممکن است بین آن‌ها برقرار باشد.

۲) چه شرایطی باید بین اشیاء و صفات آن‌ها برای برقرار بودن یک وضعیت موردنظر، وجود داشته باشد.

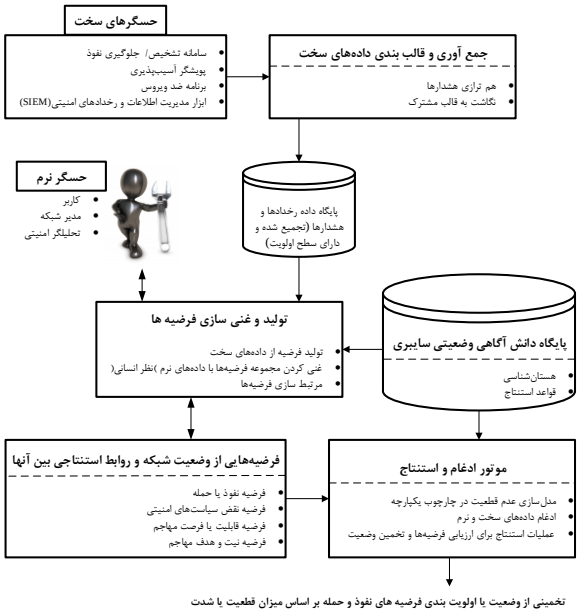
برای نیازمندی اول مدل هستان‌شناسی قابل استفاده است، و برای نیازمندی دوم، قواعد تعریف‌شده بر پایه مدل هستان‌شناسی، جواب‌گو است. بنابراین، برای نمایش دانش امنیت از قبیل دانش مربوط به آسیب‌پذیری‌ها، حملات و روابط بین آن‌ها، و همچنین استنتاج وضعیت امنیتی اجزای شبکه از روی داده‌ها می‌توان از هستان‌شناسی و قواعد استنتاج مبتنی بر آن استفاده کرد.

در فضای سایبری، انواع مختلفی از داده‌ها و اطلاعات توسط منابع متنوع ماشینی و انسانی تولید می‌شود که در تحلیل امنیت سایبری قابل استفاده است. این داده‌ها به موجودیت‌های مختلفی در فضای سایبری اشاره دارند و علاوه بر اینکه داده‌های مربوط به موجودیت‌های یکسان قابل همبسته‌سازی هستند، بر اساس وابستگی و ارتباط معنایی بین موجودیت‌های مختلف هم می‌توان به مرتبط سازی داده‌ها پرداخت. ارتباط معنایی بین موجودیت‌های فضای سایبری در قالب هستان‌شناسی قابل مدل کردن است.

از طرف دیگر داده‌های فضای سایبری دارای سطوح متفاوتی هستند. مثلاً یک هشدار خام IDS یک اطلاع سطح پایین از وضعیت امنیتی شبکه ارائه می‌دهد، ولی با پردازش داده‌های مختلف می‌توان اطلاعات سطح بالاتری مثل وضعیت به خطر افتادن یک میزبان، میزان قابلیت مهاجم، و پیامدهای حمله انجام‌شده را به‌دست آورد. ارتباط بین مفاهیم سطوح مختلف را هم می‌توان در هستان‌شناسی مدل کرد.

در این مقاله برای اشاره به انواع مختلف داده‌ها و اطلاعات مورد استفاده در فرآیند ادغام که شامل اطلاعات استنتاجی هم می‌شود، از اصطلاح فرضیه استفاده می‌شود. پس فرضیه سایبری، هر اطلاع سطح پایین یا سطح بالا از فضای امنیت سایبری را شامل می‌شود. مثال‌هایی از فرضیه سایبری عبارت‌اند از: وجود یک آسیب‌پذیری در یک سرویس، یک فعالیت مشکوک، یک اقدام نفوذ یا حمله در شبکه، یک تغییر در پیکربندی شبکه، بالا رفتن ترافیک یک سرور، کند شدن عملکرد یک میزبان، ظاهر شدن پیغام‌های غیرنرمال در صفحه یک کاربر، به خطر افتادن یک میزبان یا سرویس، تخمینی از قابلیت یا فرصت مهاجم، تخمینی از نیت مهاجم، و ...

