دانشگاه صنعتی شریف
دانشکده مهندسی کامپیوتر

پایان‌نامه کارشناسی ارشد
رایانش امن

تولید رفتار جعلی بر اساس هستی‌شناسی برای حفظ حریم خصوصی در خانه هوشمند

نگارش

بهزاد دارا

استاد راهنما

دکتر مرتضی امینی

آذر ۱۴۰۲


بسم الله الرحمن الرحیم

به نام خدا

دانشگاه صنعتی شریف

دانشکده مهندسی کامپیوتر

پایان‌نامه‌ی کارشناسی ارشد

این پایان‌نامه به عنوان تحقق بخشی از شرایط دریافت درجه‌ی کارشناسی ارشد است.

عنوان: تولید رفتار جعلی بر اساس هستی‌شناسی برای حفظ حریم خصوصی در خانه هوشمند

نگارش: بهزاد دارا

کمیته‌ی ممتحنین

استاد راهنما: دکتر مرتضی امینی امضاء:

استاد داور داخلی: دکتر ... امضاء:

استاد داور مدعو: دکتر ... امضاء:

تاریخ:

اظهارنامه

(اصالت متن و محتوای پایان‌نامه کارشناسی ارشد)

عنوان پایان‌نامه: تولید رفتار جعلی بر اساس هستی‌شناسی برای حفظ حریم خصوصی در خانه هوشمند

استاد راهنما: دکتر مرتضی امینی استاد مشاور: –

این‌جانب بهزاد دارا اظهار می‌دارم:

۱. متن و نتایج علمی ارائه شده در این پایان‌نامه اصیل بوده و زیر نظر استادان نام‌برده شده در بالا تهیه شده است.

۲. متن پایان‌نامه به این صورت در هیچ جای دیگری منتشر نشده است.

۳. متن و نتایج مندرج در این پایان‌نامه، حاصل تحقیقات این‌جانب به‌عنوان دانشجوی کارشناسی ارشد دانشگاه صنعتی شریف است.

۴. کلیه مطالبی که از منابع دیگر در این پایان‌نامه مورداستفاده قرار گرفته، با ذکر مرجع مشخص شده است.

نگارنده: بهزاد دارا

تاریخ:

امضا:

نتایج تحقیقات مندرج در این پایان‌نامه و دستاوردهای مادی و معنوی ناشی از آن (شامل فرمول‌ها، توابع کتابخانه‌ای، نرم‌افزارها، سخت‌افزارها و مواردی که قابلیت ثبت اختراع دارد) متعلق به دانشگاه صنعتی شریف است. هیچ شخصیت حقیقی یا حقوقی بدون کسب اجازه از دانشگاه صنعتی شریف حق فروش و ادعای مالکیت مادی یا معنوی بر آن یا ثبت اختراع از آن را ندارد. همچنین، کلیه حقوق مربوط به چاپ، تکثیر، نسخه‌برداری، ترجمه، اقتباس و نظایر آن در محیط‌های مختلف اعم از الکترونیکی، مجازی یا فیزیکی برای دانشگاه صنعتی شریف محفوظ است. نقل مطلب با ذکر مأخذ بلامانع است.

استاد راهنما: دکتر مرتضی امینی نگارنده: بهزاد دارا

تاریخ: تاریخ:

امضا: امضا:

سپاس بر

پروردگار که در تمامی لحظات زندگی، حضور نعمت‌های بی‌کران او را دیده‌ام و ستایش به درگاه او که مرا در انجام این پژوهش یاری کرد تا ذره‌ای از آنچه در مکتب اساتید آموخته‌ام به عنوان ره‌آوردی مختصر ارائه نمایم.

