دانشگاه فردوسی مشهد
دانشکده مهندسی - گروه مهندسی کامپیوتر
آزمایشگاه فناوری وب

پایان‌نامه کارشناسی ارشد

ارائه یک مدل داده‌آمیزی معنایی مبتنی بر JDL

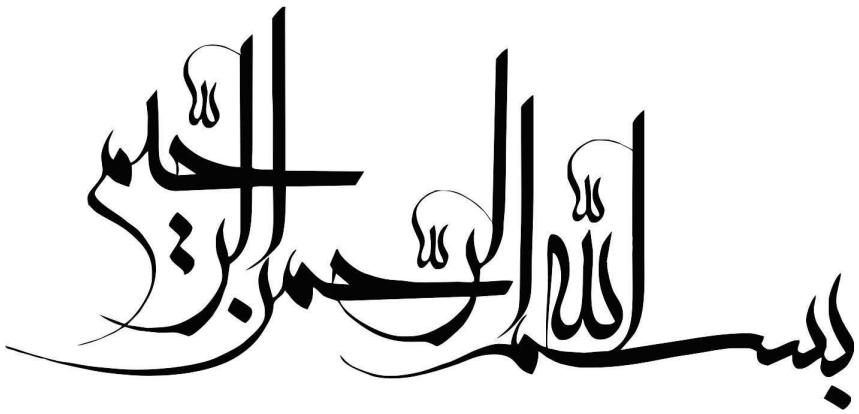
نگارش :

حوا علیزاده نوقابی

استاد راهنما:

دکتر محسن کاهانی

بهمن ماه ۱۳۹۰


بسم الله الرحمن الرحيم

تقدیر و سپاس

بدینوسیله بر خود لازم می‌دانم که از زحمات بی‌دریغ استاد گرامی، جناب آقای دکتر کاهانی که راهنمایی‌های ارزشمند ایشان در تمام مراحل انجام این این پایان‌نامه راه‌گشای من بوده است تشکر نمایم.

همچنین از تلاش‌های بی‌وقفه و پشتیبانی‌های پدر و مادر عزیزم و همسر خوبم کمال سپاس‌گزاری را دارم.

در انجام این پایان‌نامه از همکاری اعضای آزمایشگاه فناوری وب[1] بهره‌ی زیادی برده‌ام که در اینجا لازم است از آن‌ها تشکر نمایم.

[1] http://wtlab.um.ac.ir

چکیده

به واسطه افزایش جریان‌های داده ای و اطلاعاتی، بحث داده آمیزی به عنوان یکی از مهم ترین زمینه‌های تحقیقاتی و عملیاتی محسوب می‌شود. در حوزه هایی که اطلاعات زیادی وجود دارد و تصمیم گیری های ضعیف باعث پیامدهای جدی می‌شود، مسئله داده آمیزی بسیار حیاتی است. داده آمیزی اطلاعات چندین منبع را یکپارچه می‌کند و این عمل به منظور فراهم آوردن داده‌های مشخص و قابل درک درباره موجودیت ها و روابط بین آن ها صورت می‌گیرد و نهایتا منجر به استخراج دانش جدید نیز خواهد شد.

با ظهور وب معنایی و همه‌گیر شدن آن، ضرورت درک اطلاعات توسط ماشین بر هیچ کس پوشیده نیست، در این تحقیق برآنیم تا تکنولوژی های وب معنایی را وارد حوزه داده امیزی نماییم. با توجه به اینکه یکی از رایجترین مدل های داده آمیزی، مدل چندسطحی JDL می باشد، با افزودن معنا و گنجاندن آنتالوژی و سایر تکنولوژی‌های وب معنایی به این مدل، یک ساختار داده آمیزی معنایی ارائه می شود. ساختار ارائه شده یکی از چالش های اصلی موجود در سیستم‌های داده آمیزی، که چالش معنایی می باشد را برطرف خواهد نمود و ناهمگونی های موجود در نحو، ساختار و معنا را از سیستم های داده امیزی حذف خواهد نمود.

کلید واژه- داده‌آمیزی، وب معنایی، مدل JDL، انتولوژی، RDF

فهرست مطالب

فهرست جداول

فهرست شکل‌ها

فهرست علائم و اختصارات

تکمیل نشده

RDF:	Resource Description Framework.
RDFS:	Resource Description Framework Schema.
SPARQL:	SPARQL Protocol and RDF Query Language.
JDL:	

فصل ۱– مقدمه

امروزه به واسطه افزایش جریان‌های داده ای و اطلاعاتی، بحث داده آمیزی[1] به عنوان یکی از مهم ترین زمینه‌های تحقیقاتی و عملیاتی محسوب می‌شود و در این تحقیق برآنیم تا پیرامون این موضوع، به حل یکی از مسائل و چالش های اصلی آن بپردازیم.

در این فصل، در بخش ۱-۱ به موضوع اصلی تحقیق و تعریف مسئله پرداخته می شود و در بخش ۱-۲ اشاره ای به راه حل پیشنهادی مسئله خواهیم داشت و سپس در بخش ۱-۳ ساختار سایر فصل های تحقیق بیان خواهد شد.

۱–۱– تعریف مسئله

در حوزه هایی که اطلاعات زیادی وجود دارد و تصمیم گیری های ضعیف باعث پیامدهای جدی می شود، مسئله داده آمیزی بسیار حیاتی می شود و نیاز است که داده های فراوان از منابع مختلف به صورت مناسب با یکدیگر ترکیب شوند.

در مقالات گوناگون تعاریف متفاوتی از داده آمیزی ارائه شده است، اما در تعریفی کلی میتوان گفت، داده آمیزی عبارتست از استفاده از تکنیک هایی که داده های منابع مختلف را جمع آوری و ترکیب کند تا بتواند عمل استنباط را به نحوی انجام دهد که نتیجه بدست آمده بسیار موثرتر از نتیجه بدست آمده از یک منبع باشد، در واقع داده آمیزی اطلاعات چندین منبع را تفسیر و یکپارچه می کند و این عمل به منظور فراهم آوردن داده های مشخص و قابل درک درباره موجودیت ها و روابط بین آن ها صورت می گیرد و نهایتا منجر به استخراج دانش جدید نیز خواهد شد.[2][1]

[1] Data fusion

کاربردهای گسترده داده آمیزی در حوزه های مختلف (شبکه های سنسوری، پیش بینی وضع هوا، کنترل ترافیک وسایل نقلیه، پیمایش ربات، سیستم‌های کنترل و فرمان نظامی، مدیریت بحران، ...) باعث شده است که تا کنون مدل ها و معماری های بسیاری برای داده آمیزی ارائه شود. این مدل ها را در سه دسته کلی میتوان طبقه بندی نمود: مدل های مبتنی بر اطلاعات، مدل های مبتنی بر فعالیت، مدل های مبتنی بر نقش.[3]

مدل های مبتنی بر اطلاعات به داده های تولید شده در طول پروسه آمیزش، متمرکز می شوند. از جمله مدل های رایج مبتنی بر اطلاعات، میتوان به مدل معروف JDL و مدل DFD[4] اشاره کرد. تمرکز مدل های مبتنی بر فعالیت، به گام ها و فعالیت هایی است که برای پروسه آمیزش بایستی انجام شود. از جمله مدل های مبتنی بر فعالیت، میتوان مدل های Boyd Control Loop[5] و مدل Omnibus[6] را نام برد. مدل های مبتنی بر نقش به چگونگی طراحی و مدل کردن داده‌امیزی توجه می کنند و نقش‌های لازم را در نظر گرفته و ارتباطات آن‌ها را مدل می‌کنند. از مدل‌های مبتنی بر نقش میتوان Frankel-BedworthArchitecture[7] را نام برد. توضیح کامل برخی از مدلهای داده‌امیزی رایج در پیوست ۱ آورده شده است.

ناگفته نماند که در اکثر محیط ها حسگر مجرد وجود ندارد و اطلاعات جمع آوری شده توسط یک منبع مجرد، ناکافی و ناکامل می باشد و با تعداد زیادی منبع یکنواخت و غیریکنواخت روبه رو هستیم. از این رو استفاده از سیستم داده آمیزی در محیط های زیادی کاربرد دارد و ارائه یک مدل خوب برای داده آمیزی کمک زیادی به محیط های بزرگی که نیاز به جمع آوری اطلاعات دارند، خواهد نمود.

با توجه به اینکه در اکثر سیستم ها داده های منابع مختلف از نظر نحو[1]، شما[2] و یا معنا[3] ناهمگونی دارند، تفسیر درست داده ها کاری دشوار می شود. ناهمگونی در نحو از تفاوت در زبان، ناهمگونی در شما از تفاوت در ساختار و ناهمگونی در معنا از تفاوت در تفسیر داده ها ناشی می شود.[8][9]

علاوه بر اینها سه تداخل عمده دیگر که در سیستم های داده امیزی با آن روبه‌رو هستیم، عبارتند از:[10]

[1] syntax
[2] schema
[3] semantic

۱- تداخل های گیج کننده[1] (هنگامی که اطلاعات مانند یکدیگر به نظر می رسند اما واقعا اینطور نیست.)

۲- تداخل های مقیاسی[2](در اندازه گیری مقادیر پیش می آید مثل واحد پول کشورهای مختلف)،

۳- تداخل های نامگذاری[3] (تداخل هایی که از نامگذاری ها ناشی می شوند مانند مترادف ها).

بنابراین یکی از چالش های اصلی موجود در سیستم های داده آمیزی، چالش معنایی و عدم وجود معنا و مفهوم مشترک در سراسر این سیستم ها می باشد. در حقیقت برای حل چالش معنایی باید به مفهوم سازی حوزه موردنظر پرداخت و بیانی صریح و رسمی از مفاهیم و ویژگی های آن ها بیان نمود تا فهم مشترکی از یک زمینه ارائه شود. بنابراین با معنایی کردن ساختارهای موجود و ایجاد یک مدل داده امیزی معنایی، میتوان چالش های معنایی و همچنین ناهمگونی هایی که در روند داده آمیزی ایجاد می شوند، را برطرف نمود.

۱-۲- راه حل

همانطور که در بخش ۱-۱ توضیح داده شد، یکی از چالش های اصلی روند داده آمیزی، ناشی از عدم وجود معنا و مفهوم مشترک می باشد. از این رو با وارد نمودن تکنیک های وب معنایی در مدل های داده آمیزی موجود درصدد رفع این چالش و ایجاد سازگاری معنایی هستیم.

یکی از رایجترین مدل های داده آمیزی، مدل JDL[4] می باشد که بیشتر در کاربردهای نظامی دیده می شود. در این تحقیق برآنیم تا با افزودن معنا و گنجاندن آنتالوژی و استفاده از سایر تکنولوژی های وب معنایی در این مدل، آن را به یک مدل معنایی گسترش دهیم و یک ساختار داده آمیزی معنایی ارائه دهیم.

مدل JDL در طی سال های مختلف بسیار مورد توجه محققان قرار گرفته است و کارهای تحقیقاتی بسیاری مبتنی بر این مدل انجام شده است. کار پیشنهادی این تحقیق، ارائه ساختاری معنایی مبتنی بر این مدل می باشد، به عبارتی سطوح ساختار پیشنهادی براساس سطوح JDL خواهد بود، اما تمام سطوح معنایی می شوند،

[1] Confounding conflict
[2] Scaling conflict
[3] Naming conflict
[4] Joint Directors of Laboratories

در حقیقت فناوری های وب معنایی را با مدل JDL پایه ، ترکیب کرده ایم تا یک مدل معنایی برای داده آمیزی به دست آوریم. جزئیات کامل مدل پیشنهادی در بخش ۳-۱ توضیح داده خواهد شد.

۱-۳- ساختار پایان نامه

تکمیل خواهد شد

فصل ۲– کارهای مشابه

در این فصل کارهای مرتبط با مدل پیشنهادی مورد بررسی قرار می گیرند. با توجه به اینکه مدل پیشنهادی یک مدل داده آمیزی معنایی مبتنی بر JDL می‌باشد، کارهای مرتبط با آن را می‌توان در دو دسته بزرگ تقسیم‌بندی نمود. دسته اول به کارهای مبتنی بر مدل JDL می پردازد، که در بخش ۲-۱ آورده شده است و دسته دوم مربوط به کارهای داده آمیزی معنایی است که در بخش ۲-۲ توضیح داده می شود.

۲–۱– کارهای مبتنی بر مدل JDL

طی سال های مختلف تحقیقات فراوانی بر پایه مدل JDL انجام شده است که در این بخش به معرفی برخی از مهمترین این تحقیقات می پردازیم. در این بخش پس از معرفی مختصر مدل JDL پایه در بخش ۲-۱-۱ به دسته بندی و معرفی کارهای مبتنی بر این مدل در بخش ۲-۱-۲ می پردازیم.

۲–۱–۱– مدل پایه JDL

این مدل اولین بار در سال ۱۹۸۷ توسط گروه JDL در وزارت دفاع امریکا ارائه گردید و سطوح مختلفی در آن تعریف شد که در ادامه کلیت هر سطح را بیان می کنیم:[12][11][1]

سطح صفر: (پیش پردازش[1]) در سطح صفر عمل پیش پردازش روی داده های منابع گوناگون صورت می گیرد و داده های پاکسازی شده به عنوان خروجی این سطح برگردانده می شود.

سطح اول: (پالایش شی) در سطح ۱ مدل، تحلیل بیشتر مربوط به اهداف یا اشیاء منفرد می‌باشد. در واقع برای هر شی به صورت جداگانه اطلاعات هویتی، پارامتری و موقعیتی، جمع بندی و ارائه می شود. بنابراین خروجی این سطح شامل اشیا و صفات آنها خواهد بود.

[1] Pre-processing

سطح دوم: (پالایش موقعیت[2]) در این سطح با توجه به دانش پیشین و اطلاعات محیطی، توصیف مفهومی از روابط بین اشیا ایجاد خواهد شد. با توجه به اینکه در این سطح موجودیت ها به همراه صفاتشان موجود هستند،(خروجی سطح یک بودند) به کمک تعیین موقعیت اشیا و روابط بین آن ها تصویری کلی از موقعیت کنونی تعیین می‌گردد.

سطح سوم: (پالایش تهدید[3]) ارزیابی موقعیت جاری و پیش بینی تهدید ها و آسیب پذیری ها از وظایف این سطح است. همچنین دراوردن راه حل ها و فرصت عمل ها در موقعیت بوجود آمده برعهده این سطح می باشد.

می‌توان گفت که محصولات پردازش‌های لایه‌های ۱ تا ۳ مدل JDL عبارتند از:

- لایه ۱: نمایش‌های اشیاء،
- لایه ۲: نمایش‌های ارتباط میان اشیاء،
- لایه ۳: نمایش‌های اثرات ارتباط میان اشیاء.

سطح چهارم: (پالایش فرآیند[4]) در سطح ۴ نظارت بر کارایی سیستم و تخصیص منابع بر اساس هدف تعیین شده صورت می پذیرد. به عبارت دیگر اندازه گیری کارایی و میزان موثر بودن سیستم بر عهده این سطح است. پیوستگی نظارت، انتخاب اطلاعات مدیریتی، ایجاد کیفیت سرویس در این سطح انجام می شود.

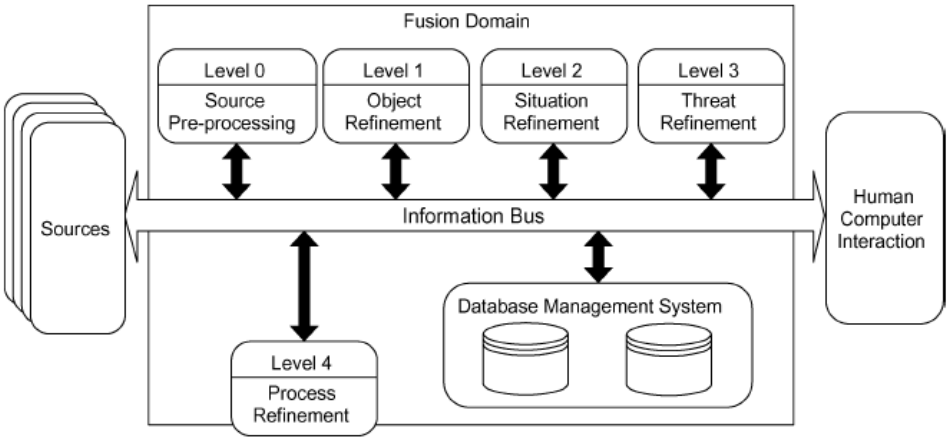
در واقع پالایش فرآیند تضمین می‌کند ترکیب داده‌ای به گونه‌ای انجام شده و پیش رود که نتایج بهینه‌ای برای سیستم حاصل شود. این مرحله می‌تواند با تنظیم پارامترها در فرآیند داده آمیزی، ایجاد دانش پیشین از اهداف و یا حرکت حسگرها جهت رسیدن به پوشش مناسبتری از ناحیه جستجو باعث بهبود نتایج نهایی سیستم شود.

---

[1] Object refinement
[2] Situation refinement
[3] Threat refinement
[4] Process refinement

سیستم مدیریت پایگاه داده : مدل JDL از یک سیستم مدیریت پایگاه داده که عملیات به روز رسانی و آماده سازی اطلاعات برای فرآیند داده آمیزی را انجام می دهد نیز بهره می‌برد. با توجه به حجم بالای داده‌ها در یک فرآیند داده امیزی، وجود چنین سیستمی ضروری به نظر می رسد.

