فصلنامه علمی - پژوهشی فرماندهی و کنترل، سال پنجم، شماره یک، بهار 1400

پردازش و ادغام اطلاعات سخت و نرم در فرماندهی و کنترل سایبری

علی جبار رشیدی[1]، سعداله سبحانی[2]

تاریخ دریافت: ۱۴۰۰/۰۵/۰۴ **تاریخ پذیرش: ۱۴۰۰/۰۷/۲۰**

چکیده

کاربردهای متنوع و تاثیرات گسترده فضای سایبر در اکثر حوزه‌های کاربردی نظامی و غیرنظامی، باعث رشد سریعِ داده‌ها، اطلاعات، دانش، فناوری، روش‌ها، ابزارها و سامانه‌های سایبری شده است. یکی از نیازهای مهم و راهبردی در چرخه فرماندهی و کنترل در حوزه سایبری، قابلیت استخراج، پردازش، ادغام و تحلیل داده‌ها و اطلاعات از منابع گوناگون برای رسیدن به آگاهی وضعیتی مطلوب از محیط عملیات سایبری می‌باشد. با توجه به جنبه‌های گوناگون این مسئله، استفاده از داده‌های نرم یعنی داده‌ها و اطلاعات قابل ارائه توسط منابع انسانی در کنار داده‌های سخت یعنی داده‌ها و اطلاعات منابع ماشینی، می‌تواند در رسیدن به تشخیص و تصمیم دقیق‌تر و مطمئن‌تر کمک کند. از اصلی ترین موضوعات تحقیقاتی در این مسئله، طراحی مدل مفهومی و فرآیند پردازشی مناسب برای تحلیل و استنتاج مبتنی بر داده و اطلاعات با امکان مرتبط‌سازی اطلاعات متنوع از منابع گوناگون با یکدیگر و مدل‌سازی انواع عدم قطعیت در داده‌های سخت و نرم و ادغام این داده‌ها است.

در این مقاله یک رویکرد مبتنی بر هستان‌شناسی برای پردازش و ادغام داده‌های سخت و نرم در فرماندهی و کنترل سایبری ارائه شده است که در آن به جنبه‌های مختلف این مسئله شامل معماری و مدل فرآیندی پردازش و استنتاج اطلاعات مبتنی بر هستان‌شناسی، روش بازنمایی عدم قطعیت و قابلیت اعتماد در داده‌های سخت و نرم و ادغام این داده‌ها، تبدیل باورها به احتمالات برای امکان تصمیم‌گیری روی فرضیه‌های مورد بررسی، طراحی مدل هستان‌شناسی برای اهداف فرماندهی و کنترل سایبری، و طراحی و پیاده‌سازی منطق استنتاج و ادغام اطلاعات مبتنی بر هستان‌شناسی پرداخته شده است. نتایج به کارگیری مدل پیشنهادی در یک سناریوی نمونه از فرماندهی و کنترل سایبری، عملیاتی بودن آن را در ادغام داده‌های سخت و نرم سایبری نشان می‌دهد. علاوه بر قابلیت مناسب برای استنتاج و ادغام، یکی از ویژگی‌های قابل توجه رویکرد پیشنهادی، قابلیت توسعه و مقیاس‌پذیری آن برای تطبیق با گسترش‌های جدید در ابزارها و نیازمندی‌های فضای فرماندهی و کنترل سایبری است.

واژگان کلیدی: ادغام اطلاعات سخت و نرم، استنتاج مبتنی بر قاعده، عدم قطعیت، فرماندهی و کنترل سایبری، هستان‌شناسی.

[1] عضو هیأت علمی (دانشگاه صنعتی مالک اشتر، دانشیار) rashidi@mut.ac.ir (نویسنده مسئول)

[2] دانشجوی دکتری (دانشگاه صنعتی مالک اشتر) ssobhani@gmail.com

1- مقدمه

کاربردهای متنوع و تاثیرات گسترده فضای سایبر در اکثر حوزه‌های کاربردی نظامی و غیرنظامی، باعث رشد سریعِ داده‌ها، اطلاعات، دانش، فناوری، روش‌ها، ابزارها و سامانه‌های سایبری شده است. یکی از نیازهای مهم و راهبردی در چرخه فرماندهی و کنترل در حوزه سایبری، قابلیت استخراج، پردازش، ادغام و تحلیل داده‌ها و اطلاعات از منابع گوناگون برای رسیدن به آگاهی وضعیتی مطلوب از محیط عملیات سایبری می‌باشد. به محض آشکار شدن یک هشدار حمله یا نفوذ توسط حسگرهای امنیتی، فرمانده یا تحلیلگر سایبری باید بتواند به اطلاعات و دانش مختلفی از آن حمله یا نفوذ، از قبیل ضعف‌ها و آسیب‌پذیری‌های مورد استفاده توسط حمله، پیامدهای احتمالی حمله، و روش‌های مقابله یا کم اثر کردن حمله دسترسی پیدا کند.

با توجه به جنبه‌های مختلف این مسئله، در کنار استفاده از داده‌ها و اطلاعات منابع ماشینی، داده‌ها و اطلاعات قابل ارائه توسط منابع انسانی می‌تواند در رسیدن به تشخیص و تصمیم دقیق‌تر و مطمئن‌تر کمک کند. این منابع انسانی شامل کاربران عادی، کارشناسان و تحلیل‌گرهای امنیتی و فرماندهان صحنه سایبری می‌تواند باشد.

انواع داده‌ها و اطلاعات ورودی قابل استفاده در سامانه‌های ادغام[2] اطلاعات به دو دسته کلی داده‌های سخت و داده‌های نرم[3] تقسیم بندی می‌شود. داده‌های سخت به داده‌ها و اطلاعات بدست آمده از حسگرهای ماشینی و داده‌های نرم به داده‌ها و اطلاعات بدست آمده از منابع انسانی گفته می‌شود [2]. در حوزه فضای فرماندهی و کنترل سایبری، داده‌های نرم شامل موارد کلی زیر می‌تواند باشد:

- گزارش کاربران یا مدیران شبکه از فعالیت‌ها و رخدادهای مشکوک و یا وضعیت عادی یا غیرعادی دارایی‌های سایبری در فضای تحت دسترسی خود
- اظهارنظر یک کارشناس یا تحلیل‌گر نسبت به متغیرهایی از فضای سایبری مورد بررسی
- اظهارنظر یک کارشناس یا تحلیل‌گر نسبت به قابلیت اعتماد یک منبع داده یا اعتبار نتایج یک پردازش

داده‌های نرم می‌تواند اطلاعات بدست 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 بدست 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.

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. با توجه به این تفاوت‌ها ادغام داده‌های نرم با داده‌های سخت دارای مشکلات و چالش‌های خاصی هم در سطح معماری و هم در سطح تکنیک‌ها و الگوریتم‌ها است. از جمله اینکه با توجه به تنوع منابع اطلاعاتی و گوناگونی دانش و اطلاعات مورد استفاده، لازم است چارچوب و فرآیندی برای استنتاج و ادغام اطلاعات طراحی و پیاده سازی شود که بتواند بستر مورد نیاز برای مرتبط سازی اقلام اطلاعاتی مختلف با یکدیگر و مدل‌سازی منطق استنتاج لازم برای پاسخگویی به پرسش‌ها و نیازهای اطلاعاتی تحلیل‌گر یا فرمانده سایبری را فراهم کند.

در کار قبلی از همین نویسندگان [1]، به ارائه یک چارچوب مبتنی بر هستان‌شناسی برای ادغام داده‌های سخت و نرم در تحلیل امنیت سایبری پرداخته شده است که در آن از هستان شناسی همراه با قواعد استنتاج مبتنی بر آن، برای مدل کردن مفاهیم و متغیرهای مسئله و استنتاج وضعیت بر اساس داده‌های دریافتی استفاده شده است. در این مقاله با تکمیل و توسعه کار قبلی برای کاربرد فرماندهی و کنترل سایبری، به موضوعات و چالش‌های مختلف مسئله ادغام داده‌های سخت و نرم پرداخته

[2]Fusion

[3]Hard data and Soft data

فصلنامهٔ علمی - پژوهشی فرماندهی و کنترل، سال پنجم، شمارهٔ یک، بهار 1400

شده است که شامل موارد زیر می‌باشد:

- ارائه معماری و مدل فرآیندی مبتنی بر هستان‌شناسی برای پردازش، استنتاج و ادغام اطلاعات
- بازنمایی یکپارچه انواع متنوع عدم قطعیت در داده‌های سخت و نرم و ادغام این داده‌ها
- در نظر گرفتن قابلیت اعتماد منابع داده‌ها در فرآیند ادغام
- تبدیل باورها به احتمالات برای امکان تصمیم‌گیری روی فرضیه‌های مورد بررسی
- طراحی مدل هستان‌شناسی برای اهداف فرماندهی و کنترل سایبری
- طراحی و پیاده‌سازی منطق استنتاج و ادغام اطلاعات برای پاسخگویی به پرسش‌ها و نیازهای اطلاعاتی تحلیل‌گر یا فرمانده سایبری.