تقدیم به

بهترین تکیه‌گاه دنیا، پدرم

بهترین حامی دنیا، مادرم

بهترین همراه دنیا، همسرم

تقدیر و تشکر

در اینجا وظیفه خود می‌دانم از تمامی افرادی که به طریقی مرا در انجام این پایان‌نامه یاری نموده‌اند، تشکر کنم. به خصوص استاد بزرگوارم، جناب آقای دکتر مرتضی امینی که همواره با گشاده‌رویی خویش پذیرای اینجانب بودند و بادقت و حوصله اینجانب را در تهیه و تدوین این پایان‌نامه یاری نموده‌اند و از هیچ‌گونه تلاش و کوششی دریغ ننموده‌اند.

چکیده

...

کلیدواژه‌ها: اینترنت اشیاء، حریم خصوصی، خانه‌ی هوشمند، هستی‌شناسی، سکوهای اینترنت اشیاء

فهرست مطالب

فهرست جداول

فهرست تصاویر

فصل ۱

مقدمه

خانه‌های هوشمند[۱] نمونه‌ای از کاربردهای مبتنی بر تکنولوژی‌های نوین، برای کمک به زندگی مستقل جمعیت مسن رو به رشد در جهان و همچنین بالاتر بردن کیفیت زندگی انسانها از بعد راحتی و آسایش و امنیت[۲] هستند. با اینکه از تعریف اولیه این مفهوم بیش از ۲۰ سال می‌گذرد اما با بالاتر رفتن سرعت هوشمند شدن حسگرها[۳] و کوچکتر شدن اندازه آن‌ها و همچنین ارزانتر شدن هزینه‌های استفاده از آن‌ها در خانه‌های هوشمند، تحقیقات و پیشرفت‌های این حوزه در حال سرعت گرفتن است. خانه‌های هوشمند در کنار کاربردهای دیگر اینترنت اشیاء[۴] مانند کشاورزی هوشمند، سیستم‌های مبتنی بر سلامت هوشمند و امثالهم، باعث افزایش تعداد حسگرها و عملگرهای اینترنت اشیاء به کارگرفته شده در جهان شده‌اند. طبق برآورد صاحب نظران این حوزه تا سال ۲۰۲۵ تعداد دستگاه‌های اینترنت اشیاء در حال استفاده در جهان، به بیش از ۵۷ میلیارد خواهد رسید [۱].

عمده تحقیقات و پیشرفت‌های صورت گرفته در حوزه خانه‌های هوشمند، جدا از بهبودها و پیشرفت‌های صورت گرفته در حوزه حسگرها، در زیرحوزه‌های سلامت انسان‌ها مانند پایش اطلاعات حیاتی مرتبط با بیماران، پایش رفتار افراد مسن، خودکارسازی[۵] رفتارها، کنترل انرژی‌های مصرفی در خانه و دسترسی به سرویس‌های از راه دور صورت گرفته است که همه آن‌ها متکی بر شناسایی و کلاس‌بندی رفتار کاربران است.

با پیشرفت‌های هرچه بیشتر در این حوزه، به تدریج مشکلات بیشتری نیز مربوط به چگونگی حفظ حریم خصوصی کاربر شناسایی و معرفی می‌گردد. برای مثال با بیشتر شدن استفاده از تجهیزات هوشمند

[۱] Smart homes
[۲] Safety
[۳] Sensors
[۴] Internet of things
[۵] Automation

بی‌سیم در خانه‌های هوشمند، آسیب‌پذیری‌های ذاتی این تجهیزات و حملات مرتبط با آن‌ها از جمله حمله کانال جانبی[۶] که مربوط به شنود ترافیک ارسالی و استنتاج اطلاعات حساس به صورت غیر مستقیم از ترافیک ارسالی است، حریم خصوصی کاربر را بیشتر در معرض خطر قرار داده است [۲].

از سوی دیگر، تجهیزات اینترنت اشیاء موجود در بازار که قابل استفاده در خانه‌های هوشمند هستند همگی ساخت یک تولید کننده خاص نیستند و لذا با مشکل عدم امکان تعامل با یکدیگر روبرو هستند. این مشکل، کاربران را متمایل به استفاده از سکوهای اینترنت اشیاء می‌نماید چرا که این سکوهای اینترنت اشیاء[۷]، کاربران را قادر می‌سازند تا با اتصال دستگاه‌ها و سرویس‌های برخط گوناگون به یکدیگر، قواعد خودکارسازی دلخواه خود را اعمال کنند و از سرویس‌های متنوع ارائه شده توسط این سکوها بهره‌مند شوند برای مثال یکی از سرویس‌های مورد استقبال کاربران در این حوزه، شناسایی رفتار فعلی کاربر و ارائه پاسخ دقیق به کاربر در مقابل رفتار مشاهده شده است.