تعامل انسان با کامپیوتر[1] : بخش دیگر این مدل، تعامل انسان با کامپیوتر می باشد که بسیار حیاتی بوده و در واقع یک واسطه ارتباطی برای ورودی انسانی و مخابره نتایج داده امیزی برای کاربران را فراهم می‌کند.



شکل ۱- مدل پایه JDL [11]

۲-۱-۲- دسته بندی کارهای مبتنی بر JDL

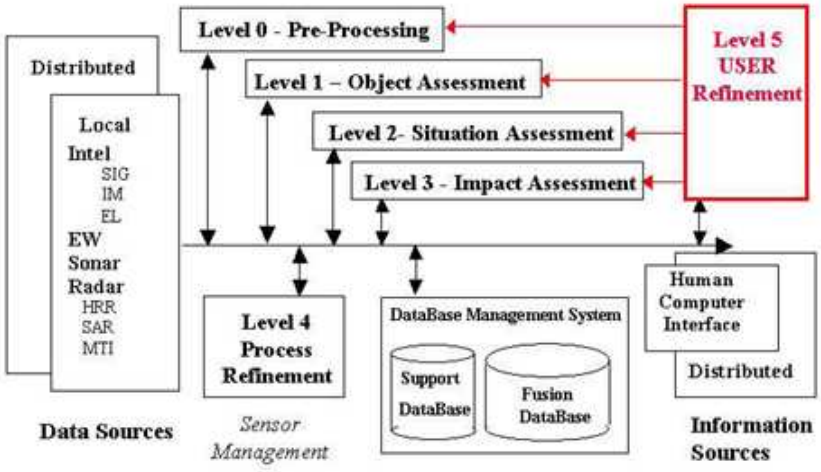
ما در این بخش ابتدا یک دسته بندی برای کارهای مبتنی بر JDL ارائه می کنیم و سپس هر دسته را با مثال توضیح می دهیم. دسته بندی پیشنهادی ما که در شکل ۲ آورده شده است، کارهای مرتبط در این زمینه را به سه دسته کلی، تقسیم می کند.

شکل ۲- کارهای مبتنی بر مدل JDL

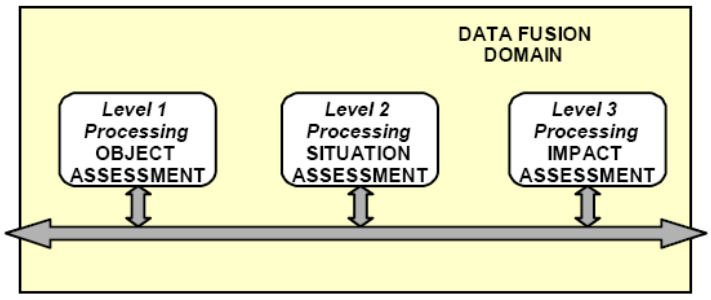
[1] . Human- computer interaction (HCI)

دسته اول شامل کارهایی می باشد که در "سطوح" مدل JDL تغییرات ایجاد می کنند، مثلا سطوح موجود در مدل پایه را با یکدیگر ترکیب می کنند یا سطح جدیدی را به مدل اضافه می کنند. از جمله کارهایی که در این دسته قرار گرفت، افزودن سطح پنجمی به این مدل با نام "user refinement" می باشد که نقش کاربر را در مدل پررنگ می نماید.[15][14] [13] شکل ۳ جایگاه سطح پنجم را در مدل پایه نشان می دهد.

از دیگر کارهایی که در رابطه با تغییر سطوح انجام شد، ارائه مدلی با نام λ-JDL بود که در آن سطح صفر و ۱ ادغام شدند و سطح ۴ نیز به صورت ضمنی در تمام سطوح قرار می گرفت.[16][17] نتیجه این ادغام ها یک مدل ۳ سطحی شد که در شکل ۴ آورده شده است.



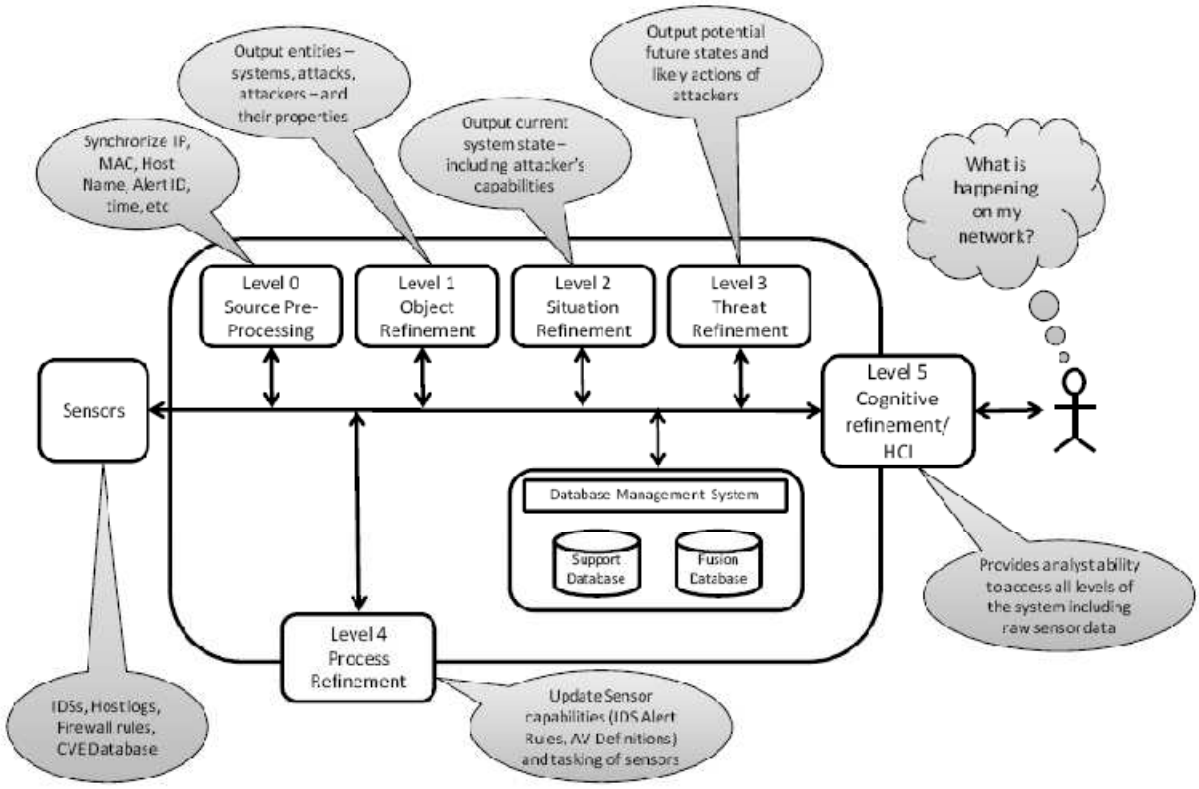
شکل ۳- مدل JDL دارای سطح پنجم یا سطح کاربر[13]



شکل ۴- مدل λ-JDL [16]

در دسته دوم کارهای مبتنی بر JDL، محققان روال داده آمیزی این مدل را به دقت مورد بررسی قرار می دهند و جهت بهبود این روند از جهات مختلف، راه حل هایی پیشنهاد می کنند. به عنوان مثال، بررسی مجدد مدل JDL می باشد که در آن به بهبود فرایند داخلی هر نود داده امیزی پرداخته شد و محققان تلاش کردند تا مواردی مانند کیفیت سرویس، قابلیت اعتماد و سازگاری در روند داده امیزی را تامین کنند.[18] کارهای که به دقت به فرایند داده آمیزی این مدل می پردازند، و ان را از جهات گوناگون بهبود می دهند، فراوان هستند و این باعث شده که از سال ارائه مدل پایه تا کنون شاهد تغییرات و بهبود در این مدل باشیم.[11][19][20][21]

در کارهایی که در دسته سوم قرار می گیرند، به این موضوع پرداخته می شود که نگاشت مناسبی بین مدل JDL و حوزه تخصصی دیگری انجام شود و با تطبیق مناسب، از این مدل در حوزه های دیگر بهره ببرند. یکی از کارهای مرتبط با این دسته، بکارگیری مدل JDL در حوزه امنیت سایبر است،[22] که تعریف هر سطح و ورودی و خروجی سطوح مختلف مدل پایه تغییر می کند و روند داده آمیزی به گونه ای نگاشت داده می شود که به تشخیص حملات امنیتی منتهی شود. شکل ۵ تطبیق مدل JDL را برای کشف حملات امنیتی نشان می‌دهد.



شکل ۵- بکارگیری مدل JDL برای تشخیص حملات امنیتی[22]

از دیگر مواردی که در دسته سوم قرار می گیرند، کاری است که مدل JDL را با انجام نگاشتی مناسب برای حوزه بیوانفورماتیک به کار گرفت.[23] و یا کار دیگری که از JDL به عنوان مدل پایه برای ساخت مجازی[1] استفاده کرد.[24]

مشاهده شد که با توجه به اهمیت مدل JDL و رایج بودن آن نسبت به سایر مدل ها، تا کنون کارهای زیادی براساس این مدل پایه انجام شده است، که میتوان انها را در سه دسته بزرگ جای داد.

۲-۲- داده آمیزی معنایی

روش‌های تلفیق داده و اطلاعات معمولاً از موضوعات علمی مختلفی استنتاج شده‌اند و سیستم‌های تلفیق داده و اطلاعات اغلب شامل چندین روش از چندین زمینه علمی مختلف می‌باشند. این روش‌های چند گانه

[1] Virtual manufacturing

معمولاً بصورت مشترک با یکدیگر و یا بشکل متوالی اعمال می‌شوند. و اما امروزه یک روند رو به رشد از تکنولوژی های وب معنایی در حال شکل‌گیری می باشد که در زمینه های مختلفی از علوم از جمله داده امیزی وارد شده است.

در بخش ۲-۲-۱ تکنیک های کلیدی وب معنایی به اختصار توضیح داده می شود و در ادامه در بخش ۲-۲-۲ نمونه هایی از کارهای داده آمیزی که از تکنولوژی های وب معنایی استفاده کردند، توضیح داده می‌شوند.

۲-۲-۱- تکنولوژی های کلیدی وب معنایی

وب معنایی، یک ابتکار با هدف بهبود وضعیت فعلی شبکه جهانی وب و رفع مشکلات آن می باشد که ایده اصلی آن استفاده از اطلاعات قابل پردازش توسط ماشین است. تکنولوژی های کلیدی وب معنایی عبارتند از ابرداده صریح ، انتولوژی، منطق و عوامل هوشمند.[25]

ابرداده صریح[1]: وب معنایی پیشنهاد می‌کند که علاوه بر اطلاعاتی که به ایجاد یک سند برای مخاطبان انسان کمک می‌کند، اطلاعاتی در مورد محتوا نیز ذخیره شود. این کار به مراتب پردازش توسط ماشین‌ها را اسان تر می کند. لغت ابرداده به "داده‌هایی درباره داده‌ها " گفته می‌شود. ابرداده بخش معنای داده‌ها را در خود دارد.

انتولوژی: به طور کلی، انتولوژی یک حوزه از سخن را بطور رسمی توصیف می‌کند. . به طور معمول ، انتولوژی شامل یک لیست محدود از اصطلاحات و روابط بین آنهاست که مفاهیم مهم را در یک دامنه، مشخص می‌کند. انتولوژی، فهم مشترکی از یک زمینه ارائه می‌دهد، چنین درک مشترکی برای غلبه بر تفاوت در اصطلاحات لازم است. در حال حاضر مهمترین زبان‌های تعریف انتولوژی RDF[2]، RDFS[3]، OWL[4] می باشد.

منطق[1] : منطق، زمینه علمی است که در مورد اصول اولیه استدلال مطالعه می‌کند. به طور کلی منطق، اول یک زبان رسمی برای توصیف دانش را ارائه می‌دهد . دوم، درک خوبی از معانی رسمی فراهم می‌کند. و سوم،

[1] Explicit metadata
[2] Resource Description Framework
[3] RDF Schema
[4] Web Ontology Langauage

استدلال‌گر خودکار می‌تواند نتایج را از دانش داده شده استنباط کند، بنابراین اطلاعات گنگ و مبهم را روشن خواهد ساخت. قابل توجه است که منطق می‌تواند برای دانشی که به طور ضمنی در انتولوژی آمده است، استفاده شود . انجام این کار، همچنین می‌تواند به کشف روابط غیر منتظره و تناقضات کمک کند.

عامل ها[۲] : عامل‌ها بخشی از نرم‌افزار هستند که بطور مستقل و فعالانه عمل می‌کنند. بطور مفهومی، آنها از مفاهیم برنامه نویسی شیء گرا و توسعه نرم افزار های مبتنی بر مولفه بدست آمده‌اند. یک عامل برخی وظایف و اولویت‌ها را از شخص دریافت می‌کند، ، به دنبال اطلاعات از منابع می‌گردد، با عوامل دیگر ارتباط برقرار می‌کند، اطلاعات را در مورد شرایط و اولوبت‌های کاربر مقایسه می‌کند، گزینه های خاص را انتخاب می‌کند و به کاربر پاسخ می‌دهد. عامل می‌تواند از تمام فن‌آوری‌های فوق استفاده کند:

- فراداده برای شناسایی و استخراج اطلاعات از منابع استفاده خواهد شد.
- آنتولوژی برای کمک در جستجو و تفسیر اطلاعات بازیابی شده ، برای برقراری ارتباط با دیگر عوامل استفاده خواهد شد.
- منطق، برای پردازش اطلاعات بازیابی شده و بدست آوردن نتیجه مورد استفاده قرار خواهد گرفت.

۲-۲-۲- کارهای داده آمیزی معنایی

برخی از کارهای معنایی حوزه داده آمیزی، درصدد ارائه چارچوب معنایی برای این منظور بودند و در چارچوب خود جایگاه تکنولوژی های وب معنایی را مشخص می کردند. از جمله این موارد، چارچوب معنایی است که در سال ۲۰۰۸ برای ترکیب داده های سنسوری پیشنهاد شد.[26] هدف چارچوب این است که داده‌ها را از حسگرها جمع‌آوری کند و عملیات‌های سطح بالا را انجام دهد و دیدی متحد از شبکه حسگر به کاربر ارائه می‌دهد. چارچوب پیشنهادی که در شکل ۶ نشان داده شده است، سعی دارد برنامه انعطاف‌پذیری را برای

---

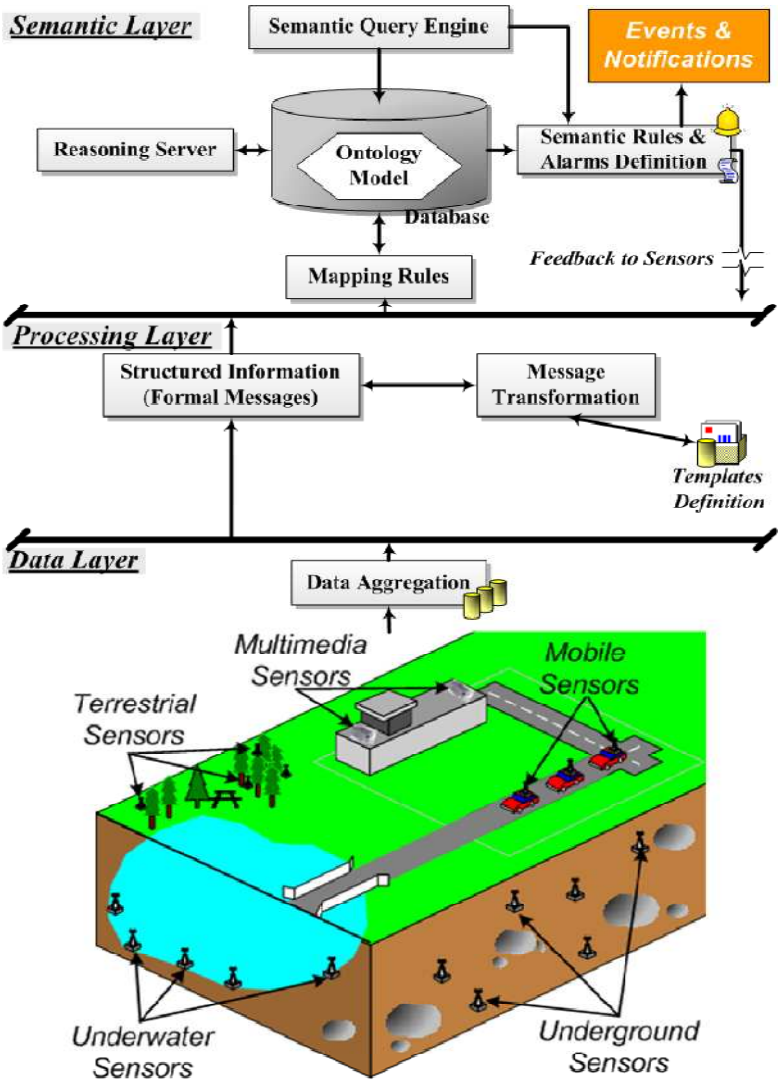
[1] logic
[2] agents

استفاده ساده و مدیریت شبکه‌های حسگر مقیاس بزرگ که متشکل از منابع داده ناهمگن هستند ارائه کند. این معماری متشکل از سه لایه است: لایه داده‌ها برای کشف، گردآوری و جمع‌آوری داده‌ها، لایه پردازش برای پردازش داده‌های جمع‌آوری شده و لایه معنایی برای تامین سازگاری مفهومی وحاشیه نویسی محتوا[1] به کمک آنتولوژی ها.