2- کلیات

2-1. سابقه تحقیقات

در تحقیقات مختلفی از جمله [3]، [4]، [5]، [6] و [7] ، به اهمیت استفاده از داده‌های نرم در کنار داده‌های سخت، چالش‌های ادغام این داده‌ها و ارائه راهکار در این زمینه پرداخته شده است. در [1] با بررسی این تحقیقات، به مقایسه آنها از جنبه‌های مختلف از جمله مسئله مورد بررسی، معماری طراحی شده برای پردازش و ادغام اطلاعات و چارچوب نظری انتخابی برای مدل‌سازی عدم قطعیت و ادغام داده‌ها پرداخته شده است. در تحقیقات اخیر نیز موضوع بهره‌گیری از داده‌های سخت و نرم در کنار هم و ادغام و یکپارچه‌سازی آنها در کاربردهای مختلفی مورد توجه قرار گرفته است. از جمله در [16] به ادغام داده‌ها از منابع مختلف سخت و نرم در مدیریت شرایط اضطراری با هدف بهبود تصمیم‌گیری‌های مدیریتی و بهره‌وری منابع پرداخته شده است. در [17] ادغام تصمیمات انسان و ماشین در شناسایی بیومتریک حساس به هزینه مورد بررسی قرار گرفته و برای مسئله شناسایی امضاء دستی روی چک‌های بانکی با هدف کمینه کردن هزینه‌های تشخیص اشتباه، از نیروی خبره انسانی در کنار یک ماشین تولید کننده امتیاز بهره گرفته شده است. برای ادغام تصمیم‌ها از روش ترکیب طبقه‌بندها مبتنی بر رأی اکثریت و عملگرهای منطقی ترکیب استفاده شده است. در [18] یک چارچوب استدلالی مبتنی بر نظریه دمپستر-شفر برای ادغام داده‌ها و تصمیم‌گیری در حل مسئله بهره‌وری انرژی در بناهای تاریخی ارائه شده است. مدل ارائه شده، داده‌های سخت از قبیل مدل‌های انرژی و پیش‌بینی های اقتصادی و داده‌های نرم شامل نظرات انسانی در بررسی فاکتورهای سیاسی-اقتصادی را برای ارزیابی گزینه‌های تصمیم به کار گرفته است.

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 مفاهیم و متغیرهای مسئله و روابط بین آنها
- مناسب بودن چارچوب ادغام ارائه شده برای مسائل محدود با ویژگی‌های خاص: مثلا شبکه‌های بیزین فازی (یکی از روش‌های مورد استفاده در تحقیقات پیشین)، برای حل 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 فرماندهی و کنترل سایبری احساس می‌شود. در این مقاله به ارائه یک راهکار مبتنی بر هستان‌شناسی برای پردازش و ادغام داده‌های سخت و نرم در مسئله فرماندهی و کنترل سایبری پرداخته شده است. در کارهای قبلی در حوزه ادغام داده‌های سخت و نرم، از هستان‌شناسی به عنوان ابزاری برای مدل‌سازی مفهومی مسئله و مبنای مرتبط‌سازی و ادغام داده‌ها استفاده نشده است.

از طرف دیگر در حوزه‌های مرتبط با فرماندهی و کنترل سایبری و آگاهی وضعیتی سایبری، تحقیقات گسترده‌ای انجام شده است که بر اساس هدف‌گذاری و رویکرد اصلی این مقاله یعنی استنتاج مبتنی بر هستان‌شناسی، به برخی از آنها اشاره می‌شود. بر اساس [1]، یک نیازمندی مهم در کسب آگاهی وضعیتی 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 این زمینه در کارهای تحقیقاتی مختلفی از جمله [8]، [9]، [10] و [11] از هستان‌شناسی و قواعد استنتاج مبتنی بر آن، برای اهداف و کاربردهای آگاهی وضعیتی سایبری از قبیل رده‌بندی خودکار آسیب‌پذیری‌ها و هشدارهای حمله، مدل کردن اطلاعات و عملیات سامانه‌های مدیریت اطلاعات و رخدادهای امنیتی، تعیین خط‌مشی‌های امنیتی، تشخیص نفوذ، ارزیابی امنیتی شبکه و سامانه‌های کامپیوتری و ارزیابی تأثیر حمله روی سامانه‌ها استفاده شده است.

در [12] به عنوان محصول یک تلاش مشارکتی برای تعریف و توسعه یک زبان استاندارد برای نمایش ساخت‌یافته اطلاعات تهدیدهای سایبری، معماری و زبان STIX[4]ارائه 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. در [13] بر اساس استانداردها و طبقه‌بندی‌های موجود از مفاهیم امنیت سایبری، هستان‌شناسی امنیت سایبری برای مدل کردن یکپارچه اطلاعات و دانش تهدید طراحی و ارائه شده است. سپس بر اساس هستان‌شناسی طراحی شده، یک چارچوب یکپارچه‌سازی اطلاعات تهدید به همراه الگوریتم تبدیل داده مورد نیاز ارائه شده است. چارچوب ارائه شده، امکان نمایش و ذخیره‌سازی یکپارچه و به اشتراک‌گذاری انواع مختلف اطلاعات تهدید از منابع چندگانه را فراهم می‌کند. در [14] یک رویکرد مبتنی بر هستان‌شناسی برای ساختن شبکه‌های بیزین ارائه شده است. با توجه به اینکه هستان‌شناسی امنیت دانش لازم را در مورد مفاهیم مختلف حوزه امنیت و ارتباط بین این مفاهیم فراهم می‌کند، این دانش می‌تواند در ساختن شبکه بیزین برای تعیین احتمال تهدید به کار گرفته شود.

کاستی اصلی مرتبط با موضوع این پژوهش، در تحقیقات پیشین آگاهی وضعیتی سایبری، استفاده ناکافی از داده نرم در کنار داده‌های سخت است. همچنین در رویکردهای مبتنی بر هستان‌شناسی ارائه شده در این حوزه، به موضوع ادغام داده‌ها و شواهد از منابع چندگانه و به ویژه با عدم قطعیت‌های متنوع، در فرآیند استنتاج و تخمین وضعیت سایبری، خوب پرداخته نشده است.

2-2. مفاهیم و داده‌ها در فرماندهی و کنترل سایبری

فرماندهی و کنترل سایبری بصورت کلی شامل کسب آگاهی لازم از وضعیت فضای سایبری مورد بررسی و گرفتن تصمیم مناسب برای اقدام در جهت امن سازی یا مقاوم سازی فضای سایبری خودی، مقابله با حملات و تهدیدات سایبری دشمن، و یا اهداف تعیین شده دیگر است. رسیدن به این مطلوب، نیازمند شناسایی به‌موقع حملات و تهدیدها روی دارایی‌های حیاتی یک سازمان یا یک شبکه، بدست آوردن تخمینی از وضعیت امنیت و کارکرد کل شبکه یا عناصری از آن مثلاً وضعیت امنیتی یک سرویس خاص، شناخت کافی از اهداف و انگیزه‌های دشمن و روش‌ها و ابزارهای مورد استفاده برای حمله و همچنین شناخت کافی از گزینه‌های در دسترس برای رفع، کاهش اثر و مقابله با حملات سایبری است.

فرماندهی و کنترل سایبری دارای جنبه‌های مختلف و مفاهیم متنوعی است. یک جنبه از آن شامل اهداف و عملکردهای مورد انتظار، انواع اطلاعات و دانش مرتبط با حوزه و فرآیندهای عملیاتی لازم می‌باشد. جنبه دیگری از آن به روش‌ها و ابزارهای قابل استفاده برای کسب اطلاعات و شناخت لازم از محیط عملیات سایبری و همچنین تأثیرگذاری در این محیط مربوط می‌شود. جنبه دیگر به نحوه ارزیابی از میزان رسیدن به اهداف و انتظارات تعیین شده اشاره دارد. یکی از مفاهیم نظری و فنی مرتبط با جنبه‌های فوق، مفهوم آگاهی وضعیتی سایبری است. در ادامه به برخی از این جنبه‌ها و مفاهیم پرداخته می‌شود.

یکی از جنبه‌های این مسئله، مفاهیم مطرح در این حوزه و

[4]Structured Threat Information eXpression

ارتباط معنایی بین آنهاست. بر اساس دانش ارائه شده از حوزه امنیت سایبری، 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. 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.

انواع داده‌ها در فرماندهی و کنترل سایبری

داده‌ها و اطلاعات پردازشی شامل داده‌های ورودی، نتایج میانی پردازش‌ها و نتایج نهایی هستند. نمونه‌هایی از این داده‌ها و اطلاعات در محیط عملیات سایبری و فضای مسئله این تحقیق، شامل موارد زیر خواهد بود:

- اطلاعاتی از دارایی‌ها و آسیب‌پذیری‌های آنها
  - وجود یک آسیب‌پذیری در یک سرویس
- گزارش‌هایی از فعالیت‌ها و رخدادهای شناسایی شده و وضعیت دارایی‌ها
  - گزارش یک فعالیت مشکوک یا یک اقدام نفوذ و حمله در شبکه
  - به خطر افتادن یک میزبان یا سرویس
  - گام‌های حمله یا سناریوهای حمله شناسایی شده
- اطلاعات و دانش از راهکارهای مقابله با تهدیدات
  - سیاست‌های امنیتی مناسب برای مقاوم‌سازی یک زیرساخت
  - وصله‌های امنیتی منتشر شده برای رفع آسیب‌پذیری‌های مشخص
  - اقدامات واکنشی مناسب برای مقابله با حملات مشخص
- نتایج پردازش‌ها که شامل اطلاعات سطح بالاتر از حمله یا وضعیت دارایی‌ها است
  - میزان پیچیدگی ابزارها و روش‌های مورد استفاده برای حمله و نفوذ
  - مقدار تجمیعی نمره آسیب‌پذیری‌های بهره‌کشی شده در حمله
  - مقدار تخمینی از سطح آسیب بالقوه یا بالفعل حمله
  - مقدار تخمینی برای انگیزه و نیت مهاجم
  - تخمینی از میزان پیشرفت یا موفق شدن حمله در رسیدن به اهداف احتمالی آن (مثلا میزان جلو رفتن در گام‌های زنجیره حمله)
  - سطح آسیب‌پذیری یک سیستم یا سرویس در مقابل حملات مشخص
  - احتمال به خطر افتادن یک میزبان بر اساس هشدارهای نفوذ و آسیب‌پذیری مرتبط با آن میزبان
  - میزان بحرانی بودن وضعیت امنیت یا کارایی یک سرور یا کل شبکه
    - میزان از دست رفتن امنیت یک سرور بر اساس تأثیر حمله روی وجوه امنیت (CIA)
    - میزان به خطر افتادن شبکه بر اساس تعداد میزبان‌ها و سرویس‌های به خطر افتاده و اهمیت سرویس‌ها
  - احتمال وقوع یک حمله مشخص بر روی یک سرور
  - پیشنهاد یک راهکار مقابله برای پیشگیری یا رفع اثر یک حمله
  - پیشنهاد پیاده‌سازی یک سیاست امنیتی برای مقاوم‌سازی یک زیرساخت

3-2. مدل باور انتقال پذیر

مدل باور انتقال پذیر[5](TBM) توسعه‌ای از نظریه شواهد دمپستر-شفر[6] است. نظریه دمپستر-شفر مدل‌های مختلفی را

[5]Transferable Belief Model

[6]Desmpster-Shafer theory of Evidence

پوشش می‌دهد که همگی از توابع باور[7] بهره می‌برند. هدف این مدل‌ها و مدل باور انتقال پذیر، مدل‌کردن درجه باور[8] افراد است، و درجه باور میزان یا درجه موافقت آن‌ها درمورد یک موضوع را نشان می‌دهد. مدل باور انتقال پذیر، بر اساس دو سطح توسعه یافته است که در یک سطح آن با نام سطح عقیدتی[9] به مدل کردن درجه باور منابع مختلف نسبت به فرضیه‌های مورد بررسی و ترکیب این باورها می‌پردازد و در سطح دوم که سطح شرط‌بندی[10] نامیده می‌شود، مقادیر باور به مقادیر احتمال تبدیل و به عبارت دیگر منتقل می‌شود تا در فرآیند تصمیم‌گیری مورد استفاده قرار گیرد [15].