از آنجا که سکوهای اینترنت اشیاء هیچ قابلیتی برای کنترل نشت داده‌های حسگرها، در اختیار کاربران قرار نمی‌دهند، لذا حریم خصوصی کاربر را با خطر مواجه می‌نمایند. هنگامی که از امکان نقض حریم خصوصی کاربر با دسترسی غیر مجاز به داده‌های رفتاری کاربر حاصل از حسگرهای خانه‌های هوشمند صحبت می‌کنیم در واقع به این موضوع توجه داریم که تجهیزات یک خانه هوشمند همچون حسگرها و عملگرها، طیف وسیعی از داده‌های رفتاری ساکنان خانه هوشمند را به طور منظم جمع‌آوری می‌نمایند. به عبارت دیگر تجهیزات هوشمند امروزی مانند گوشی‌های تلفن همراه، ساعت‌های هوشمند، تجهیزات پوشیدنی[۸] هوشمند و بسیاری از تجهیزات الکترونیکی مدرن، قابلیت تولید داده دارند. چون این داده‌ها در قالب‌های خام و اولیه خود شامل اطلاعات حساسی درباره ساکنان خانه هوشمند هستند و همچنین با توجه به اینکه در زمانی زندگی می‌کنیم که جرایم سایبری هر روز گسترده‌تر، ویرانگرتر و پیچیده‌تر می‌شود، لذا جمع‌آوری داده‌ها بدون توجه کافی به نوع و مفهوم داده‌های ارسالی از دیدگاه مهاجمین، تبعات حتمی نقض حریم خصوصی کاربر و استفاده غیر مجاز از این داده‌ها را به دنبال خواهد داشت. به همین دلیل است که طبق مطالعات صورت گرفته اخیر، حفظ حریم خصوصی کاربران یکی از موانع بسیار اساسی در توجه و سازگار شدن عموم افراد به استفاده از تکنولوژی‌های خانه‌های هوشمند است [۳].

با توجه به مواردی که ذکر شد مشخص است که تحلیل قابل اعتماد داده‌های ارسالی حسگرها و عملگرها و به طور کلی رفتار کاربر در یک خانه هوشمند و کسب اطمینان از محافظت از این داده‌ها در مقابل دسترسی مهاجمینی که اقدام به شنود ترافیک ارسالی می‌نمایند و یا عدم ارسال داده‌های محرمانه کاربر به سکوهای اینترنت اشیاء، چالش بزرگی پیش روی ارائه کنندگان راهکارهای امنیتی در این حوزه

[۶] Side channel attack
[۷] Internet of things platforms
[۸] Bearable device

است.

در حوزه حفظ حریم خصوصی کاربر در برابر سکوهای نامعتمد اینترنت اشیاء، این سوال مطرح است که چگونه می‌توان داده‌های حسگرها را به سکوهای اینترنت اشیاء ارسال کرد و از سرویس‌های متنوع این سکوها بهره‌مند شد بدون این که به حریم خصوصی کاربر خدشه‌ای وارد شود و فعالیت‌های حساس و رفتار کاربر از دید سکو قابل شناسایی نباشد. راهکارهای ارائه شده برای پاسخ به این سوال می‌بایست توازنی در پاسخ به هر دو مسئله داشته باشند و مصالحه‌ای بین حفظ حریم خصوصی کاربر و دریافت سرویس‌های مد نظر کاربر در خانه هوشمند ایجاد نمایند. این راهکارها می‌بایست برای شناسایی رفتار کاربر در خانه هوشمند یک مدل رفتاری مناسب ایجاد نمایند و سپس قادر باشند تا با پنهان‌سازی، رفتارهای حساس کاربر را از دیدگاه سکوهای اینترنت اشیاء، مخفی نمایند.