همانطور که شکل ۶ نشان می دهد، لایه معنایی این چارچوب شامل بخش هایی می باشد که عبارتند از:

- ***قوانین***: به طور کل، دو مجموعه قانون مجزا وجود دارد، مجموعه اول دربردارنده قوانینی است که مشخص می‌کند چگونه اندازه‌گیری‌های حسگر که در قالب XML[2] می‌باشد در آنتولوژی انتخاب شده نگاشت شود که قوانین نگاشت نامیده می شوند و مجموعه دوم بر اساس حقایق موجود و دانش آنتولوژی به حقایق، عمل ها و زنگ خطرها را نشان خواهد داد که قوانین معنایی نامیده می شوند.
- ***مدل آنتولوژی***: در وب معنایی مدل آنتولوژی به صورت سه گانه (موضوع، ویژگی، هدف) در یک پایگاه داده رابطه‌ای ذخیره شده و گراف مدل را تشکیل می‌دهد. مدل را می‌توان به عنوان یک گراف RDF[27] با استفاده از یکی از روش‌های پیاده‌سازی سه‌گانه ذخیره کرد. محتوای آنتولوژی باید محیط سیستم را تشریح کند.

[1] Context annotation
[2] Extensible Markup Language



شکل ۶- چارچوب معنایی برای داده امیزی سنسوری[26]

- ***سرور استدلال***: در چارچوب معنایی، پایگاه دانش ترکیبی است از یک آنتولوژی و استدلالگر که توانایی استنتاج حقایق را از حقایق موجود در آنتولوژی دارد. استدلال‌گر بخشی ضروری است زیرا دانش جدید در همین بخش تولید می‌شود. به علاوه، استدلال‌گر مدل آنتولوژی را بازرسی کرده و ثبات، رضایت‌بخشی و طبقه‌بندی مفاهیم آن را بررسی می‌کند تا به مدلی سازگار دست یابیم که هیچ ناهماهنگی در آن وجود ندارد.

- ***پرس وجوهای معنایی:*** لایه معنایی باید امکان گرفتن اطلاعات از پایگاه دانش را فراهم کند، بنابراین یکی از مؤلفه‌های مهم این لایه یک رابطه پرس و جوی معنایی است که اجازه ایجاد و اجرای پرس و جوهای هوشمند را می‌دهد. زبان پرس و جوی SPARQL[1] می‌تواند به عنوان دریچه‌ای برای دسترسی به دانش ذخیره شده به صورت سه تایی عمل کند.

طی سال های اخیر نقش آنتولوژی ها در داده امیزی بسیار موردتوجه محققان قرار گرفته است. یکی از کارهای خوبی که در این زمینه ارائه شده است، به کاربرد آنتولوژی ها در سیستم های ترکیب داده ناهمگون۲ پرداخته شده است.[28] با توجه به اینکه در اکثر سیستم ها داده های منابع مختلف از نظر نحو، شما و یا معنا ناهمگون می باشند، و تداخل های مختلفی در روند داده امیزی پیش می اید، تفسیر درست داده ها کاری دشوار می شود. از این رو آنتولوژی ها با بیان صریح و مفهوم سازی سراسری، این ناهمگونی ها را برطرف می نماید، همچنین در محیط های چندعامله انتولوژی ها باعث سهولت در اشتراک دانش۳ و استفاده مجدد۴ می شود.[29]

به سه طریق میتوان آنتولوژی ها را در سیستم های داده آمیزی بکار برد:[9][30]

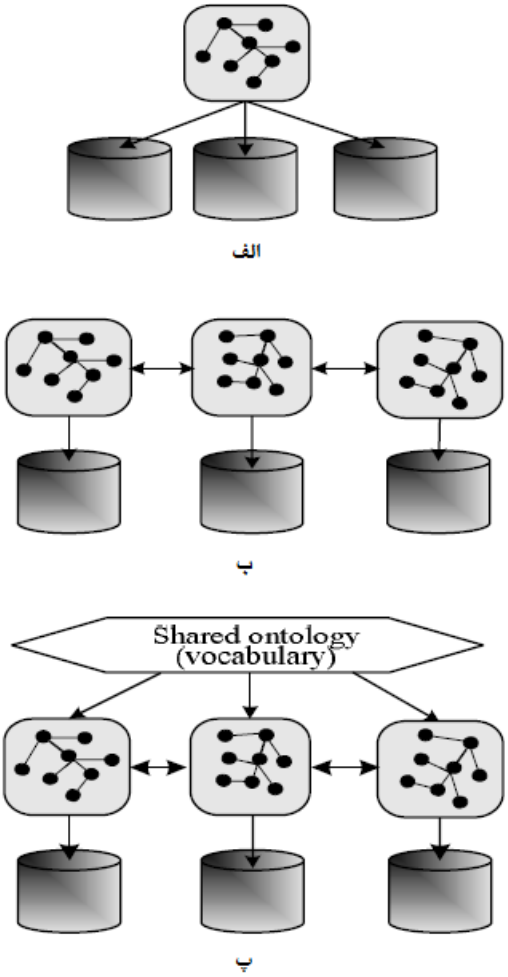
۱- **روش آنتولوژی مجرد**[5]: در این روش تنها یک آنتولوژی سراسری برای همه منابع داده وجود دارد. یک نمونه از سیستم هایی که بر این مبنا وجود دارد، SIMS [31] می باشد. شکل ۷- الف این روش را نشان میدهد.

۲- **روش آنتولوژی چندگانه**[1]: در این روش هر منبع، آنتولوژی محلی خودش را دارد و بین آنتولوژی های محلی منابع، باید نگاشت[2] انجام شود. سیستم OBSERVER [32] یک نمونه استفاده از این روش است. شکل ۷- ب این روش را نشان میدهد.

---

[1] SPARQL Protocol and RDF Query Language
[2] heterogeneous
[3] Knowledge sharing
[4] reuse
[5] Single ontology approach

۳- **روش انتولوژی ترکیبی**[۳]: این روش ترکیبی از دو روش فوق است که در ان برای هر منبع به صورت محلی آنتولوژی وجود دارد که با یک مجموعه لغات اشتراکی که دربرگیرنده لغات اساسی دامنه مدنظر هستند، در ارتباط می باشند. شکل ۷- پ این روش را نشان میدهد.

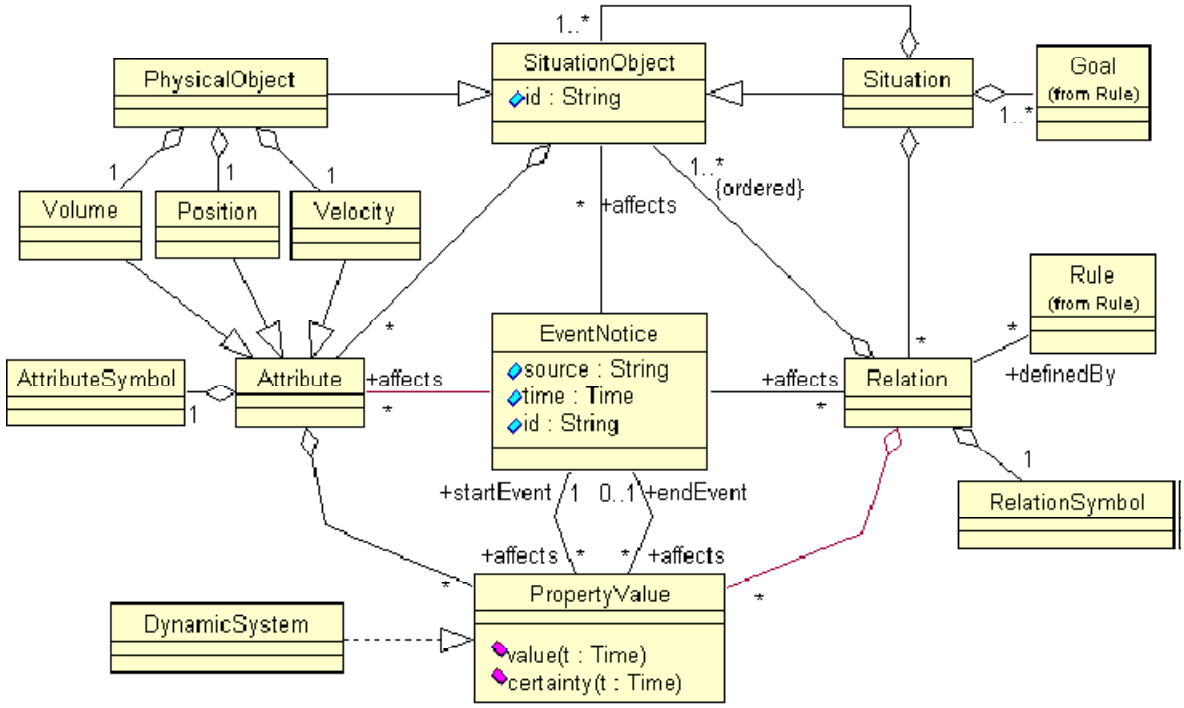


شکل ۷- روش های بکارگیری انتولوژی در سیستم های داده آمیزی[9]

برخی دیگر از کارهایی که در راستای داده آمیزی معنایی انجام می شوند، ارائه آنتولوژی سطح بالا برای بخشی از معماری ها و مدل های موجود، می باشد. به عنوان مثال، گروهی از محققان یک انتولوژی پایه برای

[1] Multiple ontology approach
[2] mapping
[3] Hybrid ontology approach

آگاهی موقعیت[1] پیشنهاد دادند که شامل تمام مولفه ها و اجزای موقعیت است. این آنتولوژی پایه در شکل ۸ آورده شده است. [33] گروهی دیگر از محققان این انتولوژی پایه را با افزودن روابط زمانی و مکانی گسترش دادند.[34] در یک کار تحقیقاتی دیگر تمام انتولوژی های سطح بالا که برای آگاهی وضعیتی ارائه شده بودند، گردآوری و مقایسه شدند.[35]



شکل ۸- آنتولوژی پایه برای آگاهی موقعیت[33]

۲-۳- خلاصه فصل

در این فصل به بررسی و معرفی کارهای پیشین مرتبط با این تحقیق پرداخته شد. کارهای پیشین در دوسته کلی معرفی شدند. دسته اول کارهای مبتنی بر مدل JDL بودند که خود در سه زیردسته جا می گیرند، زیردسته اول در سطوح مدل پایه تغییرات ایجاد کرده اند، زیردسته دوم روند داده امیزی مدل را بررسی کرده اند و زیردسته سوم با تعریف نگاشت مناسب، این مدل را به حوزه های دیگر برده اند.

[1] Situation awareness

دسته کلی دوم کارهای داده امیزی معنایی هستند که از تکنولوژی های معنایی در حوزه دادهآمیزی بهره می برند. برخی از این کارها مربوط به ارائه یک معماری معنایی برای داده آمیزی می باشند و برخی دیگر با تاکید بر استفاده از تعریف آنتولوژی ها برای دامنه خاص موردنظرشان به رفع تداخل و ناهمگونی پرداخته اند. گروهی از محققان نیز با ارائه آنتولوژی های سطح بالا، ساختاری مشترک را برای سطحی از داده آمیزی فراهم آورده بودند.

فصل ۳– روش پیشنهادی

کاربردهای گسترده داده‌آمیزی در حوزه های مختلف باعث شده است که تا کنون مدل ها و معماری های بسیاری برای داده آمیزی ارائه شود. مدل ها در واقع چارچوب مرجعی را برای توصیف، درک و تشخیص انواع مسائل مربوط به داده امیزی فراهم می کنند.

ما در این تحقیق برآنیم در جهت رفع چالش های موجود در داده امیزی، ساختاری معنایی بر مبنای مدل موجود JDL ارائه کنیم. همانطور که در بخش ۲-۱-۱ در توضیح مدل JDL اشاره شد این مدل شامل ۵ سطح می باشد که با افزودن معنا و گنجاندن آنتالوژی و استفاده از تکنولوژی های وب معنایی، آن را به یک مدل معنایی گسترش دهیم و یک ساختار داده آمیزی معنایی ارائه دهیم.

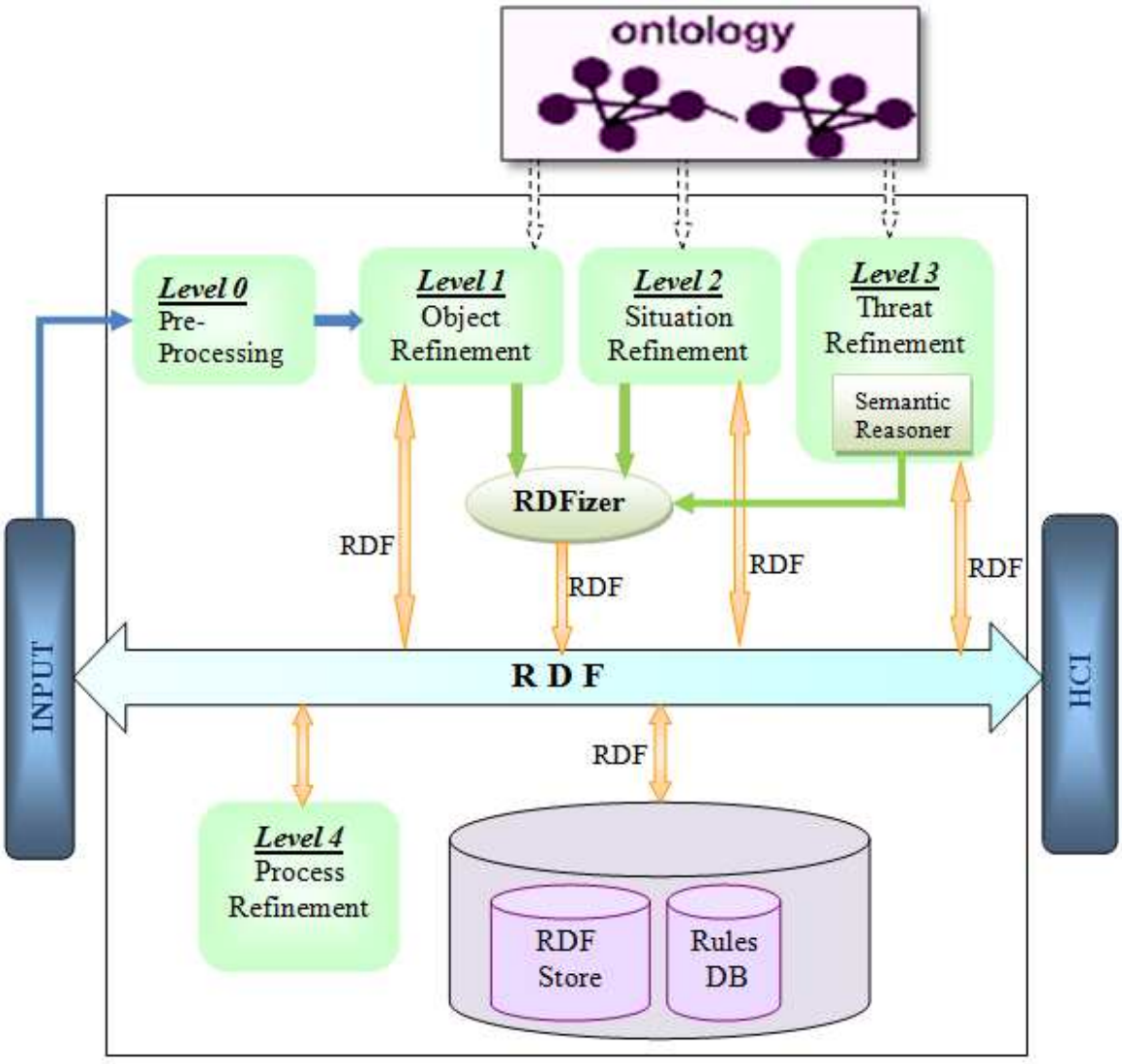
در این فصل در بخش ۳-۱ مدل پیشنهادی ارائه می شود و سپس در بخش ۳-۲ به سنجش و ارزیابی مدل می پردازیم و بخش ۳-۳ مراحل عملیاتی شدن مدل را توضیح می دهد.

۳–۱– معرفی مدل پیشنهادی

در بخش ۲-۱-۱ با مدل JDL و سطوح آن آشنا شدیم. مدل پیشنهادی ما ترکیب مدل JDL پایه با تکنولوژی های وب معنایی می باشد، نمای کلی ساختار پیشنهادی در شکل ۹ مشاهده می شود.

بطور خلاصه عملکرد ساختار را میتوان بدین شرح توصیف نمود: ابتدا سیگنالهای ورودی ارزیابی شده و پس از انجام عمل پیش پردازش برای رفع ناهنجاریها و مشکلات داده ها، به فرمت RDF تبدیل می شوند. برای این منظور با استفاده از آنتولوژی تعریف شده، مجموعه صفات هر شی به برنامه RDFizer داده می شود تا به فرمت RDF تبدیل و در پایگاه داده ذخیره شوند. سپس براساس اطلاعات موجود و دانش پیشین موقعیت اشیا و روابط

انها مشخص می شوند و سپس تهدیدها شناسایی می شوند که در این مرحله وجود یک استدلالگر معنایی[1] و مجموعه قوانین ضروری است. در ادامه جزئیات هر سطح از مدل پیشنهادی توضیح داده می شود.



شکل ۹- مدل پیشنهادی برای داده آمیزی معنایی بر مبنای مدل JDL

[1] Semantic reasoner

سطح صفر: (پیش پردازش) در سطح صفر عمل پیش پردازش روی داده های منابع مختلف صورت می گیرد و داده های پاکسازی شده در اختیار سطح بعدی که پالایش شی است، قرار می گیرد. از این رو در این سطح معنا وارد نشده است.