درواقع، در مدل باور انتقال پذیر، یک مدل ذهنی دو سطحی جهت تمایز میان دو جنبه از باورها لحاظ شده است: باورها به‌عنوان تفکرات وزن‌دار، و باورها برای تصمیم‌گیری. باورها در سطح عقیدتی با استفاده از توابع باور مقدار می‌گیرد و ممکن است مقدار آن‌ها اصلاح شود، و زمانی که تصمیمی باید اتخاذ شود، یک انتقال شرط‌بندی[11] از توابع باور به توابع احتمال صورت می‌گیرد، که حاصل آن احتمالات شرط‌بندی نامیده می‌شود. احتمالات شرط‌بندی باورها را نمایش نمی‌دهد بلکه توسط آن‌ها برانگیخته می‌شود.

در ادامه به شرح مفاهیم و توابع مورد استفاده در این مدل پرداخته می‌شود.

تعاریف پایه

تابع جرم: یک مجموعه غیر تهی متناهی در نظر بگیرید که آن را چارچوب تشخیص[12] می‌نامند و با نماد Ω نشان می‌دهند. تابع $\boldsymbol{m}: \mathbf{2}^{\Omega} \rightarrow [\mathbf{0},\mathbf{1}]$ را تابع جرم یا تخصیص باور پایه[13] گویند، اگر دارای شرایط زیر باشد:

$$\boldsymbol{m}(\emptyset) = \mathbf{0}$$
$$\sum_{\boldsymbol{A}\subseteq\Omega} \boldsymbol{m}(\boldsymbol{A}) = \mathbf{1}$$

که در آن $\mathbf{2}^{\Omega}$ به مفهوم تمام زیرمجموعه‌های ممکن Ω است. کمیت $\boldsymbol{m}(\boldsymbol{A})$ که جرم $\boldsymbol{A}$ نامیده می‌شود، درجه اطمینان از رخداد خود پیشامد $\boldsymbol{A}$ و نه هیچ زیرمجموعه مشخص دیگر آن است.

تابع باور: تابع مجموعه‌ای $\boldsymbol{Bel}: \mathbf{2}^{\boldsymbol{\Omega}} \rightarrow [\mathbf{0},\mathbf{1}]$ را تابع باور گویند اگر دارای شرایط زیر باشد:

$$\boldsymbol{Bel}(\emptyset) = \mathbf{0}$$
$$\boldsymbol{Bel}(\boldsymbol{\Omega}) = \mathbf{1}$$
$$\boldsymbol{Bel}(\boldsymbol{A_1} \cup \boldsymbol{A_2} \dots \cup \boldsymbol{A_n}) \geq \sum_{i=1}^{n} \boldsymbol{Bel}(\boldsymbol{A_i}) - \sum_{i>j} \boldsymbol{Bel}(\boldsymbol{A_i} \cap \boldsymbol{A_j}) - (-\mathbf{1})^{n} \boldsymbol{Bel}(\boldsymbol{A_1} \cap \boldsymbol{A_2} \dots \cap \boldsymbol{A_n})$$

مقدار $\boldsymbol{Bel}(\boldsymbol{A})$ درجه باوری است که بر پایه اطلاعات و شواهد موجود، به رخ‌دادن پیشامد $\boldsymbol{A}$ داده می‌شود. برای هر مجموعه $\boldsymbol{A} \subseteq \boldsymbol{\Omega}$ داریم:

$$\boldsymbol{Bel}(\boldsymbol{A}) = \sum_{\boldsymbol{X}\subseteq\boldsymbol{A}} \boldsymbol{m}(\boldsymbol{X})$$

یعنی باور اختصاص‌یافته به مجموعه $\boldsymbol{A} \subseteq \boldsymbol{\Omega}$ برابر با مجموع جرم تمام زیرمجموعه‌های آن است.

همچنین بین تابع جرم و تابع باور رابطه زیر برقرار است:

$$\boldsymbol{m}(\mathbf{A}) = \sum_{\boldsymbol{X}|\boldsymbol{X}\subseteq\boldsymbol{A}} (-\mathbf{1})^{|\boldsymbol{A}-\boldsymbol{X}|} \boldsymbol{Bel}(\boldsymbol{X})$$

تابع امکان‌پذیری یا موجه‌نمایی: تابع مجموعه‌ای $\boldsymbol{Pl}: \mathbf{2}^{\boldsymbol{\Omega}} \rightarrow [\mathbf{0},\mathbf{1}]$ را تابع امکان‌پذیری[14] گویند اگر شرایط زیر را داشته باشد:

$$\boldsymbol{Pl}(\emptyset) = \mathbf{0}$$
$$\boldsymbol{Pl}(\boldsymbol{\Omega}) = \mathbf{1}$$
$$\boldsymbol{Pl}(\boldsymbol{A_1} \cap \boldsymbol{A_2} \dots \cap \boldsymbol{A_n}) \leq (-\mathbf{1})^{n} \boldsymbol{Pl}(\boldsymbol{A_1} \cup \boldsymbol{A_2} \dots \cup \boldsymbol{A_n}) + \sum_{i>j}^{n} \boldsymbol{Pl}(\boldsymbol{A_i} \cup \boldsymbol{A_j}) - \sum_{i=1}^{n} \boldsymbol{Pl}(\boldsymbol{A_i})$$

درواقع $\boldsymbol{Pl}(\boldsymbol{A})$ درجه اطمینانی است که برای رخ‌دادن دقیقاً

[7] Belief functions
[8] Degree of belief
[9] Credal level
[10] Pignistic level
[11] Pignistic transformation
[12] Frame of discernment
[13] Basic belief assignment (bba)
[14] Plausibility function

پیشامد $\boldsymbol{A}$ و یا هر پیشامد دیگری که با $\boldsymbol{A}$ اشتراکی دارد، قائل می‌شویم. برای هر مجموعه $\boldsymbol{A \subseteq \Omega}$ داریم:

$$\boldsymbol{Pl(A) = \sum_{X \cap A \neq \emptyset} m(X)}$$

عملیات کاستن باور

در مدل باور انتقال پذیر در صورتی که باور شما نسبت به قابل اعتماد بودن یک منبع داده کامل نباشد، می‌توان مقادیر تابع جرم متناظر با داده آن منبع را متناسب با میزان قابل اعتماد بودن منبع، پایین آورد. با فرض اینکه $\boldsymbol{\alpha \leq 1}$ میزان باور شما نسبت به قابل اعتماد بودن منبع داده باشد، مقادیر جدید تابع جرم، با اعمال عملگر کاستن[15] باور، مطابق رابطه زیر بدست می‌آید.

$$\boldsymbol{m^*(A) = \alpha * m(A) \qquad A \neq \Omega}$$
$$\boldsymbol{m^*(\Omega) = \alpha * m(\Omega) + (1 - \alpha)}$$

قاعده ترکیب دمپستر و شفر

در صورتی که چندین مجموعه شواهد از منابع مختلف برای یک موضوع ارائه شده باشد، می‌توانیم با استفاده از قواعد و عملگرهای ترکیب معرفی شده، آنها را با یکدیگر ترکیب کنیم تا باور پایه در مورد شواهد مجموعه مشترک آن‌ها به‌دست آید. با فرض اینکه $\mathbf{m_1(.)}$ و $\mathbf{m_2(.)}$ توابع جرم متناظر با مجموعه شواهد دو منبع مستقل از هم در چارچوب تشخیص $\boldsymbol{\Omega}$ باشند، بر اساس قاعده ترکیب دمپستر-شفر[16] تابع جرم ترکیبی برای هر مجموعه $\boldsymbol{A \subseteq \Omega}$ با استفاده از رابطه زیر بدست می‌آید:

$$\boldsymbol{m_{1,2}(A) = [m_1 \oplus m_2](A)} = \frac{\sum_{B \cap C = A} \boldsymbol{m_1(B).m_2(C)}}{\boldsymbol{1 - \sum_{B \cap C = \emptyset} m_1(B).m_2(C)}} \qquad \boldsymbol{A \neq \phi}$$

در رابطه فوق، عبارت $\boldsymbol{\sum_{B \cap C \neq \emptyset} m_1(B).m_2(C)}$ در مخرج کسر نشان‌دهنده میزان عدم توافق یا ناسازگاری در دو مجموعه از شواهد است. مخرج کسر برای نرمال‌سازی مقادیر جرم ترکیبی استفاده می‌شود.