معماری پیشنهادی برای ادغام داده‌های سخت و نرم به‌منظور به‌کارگیری در تحلیل امنیت سایبری مطابق شکل (۲) است. معماری پیشنهادی دارای مؤلفه‌های «جمع‌آوری و قالب‌بندی داده‌ها»، «هستان‌شناسی حوزه سایبری»، «پایگاه دانش تحلیل امنیت سایبری»، «پایگاه داده فرضیه‌های نفوذ و حمله»، و «موتور استنتاج و ادغام داده‌ها» است. در ادامه به تشریح مؤلفه‌های مختلف مدل ارائه‌شده می‌پردازیم.



شکل (۲): معماری پیشنهادی برای ادغام داده‌های سخت و نرم سایبری

مؤلفه جمع‌آوری و قالب‌بندی داده‌ها، به جمع‌آوری داده‌ها از منابع مختلف و ریختن آن‌ها در قالب استاندارد تعیین‌شده می‌پردازد.

داده‌های سخت ورودی قابل استفاده، شامل هر نوع داده یا اطلاعاتی است که توسط ابزارهای امنیتی به کار گرفته‌شده در شبکه تولید می‌شود. یکی از این ابزارها، سامانه مدیریت اطلاعات و رخدادهای امنیتی (SIEM[1]) است که با تجمیع هشدارها و اطلاعات از منابع مختلف، به تولید هشدارهای سطح بالا از حملات رخ داده یا وضعیت امنیتی اجزای شبکه می‌پردازد.

بر اساس هستان‌شناسی تعریف‌شده از حوزه‌ی امنیت سایبری، گزاره‌هایی از وضعیت شبکه و رخدادهای حمله ساخته می‌شود.

مؤلفه مرتبط‌سازی داده‌ها، بر اساس اطلاعات مشترکی که

[1] Security Information and Event Management

بین دو داده وجود دارد، آن‌ها را به یک شیء یکسان مرتبط می‌کند. مثلاً دو اطلاعی که آدرس مقصد و زمان یکسانی داشته باشند، باهم مرتبط هستند و می‌توانند در استنتاج یک اطلاع جدید استفاده شوند.

۳-۱. هستان‌شناسی تحلیل امنیت سایبری

هستان‌شناسی یک دید ساختارمند از مفاهیم موجود در یک حوزه و روابط معنایی بین مفاهیم را فراهم می‌کند. همچنین با استفاده از عناصر دیگری شامل پایگاه دانش، قواعد منطق یا استنتاج، و زبان‌های پرس‌وجو، که مبتنی بر هستان‌شناسی ساخته می‌شوند، می‌توان به استدلال و استنتاج روی داده‌ها و اطلاعات پرداخت.

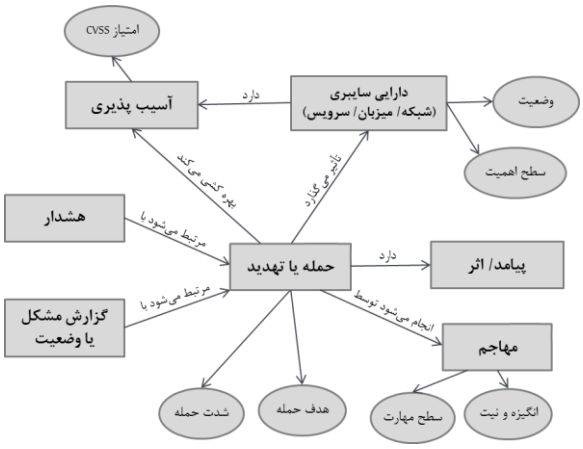
هستان‌شناسی بر اساس اهداف تعیین‌شده برای مسئله و دانش حوزه طراحی می‌شود. در این مقاله بر اساس این‌که هدف، به‌دست آوردن تخمینی از وضعیت امنیتی شبکه و میزبان‌ها یا سرویس‌های موجود در شبکه در نظر گرفته شده است، هستان‌شناسی مورد استفاده شامل مفاهیمی از قبیل پیکربندی شبکه، آسیب‌پذیری، حمله، هشدارهای ابزارهای امنیتی، و پیامدها یا اثرات حمله است.

نمونه‌هایی از دانش حوزه که نشان‌دهنده مفاهیم امنیت سایبری و روابط بین آن‌هاست و در طراحی هستان‌شناسی استفاده می‌شود، به‌صورت زیر است: یک آسیب‌پذیری موجود در یک مقصد، می‌تواند توسط یک حمله به کار گرفته شود که موجب به خطر افتادن مقصد و آسیب دیدن یک ویژگی امنیت از قبیل محرمانگی، یکپارچگی و دسترس‌پذیری منابع شود. یک رخداد (از قبیل نفوذ، حمله، و حادثه) مرتبط می‌شود با یک تهدید مفروض که از یک آسیب‌پذیری خاصی بهره‌برداری می‌کند. حسگرهای شبکه از قبیل سامانه تشخیص نفوذ، پویشگر آسیب‌پذیری، و دیواره آتش، رخدادهای مشکوک را با تجزیه و تحلیل اطلاعات بر اساس پیکربندی و آسیب‌پذیری‌های سیستم و شبکه کشف می‌کنند و بر اساس آن، هشدارهایی را تولید می‌کنند.