سطح اول: (پالایش شی) سطح ۱ نیز به طور کامل معنایی نمی شود، یه عبارتی ما در این سطح فقط اشیا و صفات انها را به فرمت استاندارد RDF تبدیل و ذخیره می کنیم، و اما وظایف دیگر این سطح مانند شناسایی اشیا[1] و نسبت دهی صفات به اشیا معنایی نشده و با الگوریتم های ریاضی و مباحث پردازش تصویر[2] انجام می شود، در واقع ما در این سطح، داده های موجود را معنایی نمایش می دهیم تا در سطوح بالاتر از آن ها استفاده شود. از این رو به کمک آنتولوژی‌های از پیش تعریف شده، اشیا به همراه صفات کشف شده‌شان به فرمت استاندارد RDF، درمی آیند و با کمک RDFizer، داده های مربوط به اشیا به همراه صفات مرتبط شان به فرمت RDF در آمده و در پایگاه داده RDFStore ذخیره می‌شوند.

بنابراین واضح است که در این مدل نیاز است برای تمام مفاهیم حوزه کاری مرتبط، آنتولوژی تعریف شود تا در این سطح با توجه به آنتولوژی تعریف شده، اشیا به همراه صفات مرتبط شان به صورت ساختار یافته مقداردهی می شوند.

سطح دوم: (پالایش موقعیت) در این سطح با توجه به دانش پیشین و اطلاعات محیطی و داده های گرداوری شده در سطح یک (اطلاعات مربوط به موجودیت ها به همراه صفاتشان به فرمت RDF)، موقعیت اشیا و روابط صریح بین آن ها تعیین می گردد. (قبلا رابطه های ممکن بین اشیا تعریف شده است).

سپس در ادامه‌ی وظایف این سطح، بر اساس روابط کشف شده، یک سری صفات مربوط به اشیا که تا کنون نامعلوم بودند، مشخص می شوند، همچنین به کمک روابطی که صریحا بین اشیا پیدا شده است، به کمک استنتاج روابط جدیدی را از روی روابط و صفات موجود استخراج می شود.[36]

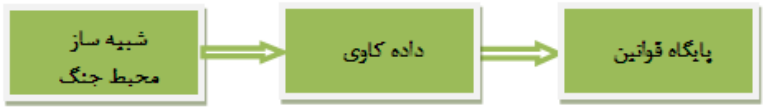
[1] Object recognition
[2] Image processing

بنابراین مشخص است که در این سطح به اطلاعات جدیدی دست می یابیم یا اطلاعات قبلی تغییر پیدا می کنند، از این رو اطلاعات استنتاجی جدید در مورد موقعیت به کمک RDFizer به فرمت RDF درامده و در RDFStore ذخیره می شود.

سطح سوم: (پالایش تهدید) ارزیابی موقعیت جاری و پیش بینی تهدید ها و آسیب پذیری ها از وظایف این سطح است. بنابراین در این سطح وجود یک استدلالگر معنایی ضروری می باشد. در این سطح با استفاده از تکنیک های استنتاج معنایی و استدلالگر تعبیه شده و با در نظر گرفتن قوانین موجود در پایگاه قوانین، وضعیت فعلی ارزیابی و استنتاج کاملی جهت پیش بینی تهدید ها انجام خواهد گرفت.

سپس با توجه به اینکه دراوردن راه حل ها و فرصت عمل ها در موقعیت بوجود آمده و برای تهدیدهای شناخته شده بر عهده این سطح می باشد، لیستی از فرصت عمل ها نیز به عنوان خروجی خواهیم داشت. در این سطح آنتولوژی هایی برای تهدیدها و فرصت عمل ها تعریف می شود.[36] خروجی های این سطح به RDFizer می رود تا پس از تبدیل به فرمت RDF در RDFStore قرار گیرد.

برای به دست آوردن قوانین یک حوزه، می توان یک شبیه ساز طراحی و پیاده سازی نموده و با استفاده از داده کاوی، قوانین را استخراج و به پایگاه قوانین اضافه نمود. به عنوان مثال برای استخراج قوانین محیط جنگ میتوان این ۳ مرحله را همانطور که در شکل ۱۰ آورده شده، انجام داد.



شکل ۱۰- استخراج قانون

و یا برای پر کردن پایگاه قوانین میتوان از افراد خبره در حوزه موردنظر کمک گرفت و نهایتا بایستی قوانین را با یک زبان استاندارد موجود برای تعریف قوانین مانند SWRL۱[37] توصیف شوند. همچنین در صورتی که

[1] Semantic Web Rule Language

به سیستم قابلیت یادگیری افزوده شود، پایگاه قوانین امکان به روز رسانی، حذف یا اضافه کردن قانونی را خواهد داشت.

سطح چهارم: (پالایش فرآیند) در سطح ۴ نظارت بر کارایی سیستم و تخصیص منابع بر اساس هدف تعیین شده صورت می‌پذیرد. در این سطح فرد خبره با مشاهده خروجی های سطح ۳ تصمیمات لازم را می گیرد و تغییرات کلی در سیستم جهت بهبود کارایی اعمال می کند.

مدیریت پایگاه داده ها : در ساختار داده آمیزی پیشنهادی چندین پایگاه داده وجود دارد. یک پایگاه داده برای ذخیره سازی اطلاعات که به شکل RDF می باشند در نظر گرفته شده است که شامل شمای RDF نیز می باشد و انتولوژی ها را در خود ذخیره دارد. برای قوانین موجود در دامنه تعیین شده نیز یک پایگاه قوانین خواهیم داشت. برای تمام این پایگاه داده ها بایستی اصول مدیریت پایگاه داده لحاظ شود و موارد لازم در زمینه مدیریت پایگاه داده ها مانند سازگاری، به روز رسانی، درج و حذف و ویرایش با سرعت مطلوب در نظر گرفته شود.

برخی دیگر از جزئیات مدل و ورودی ها و خروجی های دقیق بسته به حوزه ای که برای آن مدل پیاده سازی می شود، قابل تعریف است که در فصل ۴ برای حوزه نظامی توضیح داده خواهد شد.

۳-۲- ارزیابی مدل

در این قسمت مدل پیشنهادی بر اساس روش های سنجش مدل، در ۴ زیربخش موردارزیابی قرار خواهد گرفت.

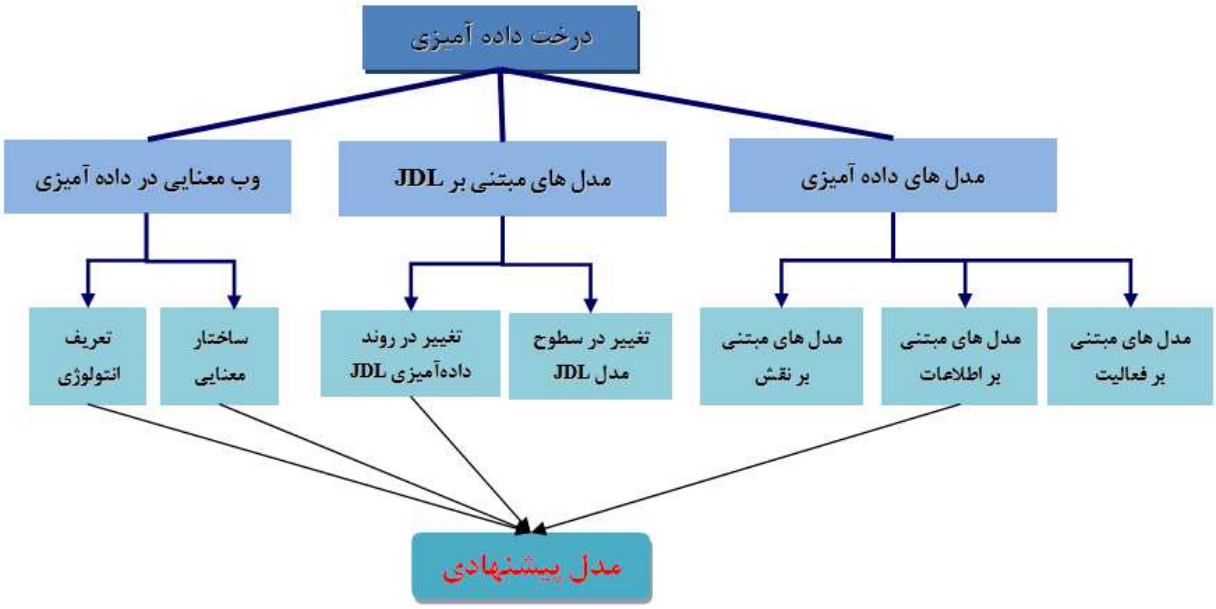
۳-۲-۱- ارزیابی مقایسه ای[1]

ابتدا برای واضح شدن جایگاه مدل پیشنهادی در مقایسه با سایر مدل ها، یک درخت داده آمیزی مرتبط با زمینه این تحقیق طراحی شده و موقعیت مدل پیشنهادی در ان بین مشخص شده است که در شکل ۱۱

[1] Comparison evaluation

مشاهده می شود. همانطور که از شکل برمیآید مدل پیشنهادی با توجه به اینکه براساس مدل JDL می باشد همانند این مدل، جزء مدل های مبتنی بر اطلاعات برشمرده می شود و در مقایسه با کارهای انجام شده مبتنی بر JDL، مدل پیشنهادی شامل سطوح مدل پایه JDL بوده و تغییری شامل ادغام یا افزودن یا حذف سطوح را اعمال نکرده است، اما مدل پیشنهادی باعث تغییر پروسه داده آمیزی می شود و روند ترکیب داده به صورت معنایی انجام می شود.

در مقایسه با کارهای معنایی داده امیزی، مدل پیشنهادی، یک ساختار معنایی برای داده آمیزی ارائه می کند که بر مبنای مدل موجود JDL است. در این مدل آنتولوژی های لازم جهت ایجاد یکپارچگی و ساخت یافتگی بایستی تعریف شوند و مراحل بعدی شامل پالایش موقعیت و تهدید مبتنی بر آنتولوژی انجام خواهند شد.



شکل ۱۱- جایگاه مدل پیشنهادی

در ادامه با توجه به اینکه مدل پیشنهادی بر پایه مدل پایه JDL می باشد، به مقایسه مدل ارائه شده با مدل JDL نیز پرداخته شده است. همانطور که در بخش ۳-۱ توضیح داده شد، مدل پیشنهادی با تعریف آنتولوژی ها،

به مفهوم سازی صریح می پردازد و داده های موجود در مدل، همه به فرمت استاندارد RDF هستند و همه داده ها همراه با ابرداده ذخیره شده اند. در جدول ۱ مدل پیشنهادی با مدل پایه JDL مقایسه شده است.

جدول1 - مقایسه مدل پیشنهادی با مدل JDL

	مدل JDL	**مدل پیشنهادی**
مفهوم‌سازی حوزه مدنظر	-	+
فرمت استاندارد برای توصیف داده ها	-	+
وجود ابرداده برای داده ها	-	+
امکان استنتاج معنایی	-	+
قابلیت استخراج دانش	-	+
حذف تداخل های معنایی در روند ترکیب داده	-	+
سازگاری در محیط‌هایی با منابع داده‌ای ناهمگون	-	+
حذف ناهمگونی ها در نحو، ساختار و معنا	-	+

۳-۲-۲- ارزیابی مبتنی بر ویژگی[1]

امروزه یک روند رو به رشد از تکنولوژی های وب معنایی در حال شکل گیری می باشد که در زمینه های مختلفی از علوم وارد شده است. از این رو در این تحقیق معنا به حوزه داده امیزی اضافه شد که ویژگی های جدید مثبت و منفی را با خود به همراه داشت.

یکی از ویژگی های اصلی مدل، ایجاد سازگاری و همگونی در نحو و ساختار و معنا می باشد، که دلیل اصلی ان مفهوم سازی حوزه مدنظر می باشد و اینکه تمام مفاهیم به صورت ساخت یافته در قالب انتولوژی طراحی و پیاده سازی می شوند. بنابراین در این مدل نیاز است که انتولوژی های حوزه موردنظر کامل تعریف و پیاده سازی شود، که این خود، مسائل قابل توجه زیادی به همراه دارد، چرا که بایستی مواردی مانند استانداردسازی آنتولوژی، مهندسی آنتولوژی، ارزیابی آنتولوژی، تطابق آنتولوژی ها را در نظر گرفت و تکامل و نگهداری انتولوژی

[1] Feature-based evaluation

نیز قابل توجه و اهمیت می باشد. درست است که انتولوژی ها در این مدل، مزایای بسیاری با خود به همراه می اورند، اما نمیتوان پیچیدگی کار با آن ها را نادیده گرفت.

در این مدل ماهیت داده ها معنایی می باشد و پرس و جو روی این داده ها به کمک زبان پرس وجوی SPARQL[38] انجام می گیرد. نکته ای که در این جا قابل ذکر می باشد، این است که کارایی این زبان پرس وجو برای مواقعی که پایگاه داده معنایی بسیار بزرگ باشد، پایین است و در مقایسه با SQL[1] کندتر عمل می کند.

ویژگی اصلی این مدل، معنایی شدن بود که سایر ویژگی ها مانند استخراج دانش و موارد جدول ۱ حاصل افزودن معنا می باشد.

۳-۲-۳-ارزیابی عملیاتی[2]

با توجه به اینکه در این تحقیق برای اولین بار این مدل پیشنهاد می شود، بایستی عملیاتی بودن ان را نیز بررسی کرد و نشان داد که این مدل جواب می دهد و کار می کند. از این رو با انتخاب حوزه نظامی و طراحی سناریو نشان خواهیم داد که خروجی های مطلوب هر سطح به صورت معنایی، قابل استخراج است و روشی برای انجام هر سطح با برنامه نویسی معنایی[3] و مبتنی بر آنتولوژی وجود دارد. در فصل ۴ به صورت مفصل به این مورد پرداخته خواهد شد.

۳-۲-۴-ارزیابی مبتنی بر کاربر[4]

با توجه به اینکه حوزه انتخابی این تحقیق برای پیاده سازی مدل پیشنهادی، حوزه نظامی می باشد که اطلاعات این زمینه را بیشتر از طریق افراد خبره و متخصص در همین زمینه میتوان کسب نمود، (به دلیل محرمانگی این حوزه، مطالب به صورت مقاله یا مستندات انتشار نمی یابد.) ما در طول انجام این کار اطلاعات لازم برای ساخت

[1] Structured Query Language
[2] Operability evaluation
[3] Semantic programming
[4] User-based evaluation

انتولوژی و خروجی سطوح را از افراد خبره کسب نمودیم و پس از اینکه پیاده سازی مدل تمام شد، این مدل در سیستم موجود در وزارت دفاع قرار داده شد و این افراد که کاربران و افراد خبره در این زمینه هستند، کارایی مدل را تایید کردند.

۳-۳- مراحل عملیاتی کردن مدل پیشنهادی

برای استفاده از مدل پیشنهادی برای یک حوزه و پیاده سازی آن، گام‌های لازم عبارتند از:

- جمع آوری مجموعه داده[1] مناسب : برای پیاده سازی پس از انتخاب حوزه، می بایست مجموعه داده ی ورودی را تامین کرد که در صورت امکان می تواند واقعی باشد و در غیر این صورت میتوان با تولید یک مجموعه داده ازمایشگاهی مدل را پیاده‌سازی نمود.
- تعریف تمام آنتولوژی های حوزه مدنظر: آنتولوژی ها بیانی صریح و رسمی از ویژگی های مفاهیم هستند. برای پیاده سازی این مدل نیاز است که تمام آنتولوژی های حوزه مدنظر تعریف شود. در گام های بعدی پیاده سازی، از این انتولوژی ها استفاده می شود از این رو بهتر است این گام با دقت کافی انجام شود تا انتولوژی های کامل در اختیار مراحل بعدی قرار بگیرد.
- تولید مجموعه قوانین و بیان قوانین با یک زبان استاندارد برای تعریف قوانین معنایی: در طول روند داده امیزی به مجموعه قوانین حوزه مدنظر نیاز داریم، از این رو به کمک شبیه ساز و داده کاوی و یا به کمک فرد خبره می بایست قوانین لازم را دراورد و سپس به زبان معنایی آنها را نوشته و در پایگاه قوانین قرار داد.
- پیاده سازی RDFizer: همانطور که تا کنون گفته شده است، در این مدل داده ها به فرمت RDF هستند بنابراین در تمام سطوح برای تبدیل خروجی های جدید نیاز به RDFizer داریم تا داده ها را به فرمت RDF درآورد تا تمام اطلاعات موجود در ساختار، معنایی و یکنواخت باشند.

[1] Data set

- پالایش شئ و موقعیت مبتنی بر آنتولوژی: این دو گام باید با کمک آنتولوژی های ساخته شده در گام تعریف انتولوژی، مبتنی بر آنتولوژی انجام شوند و به کمک برنامه نویسی معنایی، خروجی های لازم استخراج شود.
- بکارگیری استنتاج معنایی جهت پالایش تهدیدها: با توجه به اینکه ساختار ارائه شده کاملا معنایی می باشد و دانش موجود و پایگاه داده ها نیز به فرمت معنایی نمایش داده شده اند، استنتاج نیز به شیوه معنایی انجام می شود. با بکارگیری استنتاج معنایی، به نتایج لازم و نیز دانش جدید خواهیم رسید.