در مدل باور انتقال پذیر برای بدست آوردن ترکیب عطفی[17] چند مجموعه شواهد، فرآیند نرمال‌سازی وجود ندارد، بنابراین از رابطه زیر استفاده می‌شود:

$$\boldsymbol{m_{1,2}(A) = \sum_{B \cap C = A} m_1(B).m_2(C)}$$

انتقال شرط‌بندی

در صورتی که بخواهیم براساس باورهایی که به‌دست آمده است تصمیم‌گیری کنیم، باید توابع باوری که در سطح عقیدتی ایجاد شده‌اند را به توابع احتمال نگاشت کنیم. در مدل باور انتقال پذیر، عملیات نگاشت از تابع باور به تابع احتمال، تحت عنوان انتقال شرط‌بندی انجام می‌شود و تابع احتمال نتیجه با نمایش $\boldsymbol{BetP}$ یک احتمال شرط بندی نامیده می‌شود که اندازه احتمال برای اتخاذ تصمیم است. مقدار $\boldsymbol{BetP}$ به ازای هر عضو $\boldsymbol{\omega}$ از $\boldsymbol{\Omega}$ به صورت زیر بدست می‌آید.

$$\boldsymbol{BetP(\omega) = \sum_{A:\omega \subseteq A \in \Re} \frac{m(A)}{|A|(1 - m(\emptyset))}}$$

3- راه‌حل پیشنهادی

راه حل پیشنهادی برای مسئله طرح شده شامل سه جنبه است: جنبه اول آن مربوط به نحوه مدل‌سازی مفهومی از مسئله است، جنبه دوم راه حل، مربوط به طراحی معماری و چارچوب فرآیندی پردازش، استنتاج و ادغام داده‌ها است، و جنبه سوم شامل نظریه و مدل انتخابی برای نحوه بازنمایی و ادغام داده‌ها با انواع متنوع عدم قطعیت است.

راه‌حل انتخابی برای مدل‌سازی مفهومی مسئله، استفاده از مدل هستان‌شناسی است. 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. بنابراین طرح کلی مدل پیشنهادی بر این اساس است که از هستان‌شناسی به عنوان زبانی برای توصیف و مدل‌سازی مفاهیم مورد استفاده در کاربرد فرماندهی و کنترل سایبری بهره گرفته شده است.

[15] discounting

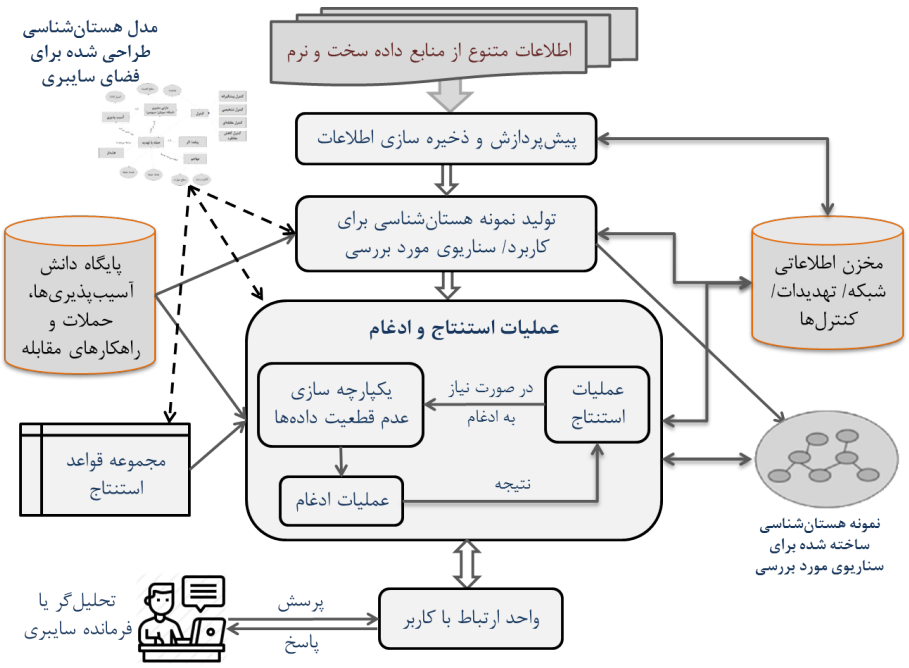
[16] Demspter-Shafer rule of combination

[17] Conjunctive combination

بر این اساس، در این مقاله دانش دامنه و منطق استنتاج موردنیاز برای اهداف موردنظر فرماندهی و کنترل سایبری، مبتنی بر هستان‌شناسی ارائه می‌شود. برای این منظور ابتدا با در نظر گرفتن نیازمندی‌های اطلاعاتی و عملیاتی این حوزه، مدل هستان‌شناسی فضای فرماندهی و کنترل سایبری توسعه داده شده است و سپس فرآیند عملیاتی لازم برای پردازش و تحلیل اطلاعات تا رسیدن به نتایج مورد انتظار در قالب یک موتور استنتاج و ادغام اطلاعات مبتنی بر هستان‌شناسی طراحی و مدل‌سازی شده است.

3-1. معماری و مدل فرآیندی پردازش و استنتاج اطلاعات مبتنی بر هستان‌شناسی

معماری پیشنهادی برای مدل‌سازی فرآیند پردازش و استنتاج اطلاعات مبتنی بر هستان‌شناسی، به منظور به‌کارگیری برای اهداف فرماندهی و کنترل سایبری مطابق شکل (1) است. معماری پیشنهادی دارای اجزاء و مؤلفه‌های «پیش‌پردازش و ذخیره سازی اطلاعات»، «مدل هستان‌شناسی فضای سایبری»، «پایگاه دانش آسیب‌پذیری‌ها، حملات و راهکارهای مقابله»، «مجموعه قواعد استنتاج»، «مخزن اطلاعاتی شبکه/ تهدیدات/ کنترل‌ها»، «تولید نمونه هستان‌شناسی برای کاربرد/ سناریوی مورد بررسی»، و «واحد عملیات استنتاج و ادغام» است. در ادامه به تشریح مؤلفه‌های مختلف مدل ارائه شده می‌پردازیم.



شکل 1. مدل فرآیندی پردازش و استنتاج اطلاعات مبتنی بر هستان‌شناسی برای فرماندهی و کنترل سایبری.

مؤلفه «پیش‌پردازش و ذخیره‌سازی اطلاعات»، با به دست آوردن داده‌ها و اطلاعات متنوع از منابع ماشینی و انسانی، به انجام پیش‌پردازش های لازم روی آنها، و ذخیره سازی آنها در قالب استاندارد تعیین شده در داخل مخزن اطلاعاتی می‌پردازد.

«مدل هستان‌شناسی فضای سایبری»، شامل توصیف مفاهیم مورد استفاده در فرماندهی و کنترل سایبری و همچنین روابط بین آنها در قالب ساختار هستان‌شناسی است. مدل هستان‌شناسی طراحی شده به عنوان اساس و مبنای ساختار یا عملیات بخش‌های دیگر معماری پیشنهادی می‌باشد. «مجموعه قواعد استنتاج»، شامل قواعد استنتاجی موردنیاز برای پاسخگویی به پرسش‌ها و نیازهای اطلاعاتی تحلیل‌گر یا فرمانده سایبری می‌باشد. این قواعد براساس زبان SWRL[18]نوشته می‌شود که

[18] Semantic Web Rule Language

فصلنامهٔ علمی - پژوهشی فرماندهی و کنترل، سال پنجم، شمارهٔ یک، بهار 1400

زبان استانداردی برای نوشتن قوانین مبتنی بر هستان‌شناسی در وب معنایی است.

مؤلفه «پایگاه دانش آسیب‌پذیری‌ها، حملات و راهکارهای مقابله»، دربرگیرنده انواع مختلف اطلاعات و دانش موجود در زمینه امنیت سایبری از قبیل اطلاعات ضعف‌ها و آسیب‌پذیری‌های کشف شده، الگوهای حملات، ابزارها و بدافزارهای مرتبط، وصله‌های امنیتی برای رفع آسیب‌پذیری‌ها، و انواع مختلف روش‌ها و راهکارهای مقابله و کاهش اثر تهدیدها خواهد بود. این پایگاه دانش با توجه به تنوع ساختارهای دانش قابل استفاده، ساختار و قالب پویایی باید داشته باشد.

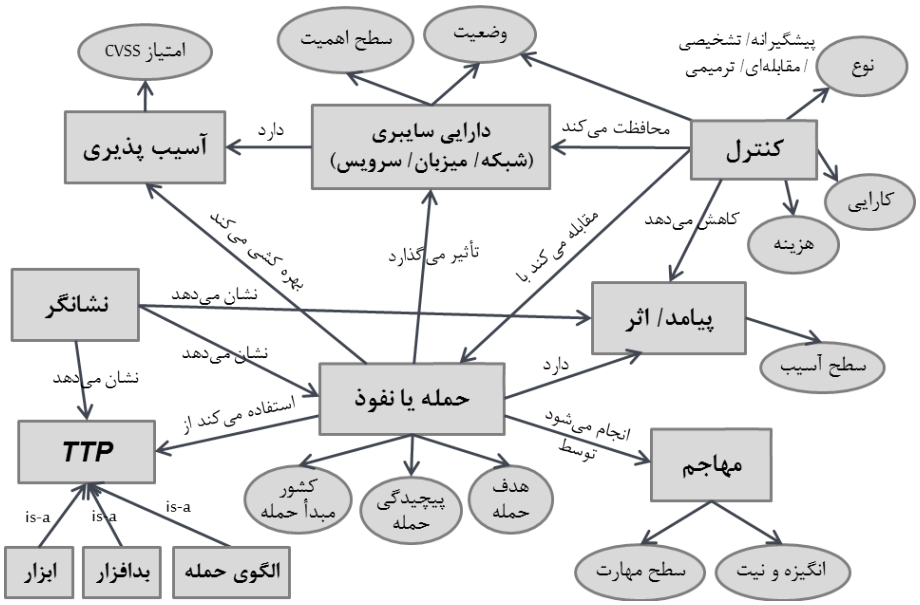
مؤلفه «مخزن اطلاعاتی شبکه/ تهدیدات/ کنترل‌ها»، پایگاه اطلاعاتی موتور استنتاج است و اطلاعات پیکربندی شبکه، آسیب‌پذیری‌های شناسایی شده در شبکه، هشدارهای حمله و نفوذ گزارش شده، کنترل‌های پیاده شده و همچنین اطلاعات وضعیتی شبکه و سرویس‌ها در آن ذخیره شده و در مراحل دیگر مورد استفاده قرار می‌گیرد.

مؤلفه «تولید نمونه هستان‌شناسی برای کاربرد/ سناریوی مورد بررسی»، بر اساس مدل هستان‌شناسی طراحی شده و اطلاعات مورد استفاده سناریوی مورد بررسی، نمونه هستان‌شناسی متناظر با آن را تولید می‌کند. نمونه هستان‌شناسی ساخته شده در فرآیند استنتاج مورد پردازش قرار گرفته و با اطلاعات استنتاجی جدید به روزرسانی می‌شود.