یکی از حالت‌ها یا انواع داده نرم مربوط به گزارش کاربران یا مدیران شبکه از مشکلات پیش‌آمده یا وضعیت کارکرد اجزای شبکه است. مشکل گزارش‌شده می‌تواند ناشی از یک حمله یا فعالیت مخرب انجام‌شده در شبکه باشد که در این صورت این گزارش می‌تواند شواهدی را برای وقوع حمله فراهم کند. بر این اساس در هستان‌شناسی، کلاسی با عنوان «گزارش مشکل یا وضعیت» در نظر گرفته شده است که با کلاس حمله در ارتباط است.

نوع دیگر داده نرم، اظهارنظر تحلیل‌گر امنیتی نسبت به اطلاعات به‌دست‌آمده از داده‌های سخت است. مثلاً تحلیل‌گر با مشاهده خروجی‌های ابزارهای مختلف از جمله گزارش‌های تحلیلی SIEM، نشانه‌هایی (قطعی یا غیرقطعی) از وقوع یک حمله (با اطلاعات مشخص) را به‌دست آورده و بر اساس دانش و تجربه خود، نسبت به درست بودن تشخیص حمله، و یا ویژگی‌هایی از حمله مثل هدف حمله یا اثر حمله، نظر خود را به سیستم ارائه می‌کند. در این حالت، مفهوم مورد اشاره داده نرم در هستان‌شناسی وجود دارد.

بر اساس توضیحات بالا، هستان‌شناسی تحلیل امنیت سایبری قابل استفاده در ادغام داده‌های سخت و نرم، مطابق شکل (۳) طراحی شده است. این هستان‌شناسی شامل کلاس‌های «دارایی سایبری»، آسیب‌پذیری، هشدار، «گزارش مشکل یا وضعیت»، «حمله یا تهدید»، «پیامد یا اثر حمله» و مهاجم است. هر کدام از این کلاس‌ها دارایی خصیصه‌هایی هستند، مثلاً کلاس حمله دارای خصیصه‌های شدت حمله و هدف حمله است و کلاس مهاجم دارای خصیصه‌های «انگیزه و نیت» و «سطح مهارت» است.



شکل (۳): هستان‌شناسی تحلیل امنیت سایبری

گزاره‌های ورودی یا استنتاجی از فضای سایبری بر اساس این هستان‌شناسی ارائه می‌شود که شامل موارد زیر می‌تواند باشد:

- وجود نمونه‌ای از یک کلاس مدل‌شده در هستان‌شناسی، مثلاً وجود یک میزبان مشخص (h1) در شبکه به‌عنوان نمونه‌ای از دارایی سایبری که با گزاره Asset(?h1) قابل بیان است.
- مقداری برای یک خصیصه از کلاس مشخص، مثلاً مشخص کردن سطح مهارت پیشرفته برای مهاجم a1 که با گزاره skill_level(?a1, "advanced") قابل بیان است.

- وجود نمونه‌ای از روابط بین کلاس‌ها مدل شده در هستان‌شناسی، مثلاً بهره‌کشی یک حمله خاص (at1) از یک آسیب‌پذیری مشخص (v1) که با گزاره exploit(?at1, ?v1) بیان می‌شود.

۳-۲. قواعد استنتاج

برای اینکه از داده‌ها و اطلاعات موجود، به نتایج موردنیاز مسئله برسیم، در کنار هستان‌شناسی، نیاز به قواعد استنتاج داریم. یک قاعده استنتاج شامل بدنه و نتیجه است که بدنه قاعده شامل گزاره‌های مرتبط با داده‌ها و گزاره‌های استنتاج شده‌ی مراحل قبلی است که با عملگرهای عطفی و فصلی باهم ترکیب می‌شوند و نتیجه قاعده، گزاره استنتاجی جدید است. در ادامه بر اساس چند مثال از فضای سایبری، نحوه به‌کارگیری قواعد استنتاج مبتنی بر هستان‌شناسی برای استنتاج گزاره‌های جدید نشان داده می‌شود.