در سال‌های اخیر، راهکارهایی در جهت حفظ حریم خصوصی کاربر در خانه‌های هوشمند در برابر سکوهای نامعتمد ارائه شده است. این راهکارها بر اساس روش، به راهکارهای مبتنی رمزنگاری، کمینه‌سازی[۹]، آشفته‌سازی[۱۰] و تولید رویداد جعلی تقسیم می‌شوند. راهکارهای مبتنی بر رمزنگاری، با استفاده از تکنیک‌های رمزنگاری، محاسبات چندجانبه امن و محیط اجرای امن داده‌های کاربر را از دید سکو پنهان می‌کنند. راهکارهای مبتنی بر کمینه‌سازی، از روش حذف داده‌هایی که در اجرای فواعد رهانا[۱۱]ـ‌کنش[۱۲] اثرگذار نیستند، اقدام به کاهش اطلاعات ارسالی به سکو و ناقص کردن دانش آن می‌کنند. راهکارهای مبتنی بر آشفته‌سازی، داده‌های کاربر را قبل از ارسال به سکو به شکل‌های مختلف تغییر می‌دهند. راهکارهای مبتنی بر رویداد جعلی، برای حفظ حریم خصوصی کاربر و اطلاعات حساس آن، از ارسال رویدادهای جعلی به سکو استفاده می‌کند؛ به نحوی که از دید سکو رویدادهای جعلی و واقعی قابل تمایز نباشند.

این پژوهش با هدف افزایش امنیت در خانه‌های هوشمند انجام شده و از هستی‌شناسی[۱۳] خانه‌های هوشمند بهره برده است. در این راستا، برای محافظت از امنیت خانه‌های هوشمند در برابر حملات مخرب و جلوگیری از نفوذ مهاجمان، اقدام به تولید سلسله رویداد جعلی شده است. این سلسله رویدادها با دقت و اصول هستی‌شناسی خانه طراحی شده‌اند به نحوی که مهاجمان قادر به تشخیص دقیق داده‌های واقعی از داده‌های جعلی نباشند و به تبع آن، نتوانند اطلاعات حساس مربوط به زندگی افراد در خانه‌های هوشمند را به دست آورند.

یکی از جوانب مهم در طراحی این راه حل این است که تنوع و تصادف در تولید سلسله رفتارها حفظ

---

[۹] Filtering
[۱۰] Randomization
[۱۱] Trigger
[۱۲] Action
[۱۳] Ontology

شده و از الگوهای قابل پیش‌بینی پرهیز گردد. برای این منظور، از عوامل تصادفی‌ساز[۱۴] بهره گرفته شده تا مهاجمین نتوانند با تحلیل تکراری بودن رفتارها به اهداف خود دست یابند. اقدام دیگری که علاوه بر تولید متنوع سلسله رفتارها برای گمراه‌سازی مهاجم انجام می‌شود، زمان انجام هر رفتار پس از رفتار دیگر است که با استفاده از عوامل تصادفی ساز، زمان انجام هر رفتار در بازه‌ای مشخص متغیر است. این پژوهش امیدوار است که با اجرای این برنامه، امنیت خانه‌های هوشمند تقویت شده و از حملات ناخواسته جلوگیری شود.

این پایان‌نامه در شش فصل ابعاد مختلف مساله را بررسی کرده و ارائه‌ی راه‌حل و ارزیابی آن را انجام می‌دهد. در فصل دوم تعاریف مفاهیم پایه‌ی مورد نیاز برای درک کامل مساله ارائه می‌شود، در فصل سوم پژوهش‌های پیشین مرتبط با این پژوهش را بررسی کرده که هر یک به بررسی یک یا چند بخش مرتبط با این پژوهش را انجام داده‌اند. در فصل چهارم راه‌حل ارائه شده برای حل این مساله را مدل‌سازی کرده و پیاده سازی کامل و جامع آن را ارائه می‌کنیم. در فصل پنجم به ارزیابی روش پیشنهادی و ارائه نتایج حاصل از ارزیابی می‌پردازیم و در فصل آخر نتیجه‌گیری این پژوهش ارائه خواهد شد.