«واحد عملیات استنتاج و ادغام» بر اساس قواعد استنتاج تعیین شده برای سناریوی مورد بررسی و با استفاده از یک موتور استدلال‌گر مبتنی بر هستان‌شناسی شروع به اجرای این قواعد می‌کند. در صورت نیاز به ادغام در فرآیند استنتاج، با توجه به انواع متفاوت عدم قطعیت در داده‌ها، یکپارچه سازی عدم قطعیت با مدل انتخابی صورت گرفته و عملیات ادغام لازم توسط مولفه ادغام انجام می‌شود و نتایج حاصل در هستان‌شناسی ذخیره می‌گردد. این کار تا رسیدن به نتایج نهایی ادامه می‌یابد و درنهایت نتایج تولید شده به واحد ارتباط با تحلیل‌گر/ فرمانده سایبری تحویل داده می‌شود. شرح نظریه و مدل انتخابی برای یکپارچه سازی عدم قطعیت و ادغام داده‌ها در بخش‌های بعدی آمده است.

3-2. مدل هستان‌شناسی طراحی شده برای کاربرد فرماندهی و کنترل سایبری

مدل هستان‌شناسی بر اساس اهداف تعیین شده برای مسئله و دانش حوزه طراحی می‌شود. بر اساس مفهوم شناسی صورت گرفته از حوزه فرماندهی و کنترل سایبری، هستان‌شناسی مورد استفاده برای اهداف این تحقیق، مطابق شکل (2) طراحی شده است.



شکل 2. مدل هستان‌شناسی طراحی شده برای اهداف فرماندهی و کنترل سایبری.

فصلنامهٔ علمی - پژوهشی فرماندهی و کنترل، سال پنجم ، شمارهٔ یک بهار 1400

این هستان‌شناسی شامل کلاس‌های «دارایی سایبری»، آسیب‌پذیری، «حمله یا نفوذ»، TTP، «پیامد یا اثر حمله»، مهاجم، نشانگر و کنترل است. هر کدام از این کلاس‌ها دارای خصیصه‌هایی هستند، مثلا کلاس حمله دارای خصیصه‌های پیچیدگی حمله و هدف حمله است و کلاس مهاجم دارای خصیصه‌های «انگیزه و نیت» و «سطح مهارت» است.

کلاس TTP نشان دهنده مجموعه روش‌ها و ابزارهای مورد استفاده برای به انجام رساندن حمله بوده و شامل زیرکلاس‌های الگوی حمله، بدافزار و ابزار است. کلاس پیامد نشان دهنده هرگونه اثر، آسیب، خسارت و نتایج غیرمجاز حمله از جمله اختلال و خرابی در کارکرد دارایی‌ها، نقض محرمانگی و صحت داده‌ها، و آسیب به دسترس‌پذیری سرویس‌ها می‌باشد. کلاس نشانگر شامل هرگونه نشانه از به وقوع پیوستن یک حمله یا فعالیت مشکوک در فضای مورد بررسی و همچنین نشانه‌های مشاهده شده از آثار و پیامدهای حمله است.

کلاس کنترل در مدل هستان‌شناسی طراحی شده، به مجموعه سیاست‌ها، راهبردها و اقداماتی اشاره دارد که در جهت حفظ اهداف طرح ریزی شده برای سازمان یا شبکه خودی و مقابله با اهداف دشمن قابل پیاده کردن است. این کنترل‌ها به دسته‌های مختلفی از جمله موارد زیر قابل شکستن است: کنترل پیشگیرانه مثلا یک سیاست در مورد رفع آسیب‌پذیری‌های کشف شده در سریع‌ترین زمان ممکن، کنترل تشخیصی مثلا استفاده از ابزارهای لازم برای شناسایی جزئیات یک حمله انجام شده، کنترل مقابله‌ای مثلا بستن ترافیک برخی از مسیرهای شبکه برای جلوگیری از انتشار حمله، کنترل ترمیمی، و کنترل کاهش مخاطره مثلا پشتیبان گیری به‌موقع از اطلاعات حساس.

گزاره‌هایی که نشانگر اطلاعات دریافتی یا حاصل استنتاج‌ها هستند، مبتنی بر هستان‌شناسی طراحی شده خواهند بود. نمونه‌هایی از گزاره‌های مبتنی بر هستان‌شناسی در فضای فرماندهی و کنترل سایبری بر اساس اطلاعات ورودی یا نتایج پردازش‌ها شامل موارد زیر است:

- اطلاعات دارایی‌های سایبری فضای مورد بررسی
  - وجود نمونه‌هایی از کلاس دارایی سایبری، مثلا گزاره Asset(?host1) برای بیان وجود میزبان host1 به عنوان یک دارایی.
- خروجی‌های پویشگر آسیب‌پذیری
  - وجود نمونه‌هایی از کلاس آسیب‌پذیری
  - وجود یک رابطه بین نمونه‌ای از کلاس دارایی با نمونه‌ای از کلاس آسیب‌پذیری، مثلا گزاره hasVuln(?host1,?vul1) برای بیان وجود آسیب‌پذیری vul1 در دارایی host1
- خروجی ابزارهای امنیتی مورد استفاده مثل IDS و دیواره آتش
  - وجود نمونه‌ای از کلاس حمله یا نفوذ
  - وجود نمونه‌ای از کلاس TTP
  - وجود نمونه‌ای از کلاس پیامد حمله
  - وجود یک رابطه بین نمونه‌ای از کلاس حمله با نمونه‌هایی از کلاس‌های دیگر
- نتایج پردازش‌ها و استنتاج‌ها
  - مشخص کردن یک مقدار برای صفت انگیزه و نیت مهاجم
  - مشخص کردن یک مقدار برای شدت یا پیچیدگی حمله
  - مشخص کردن یک مقدار برای صفت وضعیت یک دارایی

3-3. منطق استنتاج

برای اینکه از داده‌ها و اطلاعات ذخیره شده در یک نمونه هستان‌شناسی، به نتایج موردنیاز مسئله برسیم، نیاز به طراحی منطق استنتاج داریم. منطق استنتاج طراحی شده دارای سه جنبه است: استفاده از قواعد استنتاج برای نتیجه گرفتن گزاره‌های جدید بر اساس گزاره‌های موجود، بازیابی اطلاعات خاص از هستان‌شناسی مبتنی بر نیازهای تعیین شده، و گنجاندن توابع و عملگرهای ادغام در فرآیند استنتاج. در ادامه نحوه پرداختن به این جنبه‌ها در راهکار پیشنهادی مقاله بیان شده است.

در استدلال مبتنی بر هستان‌شناسی از دو زبان پایه برای

طراحی و بیان منطق استنتاج بهره گرفته می‌شود: زبان SWRL برای بیان قواعد استنتاج، و زبان SPARQL برای بیان پرس‌وجوهای بازیابی اطلاعات. SWRL یک زبان استاندارد برای نوشتن قواعد استنتاج مبتنی بر هستان‌شناسی در وب معنایی است. 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 بر اساس وجود آسیب‌پذیری مشخص در یک دارایی و امکان بهره‌کشی آن آسیب‌پذیری توسط حمله خاص، تحت خطر بودن آن دارایی توسط حمله مشخص شده را استنتاج می‌کند. گزاره استنتاج شده می‌تواند در قواعد استنتاج بعدی استفاده شود.

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

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

Attack(?a) ∧ Attacker(?b) ∧ RunAttack(?b, ?a) ∧ motivation(?b, "political-gain") ∧ skill_level(?b, "advanced") → attack_goal(?a, "destroy") ∧ attack_severity(?a, "high");

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. برای مثال در صورتی که بدنه هر دو قاعده زیر درست باشد، حالت به مخاطره افتادن دارایی a1 از دو طریق نتیجه گرفته می‌شود:

Asset(?a1), Vulnerability(?v1), has_vul(?a1, ?v1), Attack(?at1), exploits(?at1, ?v1)
→ compromised(?a1, 0.6)

Asset(?a1), Attack(?at1), Alert(?x), indicates(?x, ?at1), related_to(?x,?a1)
→ compromised(?a1, 0.7)

با ادغام میزان احتمال یا قطعیتی که در کنار استنتاج‌ها وجود دارد، می‌توان به اطمینان و قطعیت بیشتری برای نتیجه مشتق شده (در این مثال، حالت به مخاطره افتادن دارایی a1) دست پیدا کرد.

همانطور که گفته شد، در استدلال مبتنی بر هستان‌شناسی، یک نیاز دیگر، ارائه پرس‌وجوهای بازیابی اطلاعات از هستان‌شناسی است و از زبان SPARQL برای این هدف استفاده می‌شود. SPARQL زبان پرس‌وجو برای وب معنایی است و امکان استخراج اطلاعات استنتاجی موردنیاز از روی اطلاعات ذخیره شده در هستان‌شناسی را فراهم می‌کند. در این مقاله از این زبان برای جواب دادن به برخی نیازهای اطلاعاتی منطق استنتاج استفاده شده است. نمونه‌ای از کد نوشته شده با زبان SPARQL بصورت زیر است که در آن برای حملات ذخیره شده در هستان‌شناسی با در نظر گرفتن آسیب‌پذیری‌های بهره‌کشی شده توسط هر حمله، بیشترین امتیاز cvss آسیب‌پذیری را تعیین می‌کند. این مقدار در فرمول‌های تعیین پیچیدگی حمله قابل استفاده است.

```
select ?attack (max(?cvss_score) as ?max_vul_cvss_score) where {
        ?attack exploits ?vulnerability .
        ?vulnerability has ?cvss_score .
}
group by ?attack
```

4-3. ادغام داده‌های سخت و نرم

ادغام داده‌ها و اطلاعات در فضای مسئله این تحقیق، صورت‌های مختلفی دارد. یک شکل آن در واقع بر اساس یک قاعده استنتاج صورت می‌گیرد. یعنی وقتی داده‌ها و شواهدی که مربوط به متغیرهای مختلفی از مسئله هستند و با برقرار بودن شرایط تعیین شده در بدنه یک قاعده استنتاج به ازای داده‌های موجود، یک داده جدید نتیجه گرفته می‌شود، در واقع نوعی ادغام داده صورت گرفته است. ولی صورت اصلی ادغام داده‌ها در مسئله این تحقیق، مربوط به ادغام داده‌ها و شواهد متناظر با یک متغیر یا فرضیه است که در ادامه بصورت کامل بیان می‌شود.