قاعده زیر مربوط بودن دو هشدار به مهاجم یکسان را در قالب یک گزاره جدید، نتیجه می‌گیرد.

$$Alert(?x) \wedge Alert(?y) \wedge Attacker(?a) \wedge hasSource(?x,?a) \wedge hasSource(?y,?a) \rightarrow hasSameAttacker(?x,?y)$$

قاعده زیر بر اساس انگیزه و سطح مهارت مهاجم، هدف و شدت حمله را استنتاج می‌کند.

$$Attack(?a) \wedge Attacker(?b) \wedge RunAttack(?b, ?a) \wedge motivation(?b, "political\text{-}gain") \wedge skill\_level(?b, "advanced") \rightarrow attack\_goal(?a, "destroy") \wedge attack\_severity(?a, "high");$$

قاعده زیر پیش‌نیازها و پیامدهای یک حمله را بر اساس وجود نمونه‌هایی از مفاهیم هستان‌شناسی و روابط بین آن‌ها مشخص می‌کند.

$$Device(?a) \wedge Component(?b) \wedge Vulnerability(?c) \wedge sameAs(?c, CVE\text{-}2013\text{-}2251) \wedge has(?a, ?b) \wedge has(?b, ?c) \wedge AttackCodeExec(?d) \wedge exploit(?d, ?c) \wedge Adversary(?e) \wedge hasAccess(?e, ?b) \rightarrow launchAttackCodeExec\_Ins\_CVE\text{-}2013\text{-}2251(?e, ?a) \wedge hasCompromiseFiles(?e, ?b)$$

در فرآیند استنتاج ممکن است یک گزاره از طریق شواهد و قواعد مختلفی نتیجه گرفته شود که در این صورت لازم خواهد بود نتایج باهم ادغام شده و یک باور و قطعیت سراسری برای آن گزاره (فرضیه) محاسبه شود. همچنین امکان نوشتن تابعی برای ادغام داده‌ها و فراخوانی تابع در داخل قاعده وجود دارد. این نکات در مثال بخش چهار مقاله نشان داده شده است.

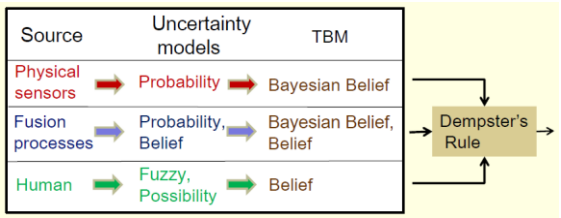
۳-۳. ادغام داده‌های سخت و نرم

زیر مسئله ادغام داده‌های سخت و نرم، به این صورت در نظر گرفته شده است که هر کدام از منابع داده سخت و نرم موجود، اطلاعاتی را در مورد متغیرهای مختلفی که در حل مسئله تعریف می‌شود، ارائه می‌کنند. در این مقاله فرض می‌کنیم که از قبل متغیرهای خاصی از مسئله انتخاب شده و با پردازش داده‌های سخت، مقادیر قابل انتساب به این متغیرها بر اساس داده‌های سخت، مشخص می‌شوند. همچنین از طریق فرم‌هایی که در اختیار افراد خبره و تحلیل گران امنیتی قرار می‌گیرد، مقادیر داده‌های نرم قابل انتساب به هر کدام از متغیرهای مسئله به‌دست می‌آیند.

نکته حائز اهمیت در این زمینه این است که در کنار داده‌های سخت، مقداری به‌عنوان قابلیت اطمینان یا احتمال درستی مربوط به هر داده وجود دارد و در کنار داده‌های نرم، مقداری به‌عنوان باور یا میزان قطعیت مربوط به هر داده ارائه می‌شود.

در ادغام داده‌های سخت و نرم ما با این مسئله مواجه هستیم که باید اطلاعات با انواع مختلفی از عدم قطعیت را با هم ادغام کنیم. اطلاعات فراهم‌شده توسط حسگرها عموماً دارای عدم قطعیت احتمالاتی هستند، در حالی که اطلاعات زبانی نوعاً دارای عدم قطعیت امکانی هستند. با توجه به این تفاوت در نوع عدم قطعیت داده‌های سخت و نرم، برای ادغام داده‌ها با هم، نیاز به یک چارچوب نظری برای نمایش و مدل‌سازی عدم قطعیت به‌صورت یکپارچه وجود دارد. مدل باور انتقال‌پذیر (TBM[1]) با توجه به ویژگی‌ها و مزایایی که دارد، گزینه مناسبی برای این هدف است.