۳-۴- خلاصه فصل

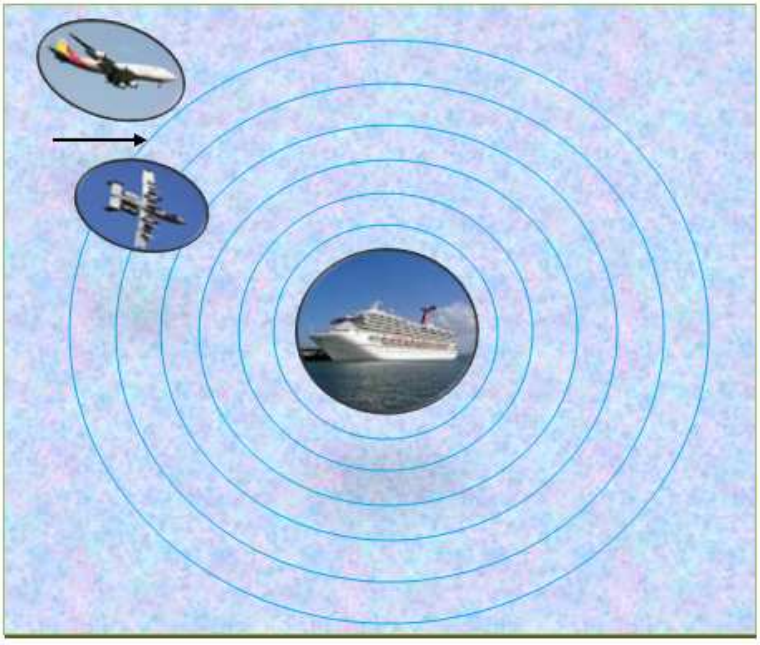
در این فصل به معرفی مدل پیشنهادی پرداخته شد و سطوح آن و معنایی شدن سطوح توضیح داده شد. سپس ارزیابی مدل در چهار بخش ارزیابی مقایسه ای، عملیاتی، مبتنی بر ویژگی و مبتنی بر کاربر بررسی شد و در ناهیت در بخش ۳-۳ گام هایی که جهت پیاده سازی مدل می بایست اجرا شود، توضیح داده شد.

فصل۴– پیاده‌سازی مدل

با توجه به این که مدل پایه JDL ذاتا برای محیط های جنگی ارائه شده است، در این فصل به پیاده سازی مدل پیشنهادی در حوزه نظامی می پردازیم و با طراحی سناریو برای این حوزه، عملیاتی شدن مدل را نشان خواهیم داد. در این فصل در بخش ۴-۱ به طراحی سناریو و تولید مجموعه داده پرداخته می شود و در بخش ۴-۲ آنتولوژی های این حوزه طراحی می شوند و ورودی و خروجی های سطوح ساختار پیشنهادی در بخش ۴-۳ بررسی و پیاده سازی می شوند و در بخش آخر این فصل پیچیدگی زمانی مربوط به اجرای مدل اورده می شود.

۴–۱– طراحی سناریو و مجموعه داده متناظر

برای پیاده سازی و نمایش عملکرد چارچوب پیشنهادی، سناریویی نظامی به کمک افراد خبره در این زمینه طراحی شد و سپس سطوح مدل برای این سناریو پیاده سازی شد. شکل ۱۴ سناریو را نمایش می دهد.



شکل ۱۲- سناریو

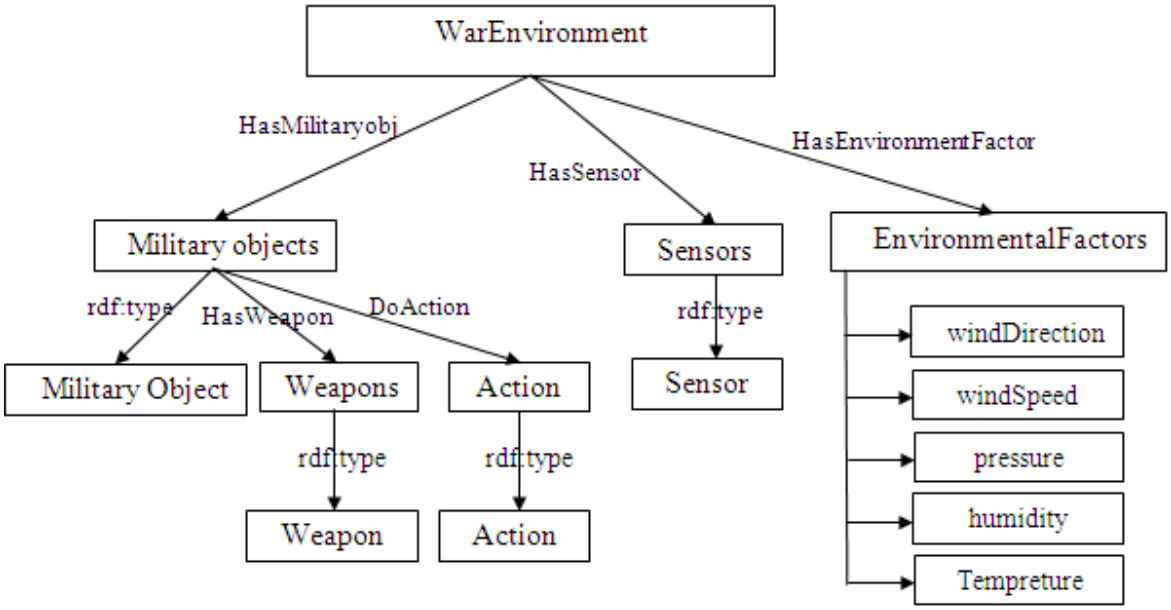
همانطور که در شکل مشخص است، در این سناریو یک کشتی درنظر گرفته شده است که ثابت می باشد و از تجهیزات خودی است که تعدادی جنگ افزار مانند موشک نیز در ان برای مواقع درگیری قرار داده شده است و دو هواپیما که به سمت راست در حال حرکت اند و هواپیمای اول دشمن است و هواپیمای نوع دوم هنوز ناشناخته می باشد. متناطر با این سناریو، مجموعه داده آزمایشگاهی تولید شد که به مدت ۱۰ ثانیه، ادامه این سناریو را در بردارد. در مجموعه داده آزمایشگاهی، هر رکورد داده که از سمت سنسورها می آید همراه با برچسب زمانی می باشد، از این رو هر رکورد که یک خط از فایل ورودی می باشد به فرمت زیر است:

date time, property, value, object, sensorID*

ابتدای هر رکورد برچسب زمانی قرار دارد و سپس صفتی که این مقدار برای آن گزارش می شود و در فیلد سوم مقدار داده قرار می گیرد و سپس شی ای که این صفت مربوط به آن می باشد و در آخرین فیلد، شناسه سنسوری که این داده را گزارش داده، آمده است.

۴-۲- طراحی و ساخت آنتولوژی ها

همانطور که در فصل گذشته گفته شد، بایستی در حوزه داده آمیزی مورد نظر، انتولوژی ها را تعریف نمود، از این رو در این سناریو نیاز به طراحی و پیاده سازی آنتولوژی های محیط جنگ وجود دارد. شکل ۱۳ آنتولوژی لایه بالای محیط جنگ را نشان می دهد که در ادامه انتولوژی های هر جزء توضیح داده خواهد شد.



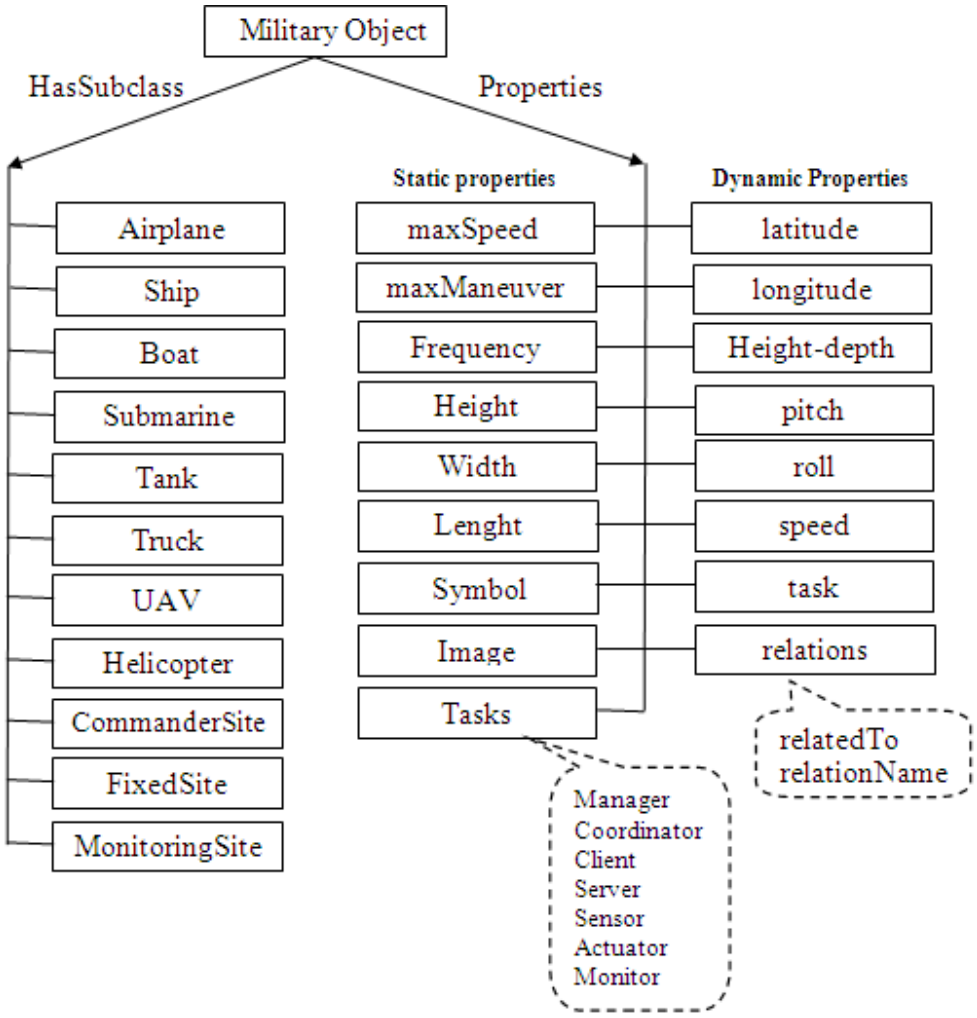
شکل ۱۳ - آنتولوژی محیط جنگ

در محیط تعدادی اشیای جنگی قرار دارند که هر کدام تعدادی جنگ افزار با خود دارند و همچنین این اشیا می توانند عملی[1] انجام دهند. تعدادی سنسور در محیط قرار دارد که مشخصات خود را دارند که در ادامه در انتولوژی مربوط به سنسور مشاهده خواهد شد. همچنین تعدادی از پارامترهای محیطی که برای شرایط جنگ اهمیت دارند، شامل سرعت و جهت باد، میزان رطوبت، فشار و دما در این انتولوژی کلی در نظر گرفته شده است. در ادامه با جزئیات آنتولوژی ها آشنا می شویم:

- آنتولوژی اشیای جنگی: در محیط جنگ در محدوده موردنظر تعدادی اشیای جنگی وجود دارد که نوع انها MilitaryObject می باشد. بنابراین کلاسی با نام MilitaryObject طراحی می شود که شامل اشیایی مانند کشتی، کامیون، هلیکوپتر و غیره به همراه صفات انها می باشد. صفات می توانند ایستا یا پویا باشند. منظور از صفات ایستا مواردی مانند طول، عرض، ارتفاع، حداکثر سرعت است که برای همه اشیایی که از این نوع باشند ثابت بوده و در طول جنگ تغییر نمی کند و بیشتر دربرگیرنده موارد شناسایی اشیا می باشد. از طرفی صفات پویا مواردی مانند طول و عرض جغرافیایی و موقعیت و سرعت

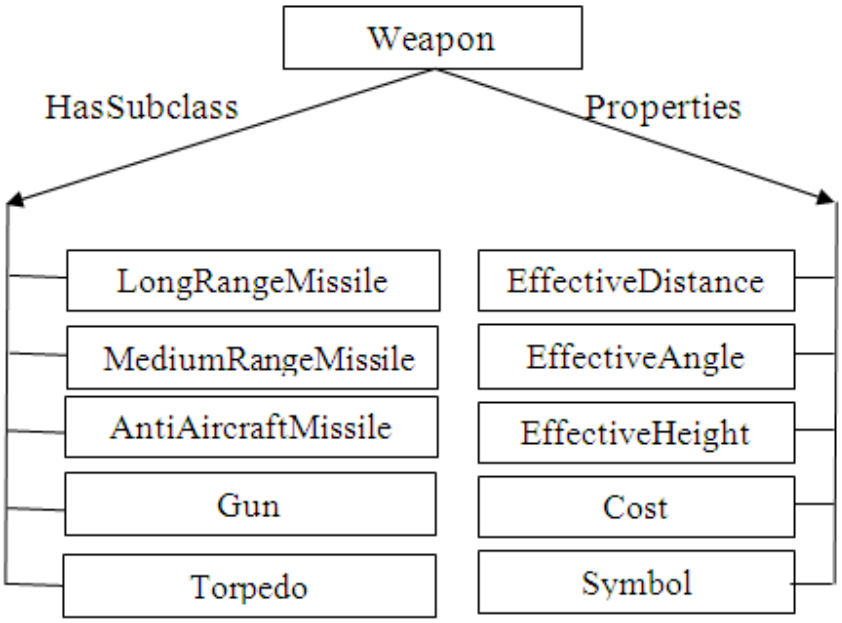
[1] action

کنونی است که در طول زمان متغیر می باشد. شکل ۱۴ انتولوژی مربوط به اشیای جنگی را نشان می دهد.



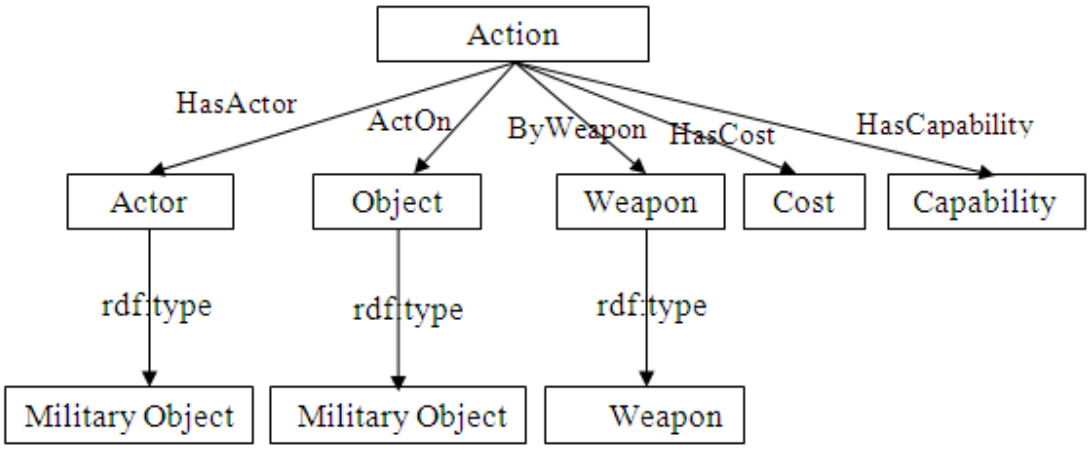
شکل ۱۴- انتولوژی اشیای جنگی

- آنتولوژی جنگ افزارها: برای جنگ افزارها و اسلحه های مختلف و صفات ان ها نیز کلاسی درنظر گرفته شده است که شامل مواردی مانند موشک و تفنگ به همراه صفاتی مانند فاصله و زاویه موثر و هزینه می باشد. شکل ۱۵ آنتولوژی جنگ‌افزارها را نشان می دهد.



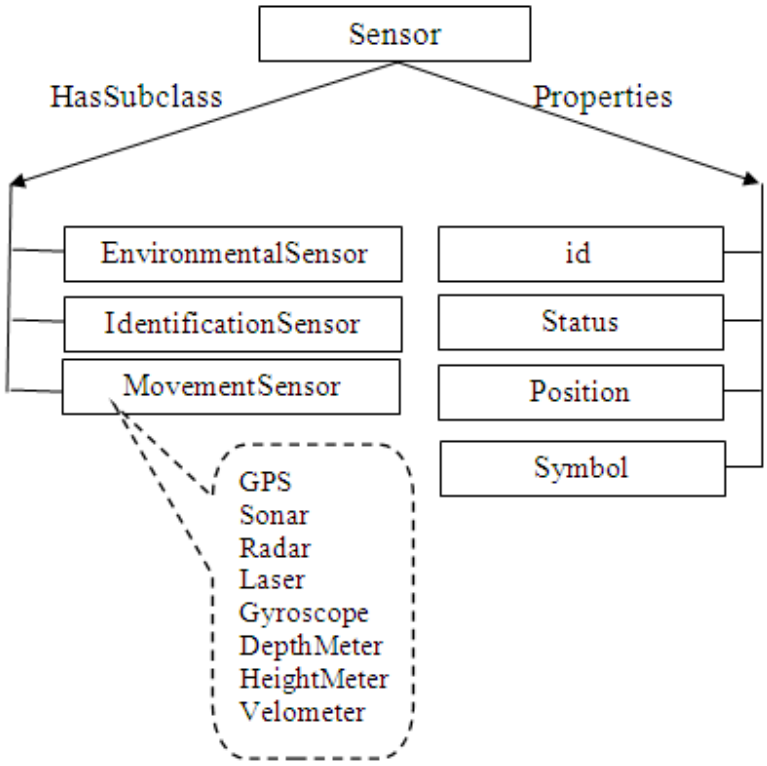
شکل ۱۵- آنتولوژی جنگ افزارها

- آنتولوژی عمل: در هر عملی که یک شی جنگی انجام می دهد و در نقش فاعل قرار می گیرد باید مواردی شامل اینکه مفعول چیست و به کمک چه جنگ افزاری، فعل انجام می شود، لحاظ شود، شکل ۱۶ کلاس عمل را نشان می دهد.