در صورتی که در طول فرآیند استنتاج، داده‌ها و شواهد مختلفی برای یک متغیر یا فرضیه از فضای مسئله داشته باشیم، با ادغام این داده‌ها و شواهد، دقت و اطمینان بیشتری در مورد مقادیر ممکن آن متغیر یا فرضیه بدست می‌آوریم و بنابراین با اطمینان بیشتری می‌توانیم تصمیم گیری لازم را انجام دهیم. یک نکته مهم در زمینه ادغام داده‌ها این است که در کنار داده‌های ارائه شده، مقداری به عنوان قابلیت اطمینان منبع داده یا میزان اعتبار و صحت داده می‌تواند وجود داشته باشد.

با فرض اینکه هر کدام از منابع داده سخت و نرم مورد استفاده، اطلاعات و مقادیری را در مورد متغیرهای تعریف شده در فضای مسئله و جنبه‌های اعتبار و قابلیت اعتماد منابع داده ارائه می‌کنند، صورت‌بندی مسئله ادغام داده‌های سخت و نرم به شکل زیر قابل طرح است:

FusionProblem = (X, DS, Data, Results)

X عبارت است از مجموعه متغیرهای فضای مسئله که در راه‌حل مبتنی بر هستان‌شناسی شامل موارد زیر خواهد بود:

- صفات و خصیصه‌های کلاس‌های هستان‌شناسی (مثلا نمره شدت آسیب‌پذیری، یا سطح مهارت یک مهاجم)
- فرضیه‌هایی در مورد وجود نمونه‌هایی از کلاس‌های هستان‌شناسی و یا مقادیر خاصی برای خصیصه‌هایی از نمونه‌های موجود (مثلا فرضیه به وقوع پیوستن یک نفوذ مشخص، فرضیه به خطر افتادن یک دارایی، و فرضیه بالا بودن پیچیدگی حمله شناسایی شده)

DS عبارت است از مجموعه منابع داده‌ها که می‌تواند بصورت کلی (از قبیل ابزار حسگری، کاربران عادی، تحلیل‌گرهای امنیتی، مدیران یا فرماندهان صحنه) یا بصورت جزئی و موردی (ابزار تشخیص نفوذ، سامانه مدیریت اطلاعات و رویدادهای امنیتی، کارشناس1، کارشناس2، مدیر امنیت1) مشخص شود.

Data عبارت است از مجموعه داده‌ها که می‌تواند مربوط به مقادیر متغیرها، مربوط به قابلیت اعتماد منابع داده‌ها و یا مربوط به میزان اعتبار و صحت داده‌های دیگر باشد. بنابراین هر عضو d از مجموعه داده‌ها به یکی از سه صورت زیر خواهد بود:

- data_source, x∈X, value
- data_source, s∈DS, reliability
- data_source, d∈Data, validity

Results شامل نتایج نهایی تولید شده در فرآیند استنتاج و ادغام خواهد بود.

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 فراهم شده توسط حسگرها یا پردازش‌های ماشینی عموماً دارای عدم قطعیت احتمالاتی هستند، در حالی که اطلاعات فراهم شده توسط انسان، معمولا دارای عدم قطعیت از نوع باور و امکان هستند. با توجه به این تفاوت در نوع عدم قطعیت داده‌های سخت و نرم، برای ادغام این داده‌ها با هم، نیاز به یک چارچوب نظری برای نمایش و مدل‌سازی عدم قطعیت بصورت یکپارچه وجود دارد. مدل باور انتقال‌پذیر با توجه به ویژگی‌ها و مزایایی که دارد، گزینه مناسبی برای این هدف است. بر اساس یک رویکرد مبتنی بر مدل باور انتقال‌پذیر، 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 بدست آمده از منابع مختلف اقدام کرد.

مراحل ادغام داده‌های سخت و نرم

بر اساس رویکرد پیشنهادی مبتنی بر مدل باور انتقال‌پذیر، مراحل ادغام داده‌های سخت و نرم به ازای چندین داده سخت و نرم ارائه شده برای یک متغیر یا فرضیه، مطابق شکل 3 خواهد بود که در ادامه شرح داده شده است.

تشکیل توابع جرم

برای تشکیل توابع جرم متناظر با داده‌ها، ابتدا چارچوب تشخیص مبتنی بر مقادیر ممکن متغیر یا فرضیه مربوطه و داده‌های ارائه شده، تعیین می‌شود. سپس مقادیر احتمالاتی داده‌های سخت و مقادیر باور داده‌های نرم با استفاده از فرمول‌های مربوطه به مقادیر تابع جرم در چارجوب تشخیص تعیین شده تبدیل می‌شود.

شکل 3. مراحل ادغام داده‌های سخت و نرم در رویکرد پیشنهادی

با فرض چارچوب تشخیص $\Omega = \{\omega_1, \omega_2, \omega_3\}$ فرمول‌های مورد استفاده در روش پیشنهادی ارائه می‌شود. به ازای چارچوب‌های تشخیص با تعداد اعضای بیشتر یا کمتر هم مشابه همین فرمول‌ها قابل استفاده است. در صورتی که داده سخت ارائه شده شامل مقادیر احتمال برای تمام اعضای چارچوب تشخیص باشد، مقادیر تابع جرم بصورت زیر بدست می‌آید:

Data: $\mathbf{p(\omega_1) = p_1}$, $\mathbf{p(\omega_2) = p_2}$, $\mathbf{p(\omega_3) = p_3}$ $\sum \mathbf{p_i = 1}$

Mass function:

$$\mathbf{m(\omega_1) = p_1}\ ,\ \mathbf{m(\omega_2) = p_2}\ ,\ \mathbf{m(\omega_3) = p_3}$$

و در صورتی که داده سخت ارائه شده شامل مقادیر احتمال برای برخی از اعضای چارچوب تشخیص (مثلا $\boldsymbol{\omega_1}$ و $\boldsymbol{\omega_2}$) باشد، مقادیر تابع جرم متناظر بصورت زیر بدست می‌آید:

Data: $\mathbf{p(\omega_1) = p_1}$, $\mathbf{p(\omega_2) = p_2}$ $\quad \sum \mathbf{p_i \le 1}$

Mass function: $\mathbf{m(\omega_1) = p_1}$, $\mathbf{m(\omega_2) = p_2}$,

$$\mathbf{m(\Omega) = 1 - (p_1 + p_2)}$$

در صورتی که داده نرم ارائه شده، بر اساس چارچوب نظریه باور، شامل مقادیر باور به ازای زیرمجموعه‌های مختلف چارچوب تشخیص باشد، مقادیر تابع جرم متناظر بصورت زیر بدست می‌آید:

Data: $\mathbf{Bel(A)}$ $\quad$ **for some** $\mathbf{A \subseteq \Omega}$

Mass function: $\mathbf{m(X) = \sum_{A|A\subseteq X}(-1)^{|A-X|}Bel(A)}$

کاستن باور

در ادامه مراحل ادغام، با اعمال عملگر کاستن باور روی توابع جرم متناظر با داده‌ها با توجه به درجه اعتماد تخصیص داده شده به منبع هر داده، توابع جرم کاسته شده را بدست آورده و به روزرسانی می‌کنیم. با فرض اینکه $\boldsymbol{\alpha \le 1}$ میزان درجه اعتماد تخصیص داده شده به یک منبع داده باشد، مقادیر جدید تابع جرم برای داده‌های آن منبع با استفاده از فرمول زیر بدست می‌آید.

$$m^*(A) = \alpha * m(A) \qquad A \neq \Omega$$
$$m^*(\Omega) = \alpha * m(\Omega) + (1 - \alpha)$$

ادغام داده‌ها

سپس توابع جرم بدست آمده برای شواهد مختلف را با استفاده از قاعده ترکیب دمپستر-شفر با هم ادغام می کنیم.

$$m_{1,2}(A) = [m_1 \oplus m_2](A)$$
$$= \frac{\sum_{B\cap C=A} m_1(B).m_2(C)}{1 - \sum_{B\cap C=\emptyset} m_1(B).m_2(C)} \qquad A \neq \phi$$

عملیات انتقال برای تصمیم

در نهایت در صورت نیاز به تصمیم روی فرضیه‌های خاصی از چارچوب تشخیص، عملیات انتقال سطح شرط‌بندی از مقادیر باور به مقادیر احتمال انجام شده و احتمالات متناظر با فرضیه‌های مورد بررسی مشخص می‌شوند.

$$BetP(\omega) = \sum_{A:\omega\subseteq A} \frac{m(A)}{|A|}$$

3-5. به کارگیری روش ادغام پیشنهادی در یک سناریو

در این قسمت بر اساس یک سناریوی نمونه، انواع مختلف داده‌های سخت و نرم قابل ارائه به سامانه و نحوه ادغام آنها با استفاده از روش پیشنهادی مقاله، نشان داده شده است. فرض کنید که با شناسایی یک حمله سایبری می‌خواهیم بر اساس پارامترهایی از قبیل کشور مبدأ حمله، میزان پیچیدگی حمله، و درجه شدت آسیب‌های بالقوه حمله، در مورد اقدامات واکنشی مناسب تصمیم گیری کنیم. برای این کار در درجه اول لازم است مبتنی بر اهداف و سیاست‌گزاری های تعیین شده و با استفاده از نظرات افراد خبره یا منابع دانشی دیگر، فرمول‌ها و قوانین موردنیاز را که تعیین کننده تصمیم‌های مناسب به ازای مقادیر مختلف این پارامترها هستند، تدوین شود. با فرض وجود این فرمول‌ها و قوانین، در مرحله بعد باید با پردازش‌های لازم روی

داده‌های ارائه شده و در صورت نیاز ادغام داده‌ها، مقادیر متناظر با پارامترهای تصمیم‌گیری را تعیین کنیم. و در مرحله آخر با اجرای قوانین تصمیم‌گیری، واکنش مناسب برای پاسخ به حمله را مشخص کنیم.