بر اساس یک رویکرد مبتنی بر مدل باور انتقال‌پذیر، مطابق شکل (۴)، می‌توانیم انواع مختلف نمایش‌های عدم قطعیت شامل نمایش احتمالی، نمایش فازی و نمایش امکانی را به یک نمایش یکسان بر اساس نظریه باور تبدیل کنیم. در این صورت می‌توان با به‌کارگیری قوانین ترکیب باور معرفی شده مانند قانون ترکیب دمپستر، به ادغام شواهد به‌دست آمده از منابع مختلف اقدام کرد.



شکل (۴): نحوه استفاده از مدل باور انتقال‌پذیر در ادغام داده‌های سخت و نرم با نمایش‌های مختلف عدم قطعیت

[1] Transferable Belief Model

در ادامه شرح مختصری از نظریه مدل باور انتقال‌پذیر و قاعده ترکیب دمپستر آورده می‌شود. مدل باور انتقال‌پذیر [۱۰] توسعه‌ای از نظریه شواهد دمپستر-شفر است. این مدل، در واقع یک مدل دو سطحی است که در سطح اول به مدل کردن درجه باور منابع مختلف نسبت به فرضیه‌های مورد بررسی و ترکیب این باورها می‌پردازد و در سطح دوم یک انتقال از مقادیر باور به مقادیر احتمال قابل استفاده در فرآیند تصمیم‌گیری صورت می‌گیرد.

با فرض این‌که $m_1(.)$ و $m_2(.)$ توابع جرم متناظر با مجموعه شواهد دو منبع مستقل از هم در چارچوب تشخیص Θ باشند، بر اساس قاعده ترکیب دمپستر، تابع جرم ترکیبی برای هر زیرمجموعه C از مجموعه Θ به‌صورت زیر به‌دست می‌آید:

$$m_{1,2}(C) = [\mathrm{m}_1 \oplus \mathrm{m}_2](\mathrm{C}) = \frac{\sum_{A\cap B=C} m_1(A)m_2(B)}{1-\sum_{A\cap B=\phi} m_1(A)m_2(B)} \quad \mathrm{C} \neq \phi \qquad (۱)$$

برای داده سخت که احتمالاتی است، تابع جرم به هر پیشامد (یک زیرمجموعه از چارچوب تشخیص) که احتمال غیر صفر دارد، به همان مقدار احتمال اختصاص می‌یابد. ولی در مورد داده نرم حالت‌های مختلف ممکن است. در یک حالت، مقادیر باور فرد نسبت به پیشامدهای مختلف ممکن در مسئله، با استفاده از فرم‌های طراحی‌شده در سامانه گرفته می‌شود. در برخی حالت‌ها هم برای نگاشت داده نرم به تابع جرم، نیاز به دانش خاصی در قالب قواعد یا جدول نگاشت داریم.

در مورد مدل و روش پیشنهادی، به‌صورت خلاصه و بر اساس جدول (۱) می‌توان گفت که مسئله مورد بررسی در این پژوهش، تحلیل امنیت سایبری بوده که در تحقیقات قبلی ادغام داده‌های سخت و نرم به آن پرداخته نشده است؛ معماری پیشنهادی برای پردازش و ادغام اطلاعات، مبتنی بر هستان‌شناسی است، به این صورت که گزاره‌های ورودی و فرضیه‌های مربوط به وضعیت بر اساس هستان‌شناسیِ طراحی‌شده بیان شده و فرآیند استنتاج و ادغام نیز بر این اساس طراحی می‌شود؛ چارچوب نظری مورد استفاده برای مدل‌سازی عدم قطعیت و ادغام، مدل باور انتقال‌پذیر و قاعده ترکیب دمپستر می‌باشد؛ و نحوه نمایش و مدل‌سازی داده‌ها نیز، همان‌طور که اشاره شد، در قالب نمونه‌هایی از مفاهیم و روابط مدل شده در هستان‌شناسی و احتمال یا باور متناظر با آن خواهد بود.