شکل ۱۶- آنتولوژی عمل

- آنتولوژی سنسور: سنسورهای محیط جنگ به سه دسته سنسورهای محیطی، سنسورهای شناسایی و سنسورهای حرکتی تقسیم می شوند. شکل ۱۷ مدل آنتولوژی سنسور و ویژگی های آن را نشان می دهد.



شکل ۱۷ - آنتولوژی سنسور

تمام آنتولوژی های فوق به کمک کتابخانه jena [39] در محیط netbeans پیاده سازی شد. jena یک کتابخانه جاوای متن‌باز است که برای پردازش داده های RDF در وب معنایی و داده های پیوندی بکار می رود. این کتابخانه علاوه بر امکان ایجاد مدل های مختلف آنتولوژی شامل RDFS و OWL، پردازش پرس‌وجوهای SPARQL و استنتاج مبتنی بر قانون را نیز حمایت می کند.

۴-۳- ورودی و خروجی سطوح مختلف مدل پیشنهادی

در سطح صفر مدل که داده امیزی سطح پایین (سطح سیگنال و پیش پردازش) می باشد تغییر معنایی ایجاد نشده است و در ادامه سطوح ۱ و۲ و۳ به صورت معنایی پیاده سازی و شرح داده می شوند.

۴-۳-۱- پالایش اشیا

در این سطح تمام اشیا و موجودیت های سناریو و ویژگی های آن ها و داده های مجموعه‌داده، به صورت سه‌تایی های RDF در RDFStore ذخیره می شوند. داده های موجود در RDFStore را با یک مثال نشان می دهیم. به عنوان مثال کلاس کامیون یک زیر کلاس از MilitaryObject می باشد و ویژگی های ایستای این کلاس که در بخش قبل آورده شد، شامل مواردی مانند حداکثر سرعت می باشد. شکل ۱۸ داده های مربوط را به صورت RDF/XML نمایش می دهد. از طرفی ویژگی های پویا مانند موقعیت کنونی را داریم که همراه با برچسب زمانی ذخیره می شوند، شکل ۱۹ یک نمونه را نشان می دهد.

```
<rdfs:Class rdf:about=http://wtlab.um.ac.ir/war/Truck>
    <rdfs:subClassOf>
        <rdfs:Class rdf:about=http://wtlab.um.ac.ir/war/MilitaryObject >
    </rdfs:subClassOf>
    <war:MaxSpeed rdf:parseType="Resource">
        <war:Unit>KilometerInHour</war:Unit>
        <war:Value rdf:datatype=http://www.w3.org/2001/XMLSchema#positiveInteger >
        200 </war:Value>
    </war:MaxSpeed>
</rdfs:Class>
```

شکل ۱۸- کلاس کامیون و یک نمونه ویژگی ایستای آن به فرمت RDF/XML

```
<war:Truck rdf:about="http://wtlab.um.ac.ir/war/truck1">
  <war:PositionX rdf:parseType="Resource">
      <war:Value rdf:datatype=http://www.w3.org/2001/XMLSchema#float >
      22</war:Value>
      <war:Time rdf:datatype=http://www.w3.org/2001/XMLSchema#dateTime >
```

```
      2011-07-20T03:32:54</war:Time>
    <war:SNRid>914</war:SNRid>
  </war:PositionX>
```

شکل ۱۹ - بیان ویژگی پویای موقعیت کامیون به فرمت RDF/XML

۴-۳-۲- پالایش موقعیت

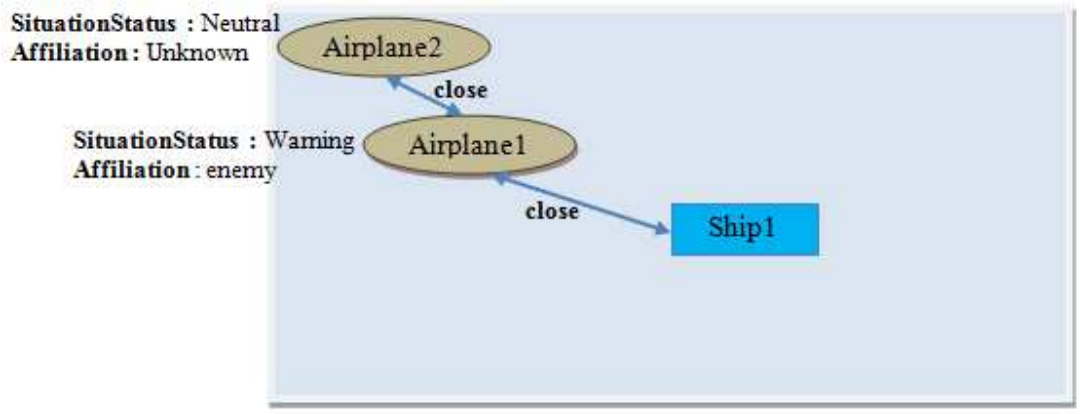
در این سطح با توجه به دانش پیشین و اطلاعات محیطی و داده های موجود در RDFStore، موقعیت اشیا و روابط بین آن ها تعیین می شود و به طور کلی یک دید کامل از موقعیت بایستی فراهم شود. در این سطح برای استنتاج و کشف روابط یا ویژگی های ضمنی و غیرصریح، قوانین معنایی روی RDFStore اعمال می شود.

اکنون به پالایش موقعیت برای سناریوی مدنظر شامل سه مورد زیر می پردازیم:

- درآوردن روابط صریح بین اشیا
- درآوردن ویژگی های نامعلوم اشیا به کمک روابط و صفات کشف شده
- درآوردن روابط جدید از روی روابط معلوم به کمک استنتاج

مورد اول با کمک محاسباتی مانند فاصله، روابط بین اشیا مانند رابطه نزدیک شدن استخراج می شود. خروجی برای این سناریو در این گام در شکل ۲۰ مشاهده می شود که در آن براساس حدآستانه[1] تعریف شده، رابطه نزدیک شدن بین هواپیمای ۱ و۲ و همچنین بین هواپیمای ۱ و کشتی ۱ پیدا شده است.

[1] threshold

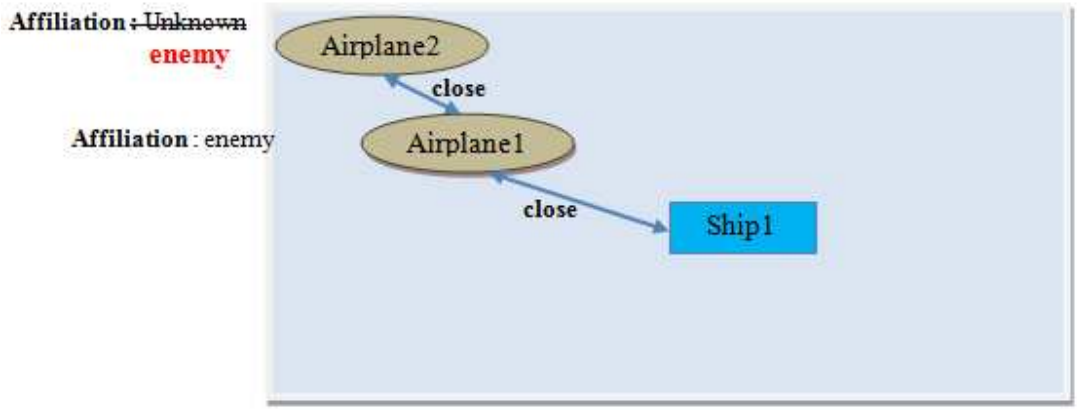


شکل ۲۰- پالایش موقعیت: کشف روابط بین اشیا

در مورد دوم می خواهیم ویژگی های نامعلوم را به کمک روابط و صفات کشف شده، مشخص نماییم. در این سناریو هواپیمای ۱، از نوع دشمن می باشد، یعنی ویژگی affiliation آن enemy می باشد اما هنوز هویت هواپیمای ۲ مشخص نیست و Unknown می باشد. به کمک قانون معنایی نمایش داده شده در شکل ۲۱ نتیجه می‌شود که هواپیمای ۲ از نوع دشمن می باشد و خروجی تا این لحظه برای سطح دوم شکل ۲۲ خواهد بود.

```
[Rule1:     (?s http://wtlab.um.ac.ir/war/IsEnemy 'true')
              (?s http://wtlab.um.ac.ir/war/Relation ?r)
              (?r http://wtlab.um.ac.ir/war/RelationName 'close')
              (?r http://wtlab.um.ac.ir/war/RelatedTo ?o)
              (?o http://wtlab.um.ac.ir/war/IsEnemy 'UnKnown')
             ->
             (?o http://wtlab.um.ac.ir/war/IsEnemy 'true')
]
```

شکل ۲۱- قانون معنایی برای استنتاج ویژگی دشمن بودن



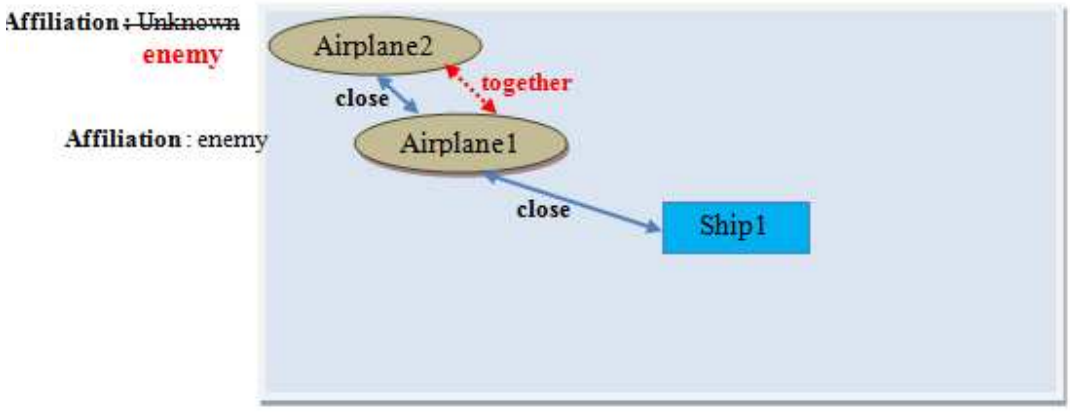
شکل ۲۲- پالایش موقعیت: کشف ویژگی نامعلوم

در مورد سوم می خواهیم روابط جدیدی را از روی روابط موجود استخراج نماییم. در این سناریو رابطه "با هم بودن یا هم گروه بودن[1]" با توجه به سایر روابط استنتاج می شود. تعریف این قانون موجود در پایگاه قوانین در شکل ۲۳ آورده شده است. خروجی نهایی سطح ۲ در شکل ۲۴ مشاهده می شود.

```
[Rule2:   (?s http://wtlab.um.ac.ir/war/IsEnemy 'true')
          (?s http://wtlab.um.ac.ir/war/Relation ?r)
          (?r http://wtlab.um.ac.ir/war/RelationName 'close')
          (?r http://wtlab.um.ac.ir/war/RelatedTo ?o)
          (?o http://wtlab.um.ac.ir/war/IsEnemy 'true')
          ->
          (?r http://wtlab.um.ac.ir/war/RelationName 'together')
]
```

شکل ۲۳- پالایش موقعیت: کشف روابط جدید با توجه به سایر روابط و ویژگی ها

[1] Together relation



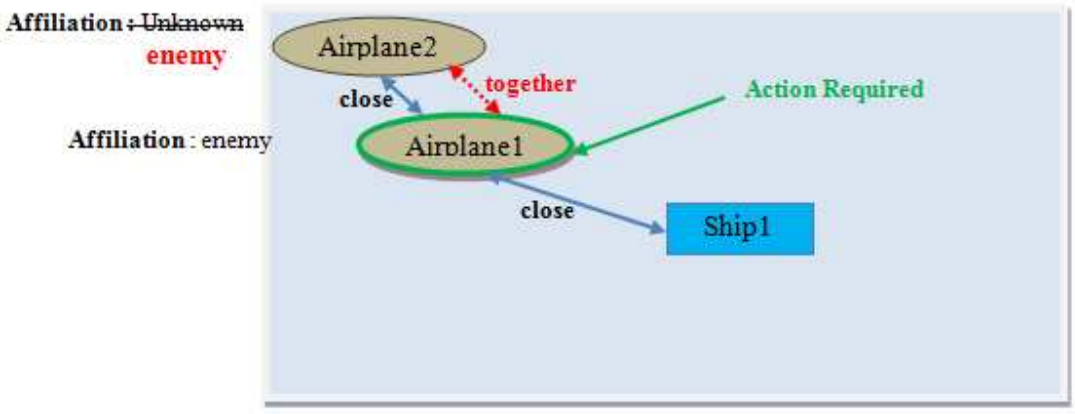
شکل ۲۴- پالایش موقعیت

۴-۳-۳-پالایش تهدید

همانطور که از قبل گفته شده است، ارزیابی موقعیت جاری و پیش بینی تهدید ها و آسیب پذیری ها و فرصت عمل ها از وظایف این سطح است. در این سناریو ابتدا با کمک یکی از قوانین پیش بینی می شود که هواپیمای شماره ۱ خطرناک است و تهدید به حساب می آید و نیاز داریم که یک عمل در رابطه با این شی اجرا شود. قانون در شکل ۲۵ مشاهده می شود و شکل متناظر با سناریو در شکل ۲۶ اورده شده است.

```
[Rule3:   (?s http://wtlab.um.ac.ir/war/IsEnemy 'true')
          (?s http://wtlab.um.ac.ir/war/SituationStatus 'Warning')
          ->
         (http://wtlab.um.ac.ir/war/Action http://wtlab.um.ac.ir/war/ActOn ?s)
]
```

شکل ۲۵- قانون **ActionRequired**



شکل ۲۶- سناریو در مرحله پالایش تهدیدها

سپس تمام اعمال ممکن یا به عبارتی فرصت عمل ها در سناریو بررسی می گردد و در جدولی به عنوان خروجی این سطح ارائه می شود. در این جدول فاعل[1] و مفعول[2] عمل به همراه سه فیلد دیگر جمع آوری شده است. یکی از این فیلدها جنگ افزاری است که در فاعل موجود بوده و می تواند عمل را با ان انجام دهد. در فیلد هزینه، پارامترهای زیادی می تواند موثر باشد و محاسبه شود، در این سناریو این فیلد با توجه به هرینه جنگ افزار محاسبه شده است. فیلد قابلیت با توجه به موقعیت فاعل و مفعول و فاصله‌موثر جنگ افزار محاسباتی انجام می دهد و اگر در فاصله موجود امکان زدن با جنگ افزار موردنظر باشد، مقدار مثبت می گیرد. جدول مربوط به این سناریو در جدول ۲ آورده شده است.

جدول2 - گردآوری فرصت عمل ها

Actor	ActOn	ActByWeapon	Cost	Capability
Ship	Airplane1	MediumRangeMissile	5	+
Ship	Airplane1	LongRangeMisssile	9	+
Ship	Airplane1	Gun	1	-

[1] actor
[2] actOn

۴-۴- زمان اجرا

زمان اجرای سطوح و گام های مختلف در جدول ۳ در کامپیوتری با حافظه اصلی ۱ گیگابایت گزارش می شود.

جدول ۳- زمان اجرای گام های سناریو

شماره گام	فعالیت انجام شده	زمان اجرا (ms)
۱	تولید شمای RDFS و ذخیره آن در RDFStore	۱۷۰۱
۲	خواندن dataset از فایل و تبدیل به فرمت RDF	۱۶۴
۳	انجام کامل سطح شناسایی موقعیت	۱۷۳
۴	انجام کامل سطح شناسایی تهدید	۶۲

زمان گزارش شده برای گام ۱ ،جهت ساخت تمام انتولوژی های این زمینه و ذخیره ان ها در فایل می باشد که بسته به تعداد انتولوژی های حوزه ای که تعریف می شود و اندازه این آنتولوژی ها، زمان اجرا متفاوت خواهد بود. زمان گام ۲ برای تبدیل داده ها به فرمت RDF توسط RDFizer است که به اندازه فایل ورودی وابسته می باشد. در گام ۳ تمام عملیات مربوط به پالایش موقعیت شامل کارهای محاسباتی و استنتاجی جهت کشف روابط انجام می شود که بسته به اندازه فایل قوانین و پیچیدگی سناریو زمان اجرا تغییر خواهد کرد. در گام ۴ پس از شناسایی تهدیدها، جدول فرصت عمل ها نیز استخراج و نمایش داده می شود که زمان گزارش شده، زمان استنتاج برای تشخیص تهدیدها و سپس در آوردن فرصت عمل ها می باشد.

فصل۵– نتیجه گیری و کارهای آینده

در این تحقیق جهت رفع ناهمگونی های معنایی در مدل داده آمیزی JDL، تکنولوژی های وب معنایی را به ان اضافه کردیم و ساختار پیشنهادی یکی از چالش های اصلی موجود در سیستم‌های داده آمیزی، که چالش معنایی می باشد را برطرف نمود و ناهمگونی های موجود در نحو، ساختار و معنا را از سیستم های داده امیزی حذف نمود. سپس مدل پیشنهادی را برای حوزه نظامی ارزیابی و پیاده سازی نمودیم.