[۱۴] Randomizing factors

فصل ۲

تعاریف و مفاهیم اولیه

پیش از مرور کارهای انجام شده در زمینه‌ی انتشار داده و حفظ حریم خصوصی و همچنین تشخیص فعالیت‌های کاربران در خانه های هوشمند نیاز است تا در ابتدا تعاریف و مفاهیمی پایه‌ای مورد نیاز ارائه گردد.

۲-۱ مفاهیم پایه

در حوزه خانه هوشمند و تشخیص فعالیت کاربران مفاهیم و اصطلاحات زیادی مطرح است که جهت شفاف‌سازی و ایجاد درک مشترک از مطالب ارائه شده، هر یک را به صورت دقیق تعریف می‌کنیم.

- همبستگی[۱]: همبستگی، ارتباط بین دو یا چند موجودیت[۲] را نشان می‌دهد که به معنی تاثیرگذاری آن‌ها روی یکدیگر است [۴].
- زمینه[۳]: با توجه به تعریف هونگ[۴] و همکاران [۵] و همچنین تعریف رودریگز[۵] و همکاران [۶]، زمینه به هرگونه اطلاعات که برای توصیف وضعیت یک موجودیت استفاده می‌شود گفته می‌شود. در این پژوهش از تعریف یسایی[۶] و همکاران [۷] استفاده شده است. این تعریف بیان می‌کند که زمینه شرایطی است که سیستم در آن کار می‌کند و آن شرایط بر نتیجه سیستم تاثیر می‌گذارد.

---

۱ Correlation
۲ Entity
۳ Context
۴ Hong
۵ Rodriguez
۶ Yasaei

- معناشناسی[۷]: شاخه‌ای از زبان‌شناسی و منطق است که تحلیل معنا و روابط بین کلمات را در خود دارد. زمانی که در سیستمی به اطلاعات معنا داده می‌شود به طوری که برای کاربران و رایانه‌ها قابل فهم و تعامل باشد، مجهز به ابزار معناشناسی است [۶].

- رخداد[۸]: داده دریافتی از حسگرها است که بیانگر حالت حسگر یا نقدار اندازه‌گیری شده توسط حسگر در لحظه‌ای از زمان است [۸].

- فعالیت[۹]: مجموعه‌ای از رخدادها که نمایانگار تاثیرات یک فعالیت انسانی، مانند ظرف شستن یا مسواک زدن، بر روی حسگرهای نصب شده در محیط باشد [۹].

- رفتار[۱۰]: در حالی که دسته‌ای از پژوهش‌ها [۱۰، ۱۱، ۱۲] دو واژه فعالیت و رفتار را هم‌معنی دانسته‌اند؛ در این پژوهش از تعریف دسته دیگر [۱۳، ۱۴] استفاده شده است که رفتار را یک سطح بالاتر و به عنوان مجموعه‌ای از فعالیت‌ها می‌دانند.

۲-۲ روش‌های داده‌محور، دانش‌محور و ترکیبی

روش‌ها در هوش مصنوعی به سه دسته‌ی داده‌محور[۱۱]، دانش‌محور[۱۲] و ترکیبی تقسیم می‌شوند [۱۵، ۱۶]:

- روش‌های داده‌محور: این روش‌ها به صورت خودکار و با استفاده از تکنیک‌های یادگیری ماشین[۱۳]، داده‌های جمع شده تا لحظه کنونی را تبدیل به مدل می‌کنند. روش‌های داده محور در محیط‌های پویا کاربردی بوده و دقت بالایی دارند اما در صورتی که نیاز به دانش با در نظر گرفتن زمینه باشد دچار مشکل می‌شود. با توجه به این مشکل امکان استفاده مجدد یک موجودیت برای موجودیت دیگر وجود ندارد و برای هر موجودیت مدلی جدا برای آموزش نیاز است. همچنین داده زیادی برای آموزش مورد نیاز است که زمانی طول می‌کشد تا به بهره‌وری برسد که اصطلاحا شروع سرد[۱۴] نام دارد. توجه شود که بعضی فعالیت‌ها به ندرت انجام شده و این فعالیت‌های مشاهده نشده نقطه ضعف این روش هستند چرا که تا زمان عدم مشاهده‌ی این فعالیت‌ها، مدل‌سازی ناقص بوده و حتی با گذشت زمان زیادی از یادگیری مدل، نمی‌توان اطمینان از کامل بودن آن داشت.

---

[۷] Semantics
[۸] Event
[۹] Activity
[۱۰] Behaviour
[۱۱] Data-driven
[۱۲] Knowledge-driven
[۱۳] Machine learning
[۱۴] Cold-start

- روش‌های دانش‌محور: در این دسته از روش‌ها فرد خبره[۱۵] با دانش پیشین از حوزه، مدل را به صورت دستی ایجاد می‌کند. این روش زمینه را در نظر می‌گیرد و قابلیت استفاده مجدد دارد. همچنین این روش مشکل شروع سرد را ندارد زیرا نیاز به داده اولیه برای آموزش ندارد و با توجه به دانش فرد خبره به خودی خود کامل است اما نیاز است تا فرد خبره دانش کامل و عمیقی داشته باشد. عیب دیگر این روش ایستا بودن آن است که تغییرات فعالیت کاربران لحاظ نمی‌شود و نیاز است به صورت دستی به‌روز شوند.

- روش‌های ترکیبی: این روش‌ها از ترکیب روش‌های داده‌محور و دانش‌محور استفاده می‌کنند تا محدودیت و نقاط ضعف این روش‌ها را برطرف نمایند و از نقاط قوت آن‌ها بهره ببرند.

۲-۳ هستی‌شناسی

هستی‌شناسی نمایش صوری[۱۶] دانش توسط مجموعه‌ای از مفاهیم[۱۷]، خصوصیات و محدودیتشان و همچنین روابط بین این مفاهیم است [۱۷]. هستی‌شناسی (تی‌باکس[۱۸]) به همراه مجموعه‌ای از نمونه‌ها[۱۹] (ای‌باکس[۲۰]) پایگاه دانش را تشکیل می‌دهند. ای باکس شامل نمونه‌هایی از عناصر تعریف شده در تی‌باکس است (به همراه روابط[۲۱]). هستی‌شناسی در حوزه‌های مختلف از جمله وب معنایی[۲۲]، موتورهای جستجو[۲۳]، تجارت الکترونیکی[۲۴]، پردازش زبان‌های طبیعی[۲۵]، مهندسی دانش[۲۶]، بازیابی اطلاعات[۲۷] و اینترنت اشیاء کاربرد دارد. از مزایای استفاده از هستی‌شناسی می‌توان به موارد زیر اشاره نمود:

- ایجاد یک فهم مشترک از ساختار اطلاعات
- امکان استفاده مجدد
- امکان تحلیل روی دانش

---

[۱۵] Expert
[۱۶] Formal
[۱۷] Concepts
[۱۸] Terminology box (TBox)
[۱۹] Instances
[۲۰] Assertion box (ABox)
[۲۱] Relations
[۲۲] Semantic web
[۲۳] Search engines
[۲۴] Electronic commerce
[۲۵] Natural Language Processing
[۲۶] Knowledge engineering
[۲۷] Data recovery

به طور کلی هستی‌شناسی شامل اجزای اصلی زیر است:

- مفاهیم: مجموعه یا کلاسی از موجودیت‌ها یا چیزهایی که درون یک حوزه وجود دارد.
- روابط: روابط یا ارتباطات برای بیان تعاملات بین مفاهیم و یا معین کردن ویژگی‌های یک مفهوم به کار می‌رود و در هستی‌شناسی دو نوع رابطه بین موجودیت‌ها وجود دارد. ارتباط رده‌بندی که سازماندهی مفاهیم در یک ساختار سلسله مراتبی را نشان می‌دهد مانند ارث‌بری کلاس‌ها در شیءگرایی و ارتباطات پیوندی که ارتباط مفاهیمی را با یکدیگر به نمایش می‌گذارد که در یک ساختار سلسله مراتبی به هم مرتبط نمی‌باشند.
- نمونه‌ها: اعضا یا نمونه‌ها همان چیزهایی هستند که توسط یک مفهوم معرفی می‌شوند مثلاً در حوزه مدارس، مدرسه‌ای با نام «مدرسه الف» عضوی از مفهوم مدرسه است. توجه باید کرد که یک هستی‌شناسی به خودی خود نمونه‌ای ندارد و صرفاً عبارت است از طراحی ساختاری از مفاهیم یک حوزه که ترکیب آن با اعضاء و نمونه‌ها، پایگاه دانش آن حوزه را ایجاد می‌نماید.
- قواعد[۲۸]: قاعده‌ها برای مقید کردن مقادیر برای کلاس‌ها یا ویژگی‌ها مورد استفاده قرار می‌گیرند. مثلاً می‌توان گفت سن یک انسان باید بیشتر از ۰ و کمتر از ۱۲۰ باشد.

تا کنون زبان‌های هستی‌شناسی زیادی توسعه یافته‌اند. این زبان‌ها عموما بر پایه زبان XML[۲۹] [۱۸] هستند که قابلیت تفسیر و سادگی معناشناسی برای ماشین را دارند. از این زبان‌ها می‌توان به RDF[۳۰] و RDF Schema [۱۹]، OIL[۳۱] + DAML[۳۲] [۲۰]، OWL [۲۱] و OWL2 [۲۲] اشاره کرد. یکی از پرکاربردترین آن‌ها OWL[۳۳] است که روی RDF و DAML + OIL توسعه یافته است و قدرت بیان بالایی دارد. OWL دارای سه زیرزبان OWL-Lite، OWL-Full و OWL-DL است و زبان توصیف قواعد SWRL[۳۴] [۲۳] امکان نوشتن قواعد را به OWL-DL اضافه کرده تا قدرت بیان آن را افزایش دهد.

[۲۸] Rules
[۲۹] eXtensible Markup Language
[۳۰] Resource Description Framework
[۳۱] Ontology Inference Layer
[۳۲] DARPA Agent Markup Language
[۳۳] Ontology Web Language
[۳۴] Semantic Web Rule Language

۲-۴ قوانین انجمنی

قوانین انجمنی[۳۵] در داده‌کاوی و یادگیری ماشین به دنبال کشف ارتباطات و تعاملات بین عناصر در مجموعه داده هستند [۲۴]. این نوع ارتباط‌ها به طور معمول بر روی داده‌های تراکنشی مانند فروش‌های خرده‌فروشی یا خریدهای آنلاین کاربرد دارند. این قوانین دارای دو تعریف اساسی پشتیبانی[۳۶] و اطمینان[۳۷] هستند:

- پشتیبانی: پشتیبانی، فرکانس درست بودن یک قانون در یک مجموعه داده معین را اندازه گیری می‌کند و نشان دهنده نسبت تراکنش هایی است که هم موارد موجود در مقدمه و هم موارد موجود در نتیجه قانون را شامل می شود و کمک می‌کند تا کاربرد یک قانون مشخص شود، و از آن برای کشف قوانین رایج یا مکرر در مجموعه داده استفاده می‌شود.
- اطمینان: اطمینان، احتمال مشروط بودن موارد موجود در نتیجه یک قانون را با توجه به اینکه موارد موجود در مقدمه درست هستند، اندازه گیری می کند. در واقع قابلیت اطمینان یک قانون را نشان می‌دهد و اینکه هر چند وقت یک‌بار حضور عناصر را در نتیجه به درستی پیش‌بینی می‌کند، در حالی که عناصر مقدمه وجود دارند.