۴. به‌کارگیری مدل پیشنهادی در یک سناریو

مثال زیر توصیفی از یک سناریوی نصب بدافزار و حمله منع سرویس را بیان می‌کند. در یک پیکربندی فرضی، در نرم‌افزار IE ماشین میزبان، آسیب‌پذیری CA-2003-22 وجود دارد. مهاجم با بهره‌کشی از این آسیب‌پذیری می‌تواند سطح دسترسی ریشه کامپیوتر میزبان را به‌دست آورده و اقدام به نصب بدافزار GT_Bot کند. میزبان‌هایی که این بدافزار روی آن‌ها نصب شده، شروع به ارسال ترافیک DoS به سروری بر روی شبکه مقصد می‌کنند. سامانه تشخیص نفوذ می‌تواند ارسال ترافیک DoS را تشخیص دهد. از طرفی ابزار امنیتی نصب‌شده روی ماشین میزبان، نصب نرم‌افزار با تولیدکننده ناشناخته و نامعتبر را گزارش کرده است (گزارش نصب از منبع نامشخص). همچنین گزارشی از طرف کاربر مبنی بر «کاهش سرعت سیستم به میزان زیاد» ارائه می‌شود (داده نرم). هدف ما به‌دست آوردن میزان قطعیت نسبت به فرضیه «به خطر افتادن ماشین میزبان» بر اساس داده‌ها و گزارش‌های دریافتی است.

اطلاعات زیر در پایگاه دانش مبتنی بر هستان‌شناسی موجود است:

Vulnerability("CA-2003-22")

Attack("GT-Bot installation")

Exploit("GT-Bot installation","CA-2003-22")

subtype("GT-Bot installation","Malware installation")

همچنین قواعد استنتاج زیر در پایگاه دانش وجود دارد:

Rule1:

Asset(?h1) & Vulnerability(?v) & hasVuln(?h1,?v) & Attack(?a) & exploit(?a,?v) → underAttack(?h1,?a);

Rule2:

underAttack(?h1, "GT-Bot installation") & Alert("sending DoS traffic",?h1) → Compromised(?h1, 0.8);

Rule3:

underAttack(?h1, "Malware installation") & Alert("نصـــــــب از منبـــــــع نامشـــــــخص",?h1) → Compromised(?h1,0.6);

Rule4:

Compromised(?h1,p) & ProblemReport(?h1, "SpeedDown", v) → Compromised(?h1, Fusion(Mass(p), Mass(v)));

با به‌کارگیری قواعد استنتاج بالا روی پیکربندی فرض شده، نتیجه می‌شود که شواهدی برای به خطر افتادن ماشین میزبان (host1) وجود دارد. با ادغام این شواهد می‌توان به نتیجه مطمئن‌تری دست یافت. قاعده۱ بر اساس وجود آسیب‌پذیری در پیکربندی ماشین میزبان، امکان حمله نصب GT-Bot بر روی

ماشین میزبان را نتیجه می‌دهد. قواعد ۲ و ۳ بر اساس هشدارهایی از حسگرهای مختلف، به خطر افتادن میزبان را با احتمال‌های ۰/۸ و ۰/۶ نتیجه می‌دهد که با توجه به یکسان بودن گزاره نتیجه شده، لازم است مقادیر احتمال و قطعیت آن‌ها باهم ادغام شود. قاعده۴ بر اساس فرضیه به خطر افتادن میزبان و یک داده نرم مرتبط با آن، میزان قطعیت فرضیه را تغییر می‌دهد که این تغییر با ادغام نتیجه به‌دست‌آمده از مرحله قبل و مقدار متناظر با داده نرم انجام می‌شود. مراحل ادغام موردنیاز با استفاده از مدل باور انتقال‌پذیر و قاعده ترکیب دمپستر-شفر به‌صورت زیر خواهد بود.

ابتدا چارچوب تشخیص و توابع جرم مربوط به شواهد را مشخص می‌کنیم و سپس با ادغام شواهد با استفاده از رابطه (۱)، تابع جرم ترکیبی را به‌صورت زیر به‌دست می‌آوریم.

چارچوب تشخیص: به خطر افتادن ماشین میزبان (comp)، سالم بودن ماشین میزبان (helth)

$$\mathrm{oD} = \{\mathrm{Compromised}, \mathrm{Healthy}\} = \{\mathrm{comp}, \mathrm{helth}\}$$

$$m_1(\mathrm{comp}) = 0.8\ , \quad m_1(helth) = 0.2$$

$$m_2(\mathrm{comp}) = 0.6\ , \ m_2(helth) = 0.4$$

$$m_{1,2}(\mathrm{comp}) = \frac{m_1(\mathrm{comp})m_2(\mathrm{comp})}{1 - (m_1(\mathrm{comp})m_2(\mathrm{helth}) + m_1(\mathrm{helth})m_2(\mathrm{comp}))}$$

$$= \frac{0.8 * 0.6}{1 - (0.8 * 0.4 + 0.2 * 0.6)} = \frac{0.48}{0.56} = 0.86$$

$$m_{1,2}(\mathrm{helth}) = \frac{0.2 * 0.4}{1 - (0.8 * 0.4 + 0.2 * 0.6)} = \frac{0.08}{0.56} = 0.14$$

برای ادغام داده نرم با نتیجه قبلی، ابتدا لازم است که تابع جرم متناظر با داده نرم (یعنی Mass(v) در قاعده۴) تشکیل شود. برای این کار می‌توان از تحلیل یک فرد خبره یا دانش ذخیره‌شده در پایگاه دانش استفاده کرد که بر اساس آن یک نگاشت از مشکل یا وضعیت گزارش شده و مقدار همراه با آن به جرم اختصاصی پیشامدهای ممکن صورت می‌گیرد.