کارهای آتی در این مدل دیده می شود که جای بررسی دقیق و پیاده سازی دارد. برخی از این موارد در زیر برشمرده می شود.

- افزودن یادگیری به مدل: اگر برای مدل پیشنهادی موضوع یادگیری افزوده و پیاده سازی شود، پیش بینی می شود کیفیت خروجی ها به مراتب بهتر شود، البته ناگفته نماند که برای یادگیری نیاز به یک مجموعه داده جمع اوری شده برای اموزش[1] می باشد.
- توسعه مدل برای محیط های چندعامله: میتوان این مدل را با شناخت کامل عامل های محیط مدنظر و وظایف هر عامل، برای محیط چندعامله توسعه داد و پیاده سازی کرد.

طراحی اولیه ای از معماری چندعامله برای جنگ های شبکه مدار[2] بر اساس مدل ارائه شده، طراحی و پیشنهاد شده است که در ان عامل های سیستم براساس تنوع ادوات جنگ تعریف می شود. از این رو ساختار داخلی هر عامل متفاوت است و هر عامل باید بتواند با سایر عاملها همگن[3] و ناهمگن[4] تبادل اطلاعات داشته باشد. این ارتباط براساس پروتکل تعریف شده ای انجام می شود و فرمت داده ها براساس RDF های دارای برچسب زمانی است. با توجه به ساختار سلسله مراتبی محیط جنگ، این

[1] training
[2] Network Centric Warfare-NCW
[3] Homogeneous agents
[4] Heterogeneous agents

چارچوب دارای یک ساختار لایه ای است که وظایف عاملها در هر لایه متفاوت خواهد بود. همچنین مجموعه ای از عامل های همگن و ناهمگن درهرلایه تحت مدیریت واحد یک عامل لایه بالاتر و براساس پروتکل های ارتباطی تعریف شده عمل می کنند. جزئیات کامل این طراحی در پیوست ۲ امده است. در ادامه کار پیاده سازی محیط چندعامله به همراه پروتکل های ارتباطی معنایی انجام خواهد شد.

- امنیت مدل: این مدل از جنبه های مختلفی نیاز به بررسی بحث های امنیتی دارد. از جمله مهمترین این موارد عبارتند از:
  - امنیت پایگاه داده ها شامل RDFStore
  - امنیت تبادل داده ها و اطلاعات
  - (در محیط چندعامله) امنیت ساختار داخلی عامل ها
  - (در محیط چندعامله) امنیت در پروتکل ارتباطی عامل ها
- پیاده سازی برای سایر حوزه ها: میتوان این مدل را برای محیط های دیگری که به گونه ای به این سطوح داده امیزی نیاز دارند، پیاده سازی نمود، مانند خانه های هوشمند، روباتیک یا مدیریت بحران.

منابع

[1] F. E. White, “Data Fusion Lexicon”, Data Fusion Subpanel of the Joint Directors of Laboratories Technical Panel for C^3, Code 4202, NOSC, San Diego, CA(1991)

[2] Hall, D. L., Llinas, J., ”An introduction to multi-sensor data fusion”, Proc. IEEE 85, (1997)

[3] E.F. Nakamura, A.A. Loureiro and A.C. Frery, “Information Fusion for Wireless Sensor Networks: Methods, Models, and Classifications”, ACM Computing Surveys, Vol. 39, No. 3(2007)

[4] Dasarathy, B. V., “Sensor fusion potential exploitation-innovative architectures and illustrative applications”, Proc. IEEE 85(1997)

[5] Boyd, J. R., “A discourse on winning and losing”, Unpublished set of briefing slides available at Air University Library, Maxwell AFB, Alabama(1987)

[6] Bedworth, M. D., O’brien, J. C., “The omnibus model: A new model for data fusion”, In Proceedings of the 2nd International Conference on Information Fusion, ISIF, Sunnyvale, 437–444(1999)

[7] C.B. Frankel, M. Bedworth, “Control, Estimation and Abstraction in Fusion Architectures: Lessons from Human Information Processing”, Proc. Of FUSION’00, Paris(2000)

[8] Jinsoo Park, and Sudha Ram, “Information systems interoperability: What lies beneath?”, ACM Transactions on Information Systems, Vol 22, No. 4, pp. 595-632(2004)

[9] M. Gagnon, “Ontology-based Integration of Data Sources”, 10th international Conference on Information Fusion, Quebec, Canada(2007)

[10] H.G. Cheng, “Representing and Reasoning about Semantic Conflicts in Heterogeneous Information Sources”, doctoral dissertation, Sloan School of Management, MIT, Cambridge, Mass(1997)

[11] Steinberg A.N., Bowman C.L., White, F.E., “Revisions to the JDL Data Fusion Model”, in Sensor Fusion: Architectures, Algorithms, and applications, Proceedings of the SPIE,Vol .3719(1999)

[12] J. Esteban, A. Starr, R. Willetts, P. Hannah, and P. Bryanston-Cross, “A review of data fusion models and architectures: Towards engineering guidelines,” Neural Computing and Applications, Vol. 14, No. 4(2005)

[13] Blasch, E and S. Plano, “JDL Level 5 fusion model: user refinement issues and applications in group tracking”, SPIE Vol 4729, Aerosense, pp. 270 – 279(2002)

[14] Blasch, E.P. and Plano, S., “Level 5: User Refinement to aid the Fusion Process”, in Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications, Proc of the SPIE, Vol. 5099(2003)

[15] E. Blasch, "User refinement in Information Fusion", Chapter 19 in Handbook of Multisensor Data Fusion 2nd Ed, Eds. D. Hall, and J. Llinas, CRC Press(2008)

[16] D. Lambert, “Grand Challenges of Information Fusion”, Proc Sixth Int. Conf. Information Fusion, pages 570- 574, Cairns, Australia, pages 213-220(2003)

[17] D. A. Lambert. Situations for situation awareness. In Proceedings of Fusion 2001, 4-th International Conference on Information Fusion(2001)

[18] J. Llinas, C. L. Bowman,G.Rogova, A. N. Steinberg, E. Waltz, F. E. White, “Revisiting the JDL data fusion model II”, In Proceedings of the 7th International Conference on Information Fusion, Stockholm, Sweden(2004)

[19]] Dr. John J. Salerno, “Where’s Level 2/3 Fusion – a Look Back over the Past 10 Years”, In 10th International Conference on Information Fusion(2007)

[20] Polychronopoulos, A. Amditis, U. Scheunert, T. Tatschke, “Revdisiting JDL model for automotive safety applications: the PF2 functional model”, The 9th International Conference on Information Fusion, Florence, Italy(2006)

[21] Steinberg, A.N., Bowman, C.L., White, F.E., “Rethinking the JDL Data Fusion Model”, to appearin NSSDF Conference Proceedings, JHAPL(2004)

[22] Giacobe, N. “Application of the JDL data fusion process model for cyber security”, Proc. SPIE 7710, 77100R(2010)

[23] J. Synnergren, J. Gamalielsson, B. Olsson, “Mapping of the jdl data fusion model to bioinformatics”, IEEE Explore, 44:1506-1511(2007)

[24] Leo J.DeVin n, MagnusHolm, AmosH.C.Ng, “The Information Fusion JDL-U model as a reference model for Virtual Manufacturing”, Journal Robotics and Computer-Integrated, Volume 26, Issue6 (2010)

[25] Grigoris Antoniou and Frank van Harmelen, “Semantic web primer”, the second edition(2008)

[26] Zafeiropoulos, N. Konstantinou, S. Arkoulis, D. Spanos, N. Mitrou, “A semantic-based architecture for sensor data fusion”, In 2nd International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies(2008)

[27] F. Manola, E. Miller, B. McBride. RDF Primer, W3C Rec. 10(2004)

[28] I.F. Cruz, H. Xiao, “The role of ontologies in data integration”, Journal of Engineering Intelligent Systems, 13 (4) (2005)

[29] N. Guarino. Formal Ontology and Information Systems. In Proceedings of the 1st International Conference on Formal Ontologies in Information Systems (FOIS 1998), pages 3–15(1998)

[30] H. Wache, T. V¨ogele, U. Visser, H. Stuckenschmidt, G. Schuster, H. Neumann, and S. H¨ubner, “Ontology-Based Integration of Information - A Survey of Existing

Approaches”, In Proceedings of the IJCAI-01 Workshop on Ontologies and Information Sharing(2001)

[31] Y. Arens, C. A. Knoblock, and C. Hsu, “Query Processing in the SIMS Information Mediator”, In The AAAI Press(1996)

[32] E. Mena, V. Kashyap, A. P. Sheth, and A. Illarramendi, “OBSERVER: An Approach for Query Processing in Global Information Systems based on Interoperation across Pre-existing Ontologies”, In Proceedings of the 1st IFCIS International Conference on Cooperative Information Systems (CoopIS 1996), pages 14–25(1996)

[33] C. J. Matheus, M. M. Kokar, and K. Baclawski, “A core ontology for situation awareness”, Proc. of the 6th International Conference on Information Fusion, Cairns, Queensland, Australia, 545-552(2003)

[34] N. Baumgartner, W. Gottesheim, S. Mitsch, W. Retschitzegger, and W. Schwinger, “BeAware!—situation awareness, the ontology-driven way”, International Journal of Data and Knowledge Engineering, 69(11):1181–1193(2010)

[35] Baumgartner, N.; Retschitzegger, W., “A Survey of Upper Ontologies for Situation Awareness”, In: Proceedings of Knowledge Sharing and Collaborative Engineering (2006)

[36] Alan N. Steinberg, " Foundations of Situation and Threat Assessment", Chapter 18 in Handbook of Multisensor Data Fusion 2nd Ed, Eds. D. Hall, and J. Llinas, CRC Press(2008)

[37] Ian Horrocks, Peter F. Patel-Schneider, Harold Boley, Said Tabet, Benjamin Grosof, Mike Dean,“SWRL: A Semantic Web Rule Language Combining OWL and RuleML”, W3C Rec(2004)

[38] E. Prud’hommeaux and A. Seaborne, “SPARQL Query Language for RDF”, W3C Candidate Rec. (2006)

[39] Shelley Powers, "jena: RDF in java ", Chapter 8 in book of Practical RDF, 1st edition(2003)

[40] GeeWah Ng, "Intelligent systems: fusion, tracking and control", Research Studies Press Ltd(2003)

[41] Shahbazian E, Blodgett D, Labbe P, "The Extended OODA Model for Data Fusion Systems", Proceedings of 4th International Conference on Information Fusion(2001)

پیوست‌ها

۶-۱- پیوست ۱- مدل های داده آمیزی

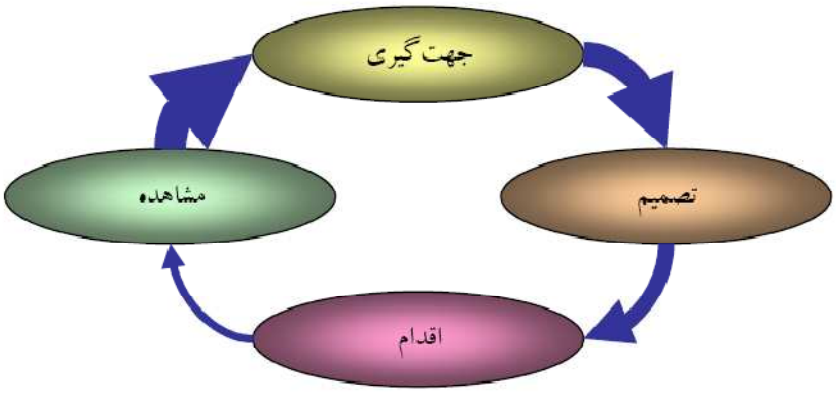
تلفیق داده یک فرایند بسیار پیچیده بوده و بنابراین به شدت نیازمند به یک مدل کلی بوده که وجوه و فعالیت‌های مختلف از فرآیند تلفیق داده را به یکدیگر مرتبط سازد. در اینجا دو مدل حلقه کنترل بوید و مدل جامع را مطرح می‌کنیم.

۱- **مدل حلقه کنترل بوید:** یک مدل چرخه ای است که دارای چهار مرحله مشاهده[1]، جهت گیری[2]، تصمیم[3] و اقدام[4] است. در مرحله مشاهده داده های حسگری جمع آوری می شوندو در مرحله جهت گیری، داده ها در جهت فهم موقعیت با هم ترکیب می شوند. در مرحله تصمیم گیری، طرحی برای پاسخ دهی فراهم می شود و در نهایت در مرحله اقدام، طرح فراهم شده اجرا می شود.[41][40][7]

همانطور که حلقه کنترلی در شکل ۲۹ نمایش داده شده است، در این شکل عرض فلش هابیان گر میزان جریان داده ها در حلقه است، هر چه که داده در حلقه پالایش شود و داده به اطلاعات تبدیل شود، میزان داده های انتقال یافته به فاز بعد کاهش می یابد.

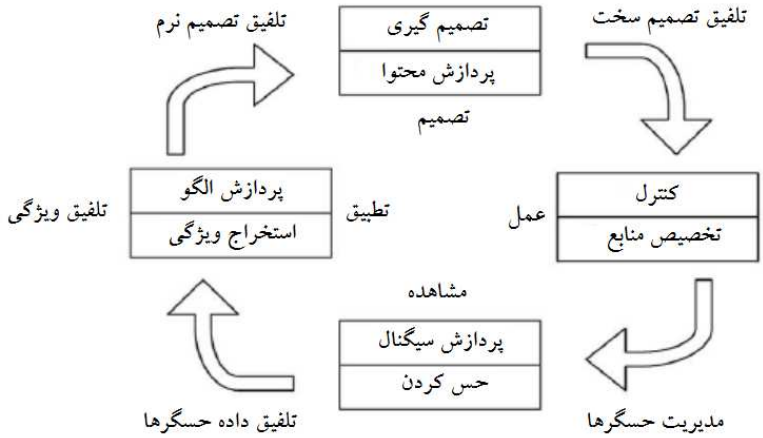
فاز مشاهده قابل مقایسه با مرحله پیش پردازش منابع در مدل JDL و فاز جهت گیری که می‌توان آنرا مشابه کاربردهای سطوح ۱، ۲ و ۳ مدل JDL در نظر گرفت. همچنین فاز تصمیم گیری را می‌توان متناظر با سطح ۴ یعنی ارزیابی فرآیند در مدل JDL در نظر گرفت. مدل JDL فاز اقدام ندارد.

[1] observe
[2] orient
[3] decide
[4] act



شکل ۲۷- حلقه کنترلی بوید[40]

۲- **مدل جامع**[1]: مدل جامع بیشتر ویژگیهای سودمند سایر روش‌ها را به کار می‌گیرد. همانطور که در شکل ۲۸ نشان داده شده است. مدل جامع از یک طبیعت دوره ای بهره می‌برد که آن را قابل مقایسه با مدل حلقه کنترل بوید می‌کند. سطوح مختلف این مدل عبارتند از: (۱) گام‌های پردازش سیگنال و فرآیند حس کردن که مشابه با فاز مشاهده در حلقه بوید است، (۲) پردازش الگو و استخراج ویژگی مشابه با فاز جهت گیری، (۳) پردازش محتوا و تصمیم گیری متناظر با فاز تصمیم گیری و (۴) کنترل و تخصیص منابع که متناظر با فاز اقدام می‌باشد. این مدل بسیار ساده بود و پیاده سازی مناسبی از آن را برای بسیاری از کاربردهای داده آمیزی غیر دفاعی و غیر نظامی می‌توان بکار برد.[6]



شکل ۲۸- مدل جامع[6]

[1] Omnibus model

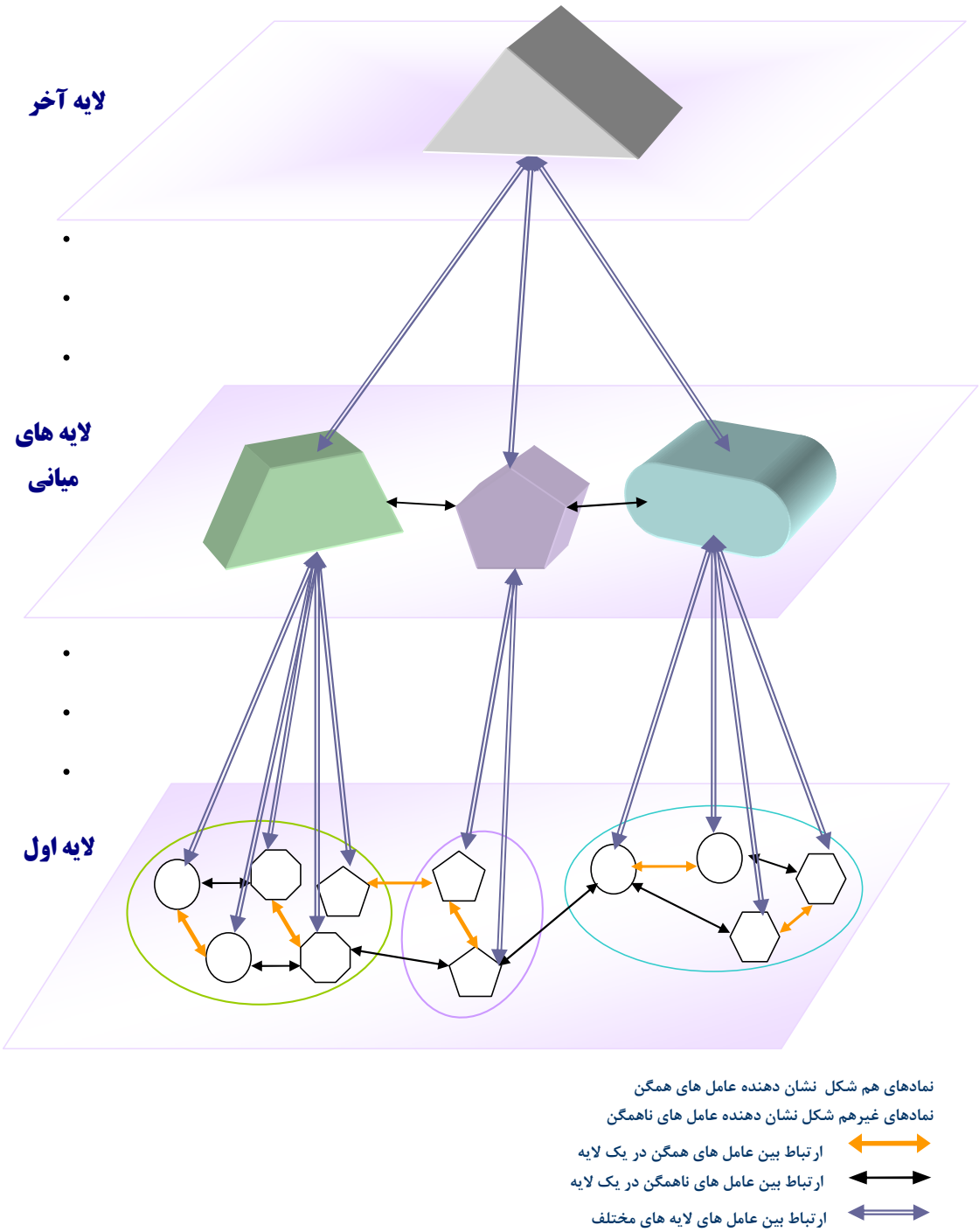
۶-۲- پیوست ۲- معرفی محیط چندعامله مدل پیشنهادی

سیستم های چندعامله از مجموعه ای از عامل ها و روابط بین آن ها تشکیل شده است. در این سیستم ها هر عامل وظیفه خود را می داند و توانایی ارتباط با دیگر عامل ها را دارد. عامل ها می توانند همگن یا ناهمگن باشند. عامل های همگن همه از یک نوع بوده و دانش حوزه ای یکسان دارند، درصورتیکه عامل های ناهمگن ساختار داخلی یکسانی ندارند. با وجود پیچیدگی هایی که برای پیاده سازی سیستم های چندعامله ناهمگن وجود دارد، توانایی این نوع سیستم ها به مراتب بیشتر از سیستم های همگن است.