در این مثال «کشور مبدأ حمله» به عنوان یک پارامتر یا متغیر مسئله در نظر گرفته شده است. برای این متغیر می‌توان بر اساس داده‌هایی که توسط حسگرهای مختلف در مورد آی‌پی‌های مبدأ یا مقصد ترافیک‌های مشکوک تولید شده است و با پردازش این داده ها، احتمالاتی در مورد کشور مبدأ حمله تعیین کرد. همچنین تحلیلگرهای امنیتی ناظر صحنه بر اساس اطلاعات و دانشی که از منابع مختلف دارند، می‌توانند به ارائه نظر خود در این زمینه بپردازند. در واقع این افراد باور و اعتقاد ذهنی خود را نسبت به حالت‌های ممکن کشور مبدأ حمله مشخص می‌کنند. از طرف دیگر ممکن است فرمانده سایبری میزان اعتماد متفاوتی نسبت به نتایج پردازش ماشینی و نظرات تحلیلگرهای مختلف داشته باشد و لازم است میزان اعتماد به منابع داده هم در ادغام اطلاعات دخالت داده شود.

مقادیر ممکن این متغیر شامل موارد زیر در نظر گرفته شده است: آمریکا، انگلیس، چین و سایر

داده سخت که نتیجه پردازش ماشینی روی داده‌های حسگری است، بصورت زیر ارائه شده است:

مقدار احتمال	کشور مبدأ حمله	نماد اختصاری
0,4	آمریکا	**a**
0,3	انگلیس	**b**
0,2	چین	**c**
0,1	سایر	**d**

تحلیلگر1 نظر و باور خود را نسبت به حالت‌های مختلف کشور مبدأ حمله بصورت زیر ارائه داده است. مقدار متناظر با نامشخص در جدول به معنی باور تحلیلگر نسبت به قابل تعیین نبودن کشور مبدأ حمله از روی شواهد موجود است. در واقع این مقدار نشان دهنده میزان ناآگاهی تحلیلگر در مورد کشور مبدأ حمله است.

کشور مبدأ حمله	میزان باور تحلیلگر1
آمریکا	0,2
آمریکا یا انگلیس	0,4
چین	0,1
نامشخص	0,3

تحلیلگر2 نظر و باور خود را نسبت به حالت‌های مختلف کشور مبدأ حمله بصورت زیر ارائه داده است.

کشور مبدأ حمله	میزان باور تحلیلگر2
آمریکا	0,4
انگلیس	0,2
چین یا سایر	0,2
نامشخص	0,2

فرمانده سایبری میزان اعتماد خود را نسبت به داده‌های ارائه شده بصورت زیر تعیین کرده است.

منبع داده	میزان اعتماد
پردازش ماشینی	0,8
تحلیلگر1	0,9
تحلیلگر2	0,7

برای ادغام داده‌های فوق ابتدا همه داده‌ها را با استفاده از نظریه مدل باور انتقال پذیر ارائه می‌کنیم. بر این اساس توابع جرم متناظر با داده‌ها بصورت زیر تشکیل می‌شود:

چارچوب تشخیص:

$$\boldsymbol{\Omega} = \{\boldsymbol{a} \equiv \text{آمریکا}, \boldsymbol{b} \equiv \text{انگلیس}, \boldsymbol{c} \equiv \text{چین}, \boldsymbol{d} \equiv \text{سایر}\}$$

توابع جرم متناظر با داده سخت:

$$\mathbf{m_1(a) = 0.4}, \quad \mathbf{m_1(b) = 0.3}, \quad \mathbf{m_1(c) = 0.2}, \quad \mathbf{m_1(d) = 0.1}$$

توابع جرم متناظر با داده نرم 1:

$$\mathbf{m_2(a) = 0.2}, \quad \mathbf{m_2(\{a,b\}) = 0.4}, \quad \mathbf{m_2(c) = 0.1}, \quad \mathbf{m_2(\Omega) = 0.3}$$

توابع جرم متناظر با داده نرم 2:

$$\mathbf{m_3(a) = 0.4}, \quad \mathbf{m_3(b) = 0.2}, \quad \mathbf{m_3(\{c,d\}) = 0.2}, \quad \mathbf{m_3(\Omega) = 0.2}$$

مقادیر درجه اعتماد به منابع داده‌ها:

$$\boldsymbol{\alpha_1 = 0.8}, \quad \boldsymbol{\alpha_2 = 0.9}, \quad \boldsymbol{\alpha_3 = 0.7}$$

ابتدا با اعمال عملگر کاستن باور معرفی شده در مدل باور انتقال پذیر، روی توابع جرم متناظر با داده‌ها با توجه به درجه اعتماد تخصیص داده شده به منبع هر داده، توابع جرم کاسته شده را بدست آورده و به روزرسانی می‌کنیم.

$m_1^*(a) = \alpha_1 * m_1(a) = 0.8 * 0.4 = 0.32$
$m_1^*(b) = 0.8 * 0.3 = 0.24$
$m_1^*(c) = 0.8 * 0.2 = 016$
$m_1^*(d) = 0.8 * 0.1 = 0.08$
$m_1^*(\Omega) = \alpha_1 * m_1(\Omega) + (1 - \alpha_1) = 0.2$

$m_2^*(a) = \alpha_2 * m_2(a) = 0.9 * 0.2 = 0.18$
$m_2^*(\{a,b\}) = 0.9 * 0.4 = 0.36$
$m_2^*(c) = 0.9 * 0.1 = 0.09$
$m_2^*(\Omega) = \alpha_2 * m_2(\Omega) + (1 - \alpha_2) = 0.37$

$m_3^*(a) = \alpha_3 * m_3(a) = 0.7 * 0.4 = 0.28$
$m_3^*(b) = 0.7 * 0.2 = 0.14$
$m_3^*(\{c,d\}) = 0.7 * 0.2 = 0.14$
$m_3^*(\Omega) = \alpha_3 * m_3(\Omega) + (1 - \alpha_3) = 0.44$

سپس توابع جرم به روز شده متناظر با داده‌های نرم را با استفاده از قاعده ترکیب دمپستر با هم ادغام می‌کنیم.

$$k = 1 - \big(m_2(a)m_3(b) + m_2(a)m_3(\{c,d\}) + m_2(\{a,b\})m_3(\{c,d\}) + m_2(c)m_3(a) + m_2(c)m_3(b)\big) = 0.8614$$

$$m_{2,3}(a) = \frac{1}{k} * (m_2(a)m_3(a) + m_2(a)m_3(\Omega) + m_2(\{a,b\})m_3(a) + m_2(\Omega)m_3(a)) = (0.18 * 0.28 + 0.18 * 0.44 + 0.36 * 0.28 + 0.37 * 0.28)/0.8614 = \frac{0.334}{0.8614} = 0.39$$

$$m_{2,3}(b) = \frac{m_2(\{a,b\})m_3(b) + m_2(\Omega)m_3(b)}{k} = 0.12$$

$$m_{2,3}(\{a,b\}) = \frac{m_2(\{a,b\})m_3(\Omega)}{k} = 0.18$$

$$m_{2,3}(c) = \frac{m_2(c)m_3(\{c,d\}) + m_2(c)m_3(\Omega)}{k} = 0.06$$

$$m_{2,3}(\{c,d\}) = \frac{m_2(\Omega)m_3(\{c,d\})}{k} = 0.06$$

$$m_{2,3}(\Omega) = \frac{m_2(\Omega)m_3(\Omega)}{k} = 0.19$$

در ادامه توابع جرم بدست آمده از ادغام داده‌های نرم را با توابع جرم متناظر با داده سخت ادغام می‌کنیم که در نتیجه آن مقادیر زیر حاصل می‌شود.

$m_{1,2,3}(a) = 0.51$
$m_{1,2,3}(b) = 0.224$
$m_{1,2,3}(c) = 0.098$
$m_{1,2,3}(d) = 0.032$
$m_{1,2,3}(\{a,b\}) = 0.057$
$m_{1,2,3}(\{c,d\}) = 0.019$
$m_{1,2,3}(\Omega) = 0.06$

ممکن است لازم باشد که بر اساس مقادیر احتمالی حالت‌های مختلف این متغیر یعنی کشور مبدأ حمله، استنتاج یا تصمیم گیری‌های خاصی صورت بگیرد. مثلا ممکن است بر اساس یک قانون ثبت شده در سیستم، درصورتی که با احتمال بالاتر از 0,8 کشور مبدأ حمله آمریکا یا انگلیس باشد، اقدام واکنشی خاصی باید انجام شود.

در مدل باور انتقال پذیر برای این منظور، با استفاده از عملیاتی با عنوان انتقال شرط‌بندی، مقادیر باور به مقادیر احتمال تبدیل می‌شود.

$$P_{bet}(a) = 0.51 + \frac{0.057}{2} + \frac{0.06}{4} = 0.553$$

$$P_{bet}(b) = 0.224 + \frac{0.057}{2} + \frac{0.06}{4} = 0.267$$

$$P(AttackSourceCountry = America\ or\ England) = 0.553 + 0.267 = 0.82$$

3-6. نکات پیاده سازی موتور استنتاج و ادغام

چارچوب پیاده سازی این تحقیق به این صورت است که ابتدا اطلاعات و داده‌های دریافتی در یک نمونه هستان شناسی ذخیره می‌شود، این نمونه‌ی هستان شناسی توسط یک موتور استنتاج (شامل مجموعه‌ای از قوانین استنتاج) پردازش شده و گزاره‌های استنتاجی تولید می‌شود. پیاده سازی برنامه با استفاده از زبان برنامه نویسی پایتون و چارچوب جنگو انجام شده است. جنگو یک چارچوب نرم‌افزاری تحت وب آزاد و متن‌باز است که امکان اجرای برنامه‌های به زبان پایتون را فراهم می‌کند.

در پیاده سازی تحقیق، از دو مؤلفه زیر استفاده شده است:

- مؤلفه owlready2 که امکان کار با هستان شناسی و عملیات استنتاج روی آن را فراهم می‌کند.
- مؤلفه pyds که امکان عملیات ادغام با روش دمپستر-شفر را فراهم می‌کند.