در این سناریو، با توجه میزان تأثیر مشکل کاهش سرعت در احتمال وقوع حمله و میزان نامعلومی یا عدم قطعیت زیاد در فضای مسئله، نتیجه‌گیری فرد خبره برای توابع جرم به این صورت است:

$$m_3(\mathrm{comp}) = 0.4\ , \ m_3(\Theta) = 0.6$$

ادغام این تابع جرم با نتایج قبلی به‌صورت زیر خواهد بود:

$$m_{1,2,3}(\mathrm{comp}) = \frac{m_{1,2}(\mathrm{comp})m_3(\mathrm{comp}) + m_{1,2}(\mathrm{comp})m_3(\Theta)}{1 - \left(m_{1,2}(\mathrm{helth})m_3(\mathrm{comp})\right)}$$

$$= \frac{0.86 * 0.4 + 0.86 * 0.6}{1 - (0.14 * 0.4)} = \frac{0.86}{0.944} = 0.91$$

$$m_{1,2,3}(\mathrm{helth}) = \frac{m_{1,2}(\mathrm{helth})m_3(\Theta)}{1 - \left(m_{1,2}(\mathrm{helth})m_3(\mathrm{comp})\right)}$$

$$= \frac{0.14 * 0.6}{1 - (0.14 * 0.4)} = \frac{0.084}{0.944} = 0.09$$

مشاهده می‌شود که گزارش کاربر مبنی بر کاهش سرعت سیستم، احتمال به خطر افتادن ماشین میزبان را بالاتر می‌برد.

۵. نتیجه‌گیری

در این مقاله مدل جدیدی جهت ادغام داده‌های سخت یعنی داده‌های به‌دست‌آمده از حسگرهای ماشینی و داده‌های نرم یعنی مشاهدات و استنتاج‌های انسانی در کاربرد تحلیل امنیت سایبری ارائه شد. در مدل پیشنهادی از هستان‌شناسی همراه با قواعد استنتاج مبتنی بر آن، برای مدل کردن مفاهیم و متغیرهای مسئله و استنتاج وضعیت بر اساس داده‌های دریافتی استفاده شده است. همچنین مدل باور انتقال‌پذیر و قاعده ترکیب دمپستر-شفر برای مدل‌سازی یکپارچه عدم قطعیت داده‌ها و ادغام داده‌ها و شواهد مختلف به کار گرفته شده است.

در هستان‌شناسی تحلیل امنیت سایبری، مفاهیمی از قبیل دارایی‌ها (شامل شبکه، میزبان و سرویس)، آسیب‌پذیری، حمله، رویداد (هشدارهای ابزارهای امنیتی) و گزارش مشکل یا وضعیت در نظر گرفته شد. ساختن گزاره‌ها از داده‌های ورودی و طراحی قواعد منطق مورد استفاده در استنتاج، مبتنی بر هستان‌شناسی مدل شده انجام می‌شود. با به‌کارگیری مدل پیشنهادی در یک سناریوی نمونه از تحلیل امنیت سایبری، نشان داده شد که روش ارائه‌شده، عملیاتی بوده و کارآمدی لازم برای ادغام داده‌های سخت و نرم سایبری را داراست.

کارهای آتی این تحقیق شامل استفاده از چارچوب‌های نظری دیگر برای مدل‌سازی عدم قطعیت و ادغام داده‌ها (به‌طور مثال نظریه اندازه فازی و عملگرهای تجمیع مبتنی بر آن)، توسعه مدل با تکمیل هستان‌شناسی و قواعد استنتاج، و به‌کارگیری مدل در مثال‌های بزرگ‌تر و واقعی‌تر می‌باشد.

۶. مراجع

[1] D. L. Hall, M. D. McNeese, J. Llinas, and T. Mullen, "A framework for dynamic hard/soft fusion," In FUSION, pp. 1-8, 2008.

[2] J. Llinas, N. Rakesh, D. Hall, and J. Lavery, "A multi-disciplinary university research initiative in hard and soft information fusion: Overview, research strategies and initial results," In Information Fusion (FUSION), 2010 13th Conference on, pp. 1-7. IEEE, 2010.