چارچوب پیشنهادی یک سیستم چندعامله است که دارای مجموعه از عامل های همگن و ناهمگن بوده و ارتباط این عامل ها براساس پروتکل های تعریف شده می باشد. همچنین، هر عامل قابلیت یادگیری دارد که این یادگیری در تعامل با یک محیط پویا بدست می آید. عامل در نتیجه برخی رفتارهایی که وضعیت محیط را بهتر می کند پاداش می گیرد و در نتیجه رفتارهایی که باعث بدشدن وضعیت محیط می شود، جریمه می شود و این وظیفه عامل است که با تجربه کردن ترکیب های مختلف رفتارها تشخیص دهد که کدام رفتار تحت چه شرایطی منجر به پاداش بیشتری می شود و انتخاب بهتری کند.

شکل ۲۹ معماری محیط چندعامله چارچوب پیشنهادی را نشان می دهد. در این شکل هریک از اشکال هندسی نمایانگر یک نوع عامل هستند. عامل های همگن با شکل هندسی یکسان نشان داده شده اند. مجموعه ای از عامل های همگن و ناهمگن تحت مدیریت عامل لایه بالاتر عمل می کنند. ارتباط بین عامل های همگن با یکدیگر، ارتباط بین عامل های ناهمگن و همچنین ارتباط بین عاملهای لایه های مختلف باهم متفاوت است.

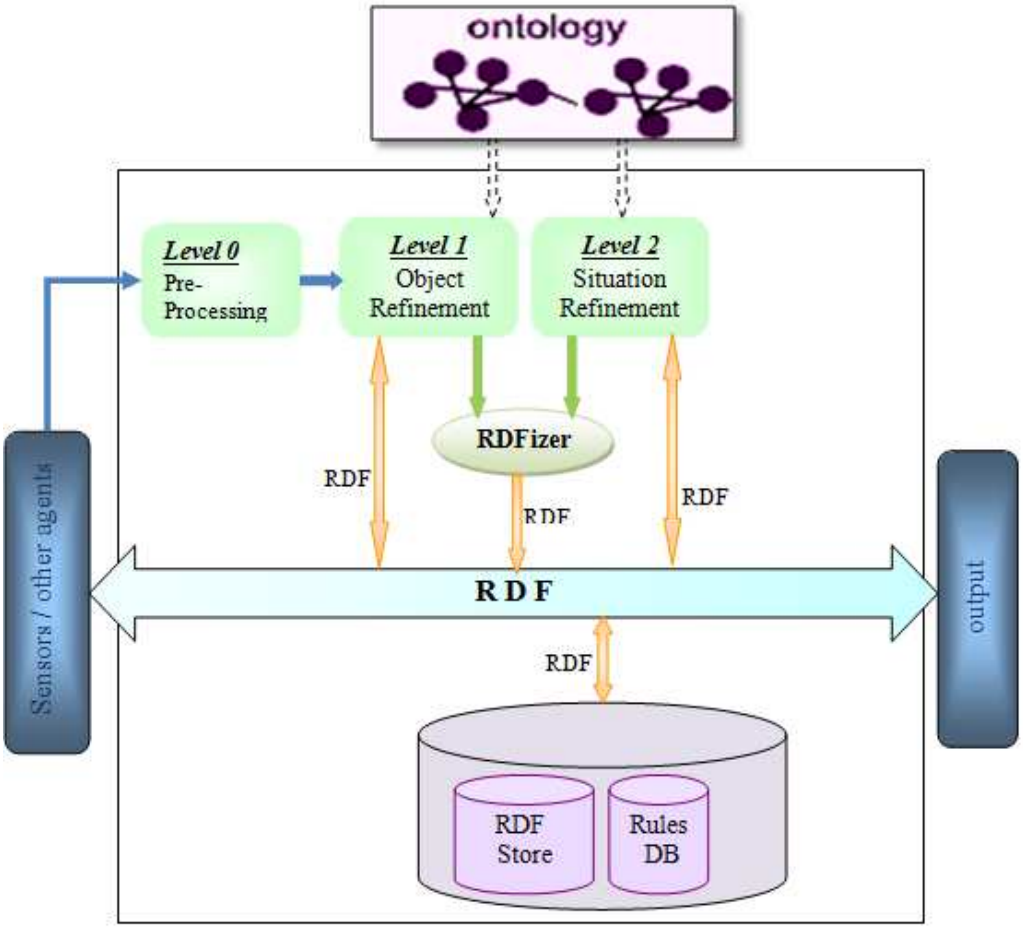
جزئیات مربوط به ساختار داخلی عامل ها در ادامه توضیح داده خواهد شد.



شکل ۲۹- معماری لایه ای محیط چندعامله چارچوب پیشنهادی

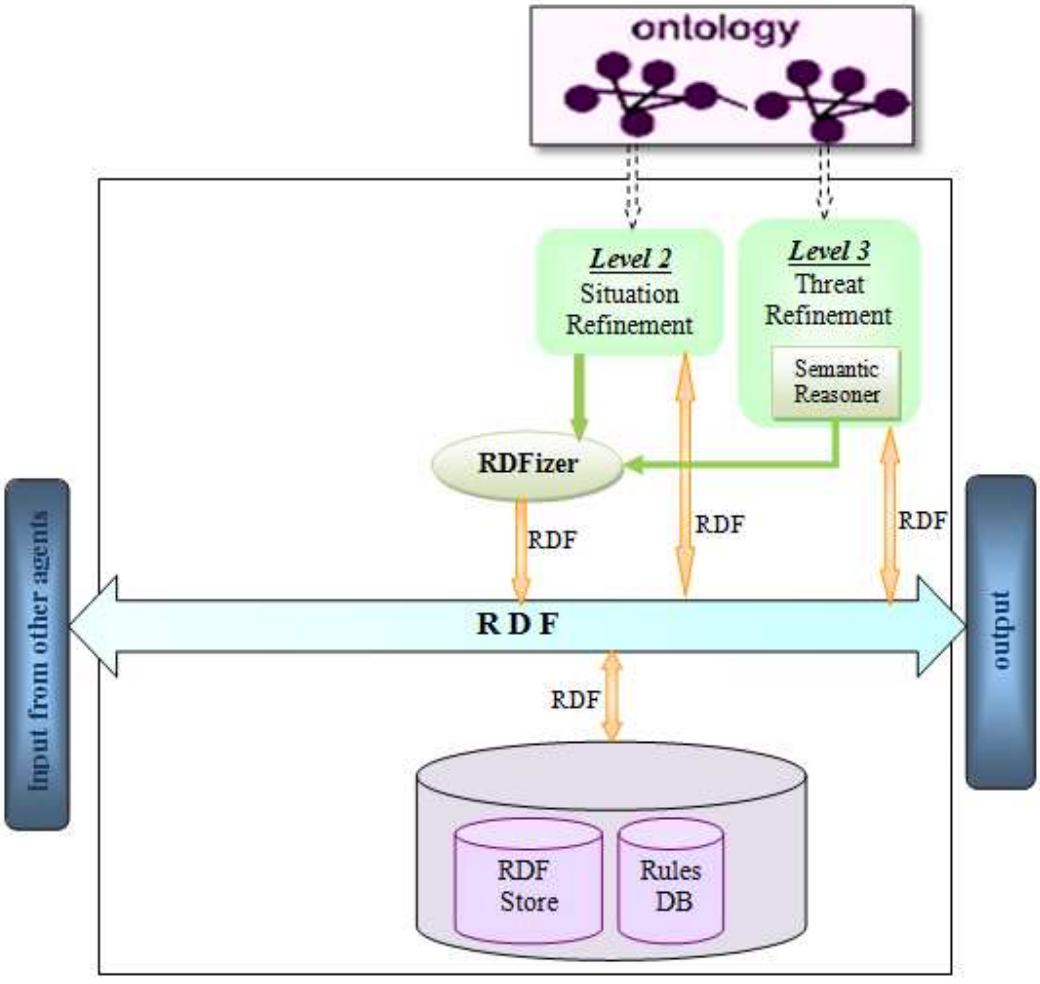
ساختار داخلی هر عامل برگرفته از مدل داده آمیزی معنایی پیشنهادی مبتنی بر JDL می باشد. با توجه به اینکه ساختار داخلی هر عامل در هر لایه متفاوت خواهد بود، در این بخش ساختار عامل ها برای هر لایه مشخص می شود.

در لایه اول سیگنالهای ورودی پس از ارزیابی و انجام عمل پیش پردازش، باید به فرمت RDF تبدیل شوند. برای این منظور با استفاده از آنتولوژی تعریف شده، مجموعه صفات هر عامل بهمراه ویژگی های تعیین کننده موقعیت آن به RDFizer داده می شود تا به فرمت RDF تبدیل شوند. سپس، سه تایی های RDF در پایگاه داده ذخیره می‌شوند. عامل های این لایه تا حدی ارزیابی موقعیت برای خودشان انجام می دهند. شکل ۳۰ ساختار داخلی عامل های لایه اول را نشان می دهد.



شکل ۳۰- ساختار داخلی عامل ها در لایه اول

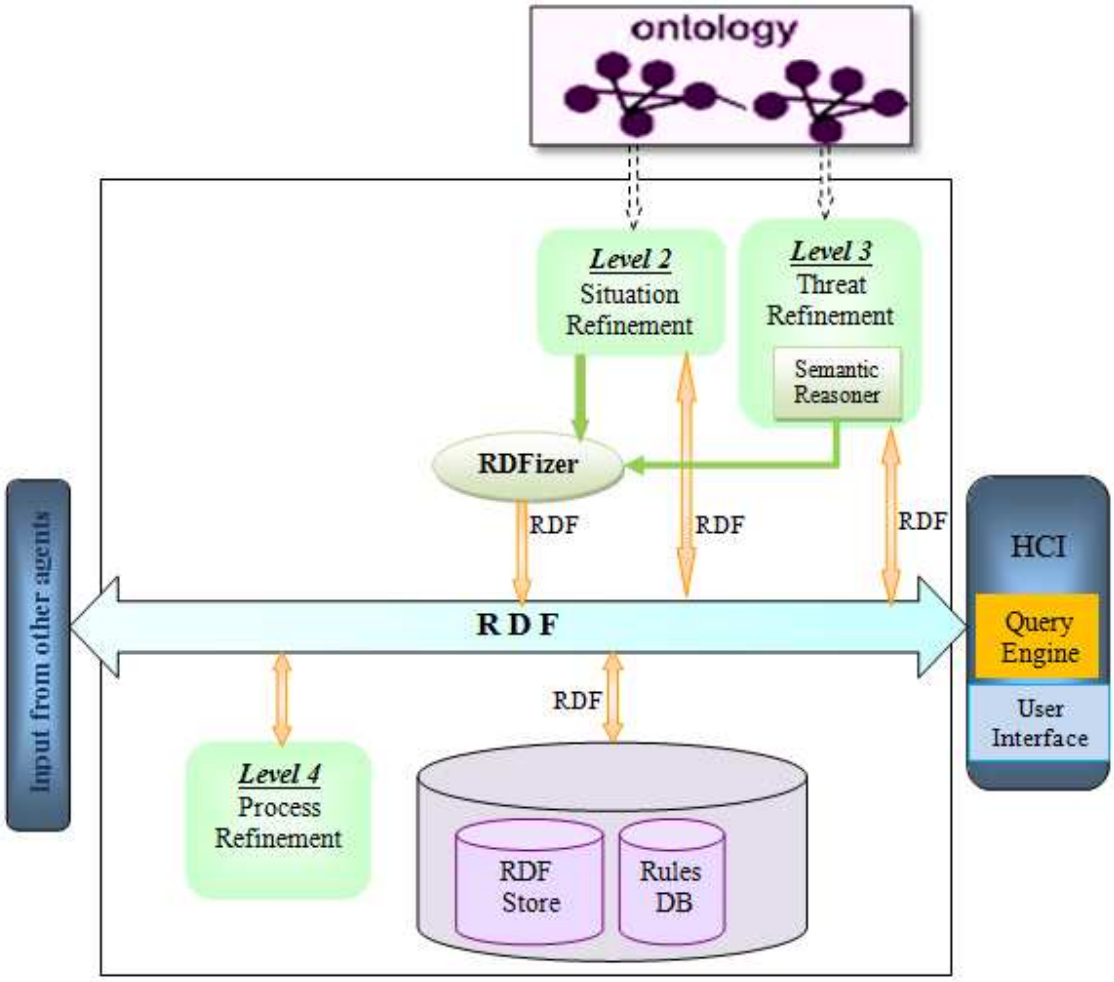
در لایه های میانی، هر عامل براساس موقعیت خود و بر مبنای اطلاعات دریافتی از سایر عامل ها تهدیدها را شناسایی می کند. همانطور که در شکل ۲۹ آمده است، در این لایه مجموعه ای از عامل های همگن و ناهمگن تحت مدیریت واحد یک عامل لایه بالاتر و براساس پروتکل های ارتباطی تعریف شده عمل می کنند. عامل های این لایه از طریق ورودی/خروجی و با استفاده از پروتکل های تعریف شده می توانند با سایر عامل ها تبادل اطلاعات داشته باشند. شکل ۳۱ ساختار داخلی عامل های لایه های میانی را نشان می دهد.



شکل ۳۱- ساختار داخلی عامل های لایه میانی

در لایه آخر، عامل ها نقش فرماندهی دارند که براساس اطلاعات دریافتی از عامل های لایه های پایین، پس از انجام ارزیابی موقعیت و تهدید در سطح وسیع تر، تصمیم گیری نهایی را انجام داده و دستورات مناسب صادر

می کنند. ورودی این لایه، اطلاعات سایر عامل هاست و عامل های این لایه علاوه بر ارائه اطلاعات به سایر عامل ها، باید به سوالات کاربران از طریق یک رابط کاربر که دارای یک موتور پرس و جو می باشد، پاسخ دهند. شکل ۳۲ ساختار داخلی عامل های لایه های آخر را نشان می دهد.



شکل ۳۲- ساختار داخلی عامل های لایه آخر

واژه‌نامه

تکمیل نشده

داده آمیزی	Data fusion

Engineering Faculty

Department of Computer Engineering

Web Technology Lab

Semantic fusion model based on JDL

By:

Havva Alizadeh Noughabi

Supervisor:

Dr. Mohsen Kahani

Dissertation submitted in pursuance of the degree of Master of computer software

January 2012

Abstract:

Data fusion becomes significant in a variety of situations both commercial and scientific and can be modeled in different methods. The most widely used fusion model is JDL model, but it suffers from semantics and syntax issues. In order to achieve semantic interoperability in a heterogeneous information system, the meaning of the information that is interchanged has to be understood across the systems. Semantic conflicts occur whenever two contexts do not use the same interpretation of the information. Using semantic technologies for the extraction of implicit knowledge is a new approach to overcome this problem. In this paper a semantic fusion model is proposed based on JDL which can resolve semantic conflicts between heterogeneous data sources.

Keywords: semantic web, JDL model, data fusion, RDF