همچنین بر اساس نیاز صورت مسئله، فیلدهای استنتاجی جدیدی به کلاس‌های هستان شناسی اضافه شده و عملیات لازم برای بدست آوردن مقدار آنها پیاده سازی شده است. برای مثال فیلد compromised_value درکلاس Asset تعریف شده است که بر اساس مقدار مجموعه‌ای فیلد compromised و در صورت نیاز ادغام مقادیر مختلف تعیین مقدار می‌شود. عملیات ادغام بر اساس عملگر ترکیب نظریه ادغام دمپستر-شفر طراحی و پیاده سازی شده است.

```
with onto:
class Asset (Thing):
def compromised_value(self):
  n = len(self.compromised)
  if (n == 0):
      return 0
  elif (n == 1):
      return self.compromised[0]
  else:
    m1 = MassFunction( {'a':self.compromised[0],
         'b':1- self.compromised[0]})
    i = 1;
    while i < n:
      m2 = MassFunction( {'a':self.compromised[i],
           'b':1-self.compromised[i]})
      m1 = m1.combine_conjunctive(m2)
      i += 1
    #  end of while
    return m1.bel('a')
```

سپس با اجرای قواعد تعریف شده با استفاده از مؤلفه pellet که یک موتور استدلال‌گر مبتنی بر هستان‌شناسی است، استنتاج‌های لازم صورت میگیرد و با فراخوانی مقادیر فیلدهای موردنیاز برای نمایش و ارائه، توابع مربوطه فراخوانی و اجرا شده و نتایج موردنظر تولید می‌شود.

4- جمع بندی

در این مقاله یک رویکرد مبتنی بر هستان‌شناسی برای پردازش و ادغام داده‌های سخت و نرم در فرماندهی و کنترل سایبری ارائه شده است که در آن به جنبه‌های مختلف این مسئله شامل معماری و مدل فرآیندی پردازش و استنتاج اطلاعات مبتنی بر هستان‌شناسی، روش بازنمایی عدم قطعیت و قابلیت اعتماد در داده‌های سخت و نرم و ادغام این داده‌ها، تبدیل باورها به احتمالات برای امکان تصمیم‌گیری روی فرضیه‌های مورد بررسی، طراحی مدل هستان‌شناسی برای اهداف فرماندهی و کنترل سایبری، و طراحی و پیاده‌سازی منطق استنتاج و ادغام اطلاعات مبتنی بر هستان‌شناسی پرداخته شده است. در رویکرد پیشنهادی از هستان‌شناسی به عنوان زبانی برای توصیف و مدل‌سازی مفاهیم مورد استفاده در کاربرد فرماندهی و کنترل سایبری بهره گرفته شده و یک معماری و فرآیند پردازشی مناسب برای تحلیل و استنتاج مبتنی بر هستان‌شناسی ارائه شده است. با راه‌حل ارائه شده در این مقاله، نشان داده شد که هستان شناسی با امکانات و ابزارهای قابل تعریف مبتنی بر آن، بستر لازم برای مرتبط سازی و ادغام داده های سخت و نرم را فراهم می کند.

داده‌ها و اطلاعات ورودی شامل داده‌های سخت یعنی داده‌ها و اطلاعات منابع ماشینی و داده‌های نرم یعنی داده‌ها و اطلاعات قابل ارائه توسط منابع انسانی در نظر گرفته شده است، تا بتوان با استفاده از مزایای هر کدام از انواع داده‌های سخت و نرم به تشخیص و تصمیم دقیق‌تر و مطمئن‌تری دست پیدا کرد. با توجه به انواع متفاوت عدم قطعیت در داده‌های سخت و نرم، برای ادغام آنها نیاز به نمایش یکپارچه انواع عدم قطعیت وجود دارد که برای این هدف از نظریه باور و مدل باور انتقال پذیر استفاده شده است و برای ادغام شواهد نیز قاعده ترکیب دمپستر-شفر به کار گرفته شده است. همچنین با توجه به نیاز به تصمیم‌گیری در چرخه فرماندهی و کنترل سایبری از قابلیت سطح شرط‌بندی در مدل باور انتقال پذیر بهره گرفته شده است که در آن باورها در انتهای کار به احتمالات شرطی موردنیاز برای اتخاذ تصمیم تبدیل می‌شود.

در هستان‌شناسی طراحی شده برای اهداف فرماندهی و کنترل سایبری، مفاهیمی از قبیل دارایی‌ها (شامل شبکه، میزبان و سرویس)، آسیب‌پذیری، حمله یا نفوذ، مهاجم، TTP (روش‌ها و ابزارهای انجام حمله)، پیامد یا اثر حمله، نشانگر (هر گونه رخداد یا وضعیت نشان دهنده حمله یا نفوذ) و کنترل (روش‌های

اقدام لازم برای پیشگیری، مقابله و کاهش اثر تهدیدها) در نظر گرفته شد. گزاره‌هایی که نشانگر اطلاعات دریافتی یا حاصل استنتاج‌ها هستند، مبتنی بر هستان‌شناسی طراحی شده خواهند بود. منطق استنتاج طراحی شده نیز شامل سه جنبه است: استفاده از زبان SWRL برای بیان قواعد استنتاج مبتنی بر هستان‌شناسی به منظور نتیجه گرفتن گزاره‌های جدید بر اساس گزاره‌های موجود، استفاده از زبان SPARQL برای بیان پرس‌وجوهای بازیابی اطلاعات از هستان‌شناسی مبتنی بر نیازهای تعیین شده، و گنجاندن توابع و عملگرهای ادغام در فرآیند استنتاج.

نتایج به کارگیری مدل پیشنهادی در یک سناریوی نمونه از فرماندهی و کنترل سایبری، عملیاتی بودن آن را در ادغام داده‌های سخت و نرم سایبری نشان می‌دهد. علاوه بر قابلیت مناسب برای استنتاج و ادغام، یکی از ویژگی‌های قابل توجه رویکرد پیشنهادی، قابلیت توسعه و مقیاس‌پذیری آن برای تطبیق با گسترش‌های جدید در ابزارها و نیازمندی‌های فضای فرماندهی و کنترل سایبری است.

از پژوهش‌های قابل انجام در ادامه این تحقیق به موارد زیر می‌توان اشاره کرد: استفاده از چارچوب‌های نظری دیگر برای مدل‌سازی عدم قطعیت و ادغام داده‌ها (برای مثال نظریه اندازه فازی و عملگرهای تجمیع مبتنی بر آن)، وارد کردن مفهوم بازنمایی‌های عدم قطعیت و یکپارچه کردن آنها در مدل هستان‌شناسی، و به‌کارگیری مدل ارائه شده در تحلیل و ارزیابی داده‌های حملات و رویدادهای معتبر از قبیل حمله استاکس نت.

5- مراجع

[1] رشیدی، علی جبار، سبحانی، سعداله، حسینی، سیدمجتبی. (1398). ارائه چارچوب مبتنی بر هستان‌شناسی برای ادغام داده‌های سخت و نرم در تحلیل امنیت سایبری. پدافند الکترونیکی و سایبری، 7(4)، 89-79.

[2] Hall, David L., Michael McNeese, James Llinas, and Tracy Mullen. "A framework for dynamic hard/soft fusion." In 2008 11th International Conference on Information Fusion, pp. 1-8. IEEE, 2008.

[3] Llinas, James, Rakesh Nagi, David Hall, and John Lavery. "A multi-disciplinary university research initiative in hard and soft information fusion: Overview, research strategies and initial results." In 2010 13th International Conference on Information Fusion, pp. 1-7. IEEE, 2010.

[4] Jenkins, Michael P., Geoff A. Gross, Ann M. Bisantz, and Rakesh Nagi. "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] Khaleghi, Bahador. Distributed Random Set Theoretic Soft/Hard Data Fusion. PhD diss. University of Waterloo, 2012.

[6] Wickramarathne, Thanuka L. An Analytical Framework for Soft and Hard Data Fusion: A Dempster-Shafer Belief Theoretic Approach. PhD diss. MIAMI UNIV CORAL GABLES FL, 2012

[7] Keyvan Golestan, Fakhri Karray, and Mohamed 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: 21046-21056, 2017

[9] Gonzalez Granadillo, Gustavo, Yosra Ben Mustapha, Nabil Hachem, and Herve Debar. "An ontology-driven approach to model SIEM information and operations using the SWRL formalism." International Journal of Electronic Security and Digital Forensics 7 4, no. 2-3: 104-123, 2012.

[10] 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) 18, no. 5: 554-562, 2013.

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

[12] Barnum, Sean. "Standardizing cyber threat intelligence information with the Structured Threat Information eXpression (STIX)." MITRE Corporation 11: 1-22, 2012.

[13] Zhao, Yishuai, Bo Lang, and Ming Liu. "Ontology-based unified model for heterogeneous threat intelligence integration and sharing." In Anti-counterfeiting, Security, and Identification (ASID), 2017 11th IEEE International Conference on, pp. 11-15. IEEE, 2017.

[14] Fenz, Stefan. "An ontology-based approach for constructing Bayesian networks." Data & Knowledge Engineering, vol.73, pp.73-88, 2012.

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

[16] McMullen, Sonya AH, Mac J. McMullen, Peter Forster, David Ison, and Patti J. Clark. "Emergency management: Exploring hard and soft data fusion modeling with unmanned aerial systems and non-governmental human intelligence mediums." In Proceedings of SAI Intelligent Systems Conference, pp. 502-520. Springer, Cham, 2016.

[17] Coetzer, Johannes, Jacques Swanepoel, and Robert Sabourin. "Dynamic fusion of human and machine decisions for efficient cost-sensitive biometric authentication." In 2020 International SAUPEC/RobMech/PRASA Conference, pp. 1-6. IEEE, 2020.

[18] Orr, S. "Data fusion to synthesise quantitative evidence, value and socio-economic factors: A framework and example of Dempster-Shafer theory." In Proceedings of the 3rd International Conference on Energy Efficiency in Historic Buildings, vol. 3, pp. 163-171. Uppsala University, 2018.