[3] G. Gross, R. Nagi, and K. Sambhoos, "Soft information, dirty graphs and uncertainty representation/processing for situation understanding," In Proceedings of the 13th International Conference on Information Fusion, Edinburgh, UK, 2010.

[4] M. P. Jenkins, G. A. Gross, A. M. Bisantz, and N. Rakesh, "Towards context aware data fusion: Modeling and integration of situationally qualified human observations to manage uncertainty in a hard+ soft fusion process," Information Fusion, vol. 21, pp. 130-144, 2015.

[5] B. Khaleghi, "Distributed Random Set Theoretic Soft/Hard Data Fusion," PhD diss. University of Waterloo, 2012.

[6] T. L. Wickramarathne, "An Analytical Framework for Soft and Hard Data Fusion: A Dempster-Shafer Belief Theoretic Approach," PhD diss. Miami Univ. Coral Gablesfl, 2012.

[7] K. Golestan, F. Karray, and M. S. Kamel, "An integrated approach for fuzzy multi-entity bayesian networks and semantic analysis for soft and hard data fusion," In Fuzzy Systems (FUZZ-IEEE), 2014 IEEE International Conference on, 2015.

[8] Xu, Guangquan, Yan Cao, Yuanyuan Ren, Xiaohong Li, and Zhiyong Feng, "Network security situation awareness based on semantic ontology and user-defined rules for Internet of Things,"IEEE Access 5, pp. 21046-21056, 2017.

[9] G. Gonzalez Granadillo, M. Yosra Ben, H. Nabil, and D. Herve, "An ontology-driven approach to model SIEM information and operations using the SWRL formalism," International Journal of Electronic Security and Digital Forensics, vol. 74, no. 2-3, pp. 104-123, 2012.

[10] P. Smets, "Data fusion in the transferable belief model," Information Fusion, FUSIon 2000, Proceedings of the Third International Conference on, vol. 1, IEEE, 2000.

[11] Gao, Jian-bo, Bao-wen Zhang, Xiao-hua Chen, and Zheng Luo, "Ontology-based model of network and computer attacks for security assessment," Journal of Shanghai Jiaotong University (Science), vol. 18, no. 5, pp. 554-562, 2013.

[12] Wu, Songyang, Yong Zhang, and Wei Cao. "Network security assessment using a semantic reasoning and graph based approach," Computers & Electrical Engineering, vol. 64, pp. 96-109, 2017.

[13] S. Barnum, "Standardizing cyber threat intelligence information with the Structured Threat Information expression (STIX)," MITRE Corporation, vol. 11, pp. 1-22, 2012.

[14] V. Akbari and S. M. Safavi Homami, "A Framework For The Status Estimation In Distributed Denial-Of-Service Attacks By Data Fusion Of Human-And-Technical Sensors Based On Fuzzy Logic," Journal of Electronical & Cyber Defence, vol. 5, no. 3, Serial No. 19, IHU. AC. IR, 2017. (In Persian)

Journal of Electronical & Cyber Defence
Vol. 7, No. 4, 2020, Serial No. 28

A New Method for Image Steganography Using Discrete Wavelet Transforms

A. J. Rashidi*, S. Sobhani, M. Hoseini

*Malek Ashtar University of Technology

(Received: 01/12/2018, Accepted: 05/03/2019)

ABSTRACT

In Cyber Security Analysis, in addition to data and information obtained from machine-based sensors like intrusion detection systems, firewalls and vulnerability scanners (hard data), human observations and conclusions from world's state including problems reported by users and network administrators, and assessments made by security analysts about network security status (soft data), can be used to obtain more accurate and more reliable estimation and decision. Hard and soft data fusion in cyber security analysis has many challenges such as framework design for problem modeling and representation of different types of uncertainty. This paper presents a new model based on ontology for fusion of hard and soft data in cyber security analysis. First, the concepts and problem variables are modeled and then the inference of assets' security status is made by using a set of rules. Also, for fusion of data and unified modeling of different uncertainties, transferable belief model (TBM) and Dempster-Shafer combination rule are used. Results of applying the proposed model in a sample scenario of cyber security analysis show its operability for hard and soft data fusion. Considering the extensibility of ontology and knowledge base, high flexibility and dynamism are characteristics of the proposed model.

Keywords: Hard and soft data fusion, Uncertainty modeling, Cyber security analysis, Ontology.

* Corresponding Author Email: rashidi@mut.ac.ir