قوانین انجمنی برای کشف ارتباطات مفهومی و معنادار بین عناصر در مجموعه داده استفاده می‌شوند و در مواردی مانند تجزیه و تحلیل سبد خرید، سیستم‌های پیشنهادی، و اتخاذ تصمیمات در داده‌کاوی و تحلیل داده مورد استفاده قرار می‌گیرند.

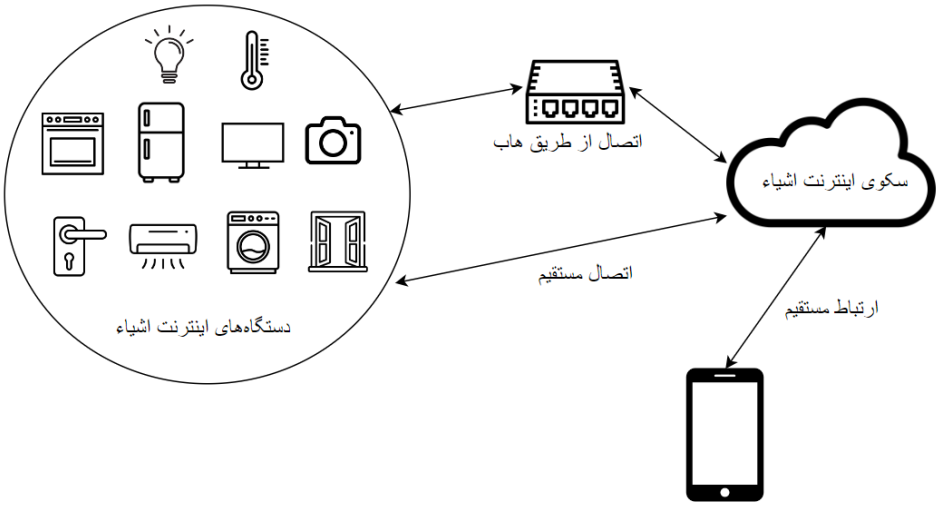
۲-۵ زیست‌بوم خانه‌های هوشمند

با استفاده از امکانات خانه‌های هوشمند کاربران می‌توانند دستگاه‌های اینترنت اشیاء را از راه دور کنترل کنند. ضمن این که کارهای مختلفی نیز می‌تواند به صورت خودکار برای سهولت زندگی انسان در این زیرساخت انجام شود. معماری رایج خانه‌های هوشمند در شکل ۲-۱ نشان داده شده است.

خانه های هوشمند شامل اجزای زیر هستند:

- دستگاه‌های اینترنت اشیاء: دستگاه‌های اینترنت اشیاء شامل حسگر و یا عملگرهایی[۳۸] هستند که حسگرها خصوصیتی را اندازه‌گیری می‌کنند و عملگرها کنشی را انجام می‌دهند. به عنوان مثال حسگر

[۳۵] Associative rules
[۳۶] Support
[۳۷] Confidence
[۳۸] Actuators



شکل ۲-۱: معماری رایج خانه‌های هوشمند [۲۵]

نور، نور محیط را اندازه‌گیری کرده و در صورت بالا بودن شدت نور، عملگر نور محیط را کم می‌کند. عملکرد یک عملگر می‌تواند به صورت خودکار یا به صورت دستی انجام شود. دستگاه‌ها در خانه هوشمند به دو دسته تقسیم می‌شوند [۲۶]:

- دستگاه‌های متصل به ابر[۳۹]: این دستگاه‌ها با تکنولوژی وای‌فای[۴۰] با زیرساخت ابری ارتباط برقرار می‌کنند. استفاده از وای‌فای به دلیل مصرف زیاد انرژی قابل استفاده در تمامی دستگاه‌ها نیست و اکثر دستگاه‌ها از تکنولوژی دیگری بهره می‌برند.
- دستگاه‌های متصل به هاب[۴۱]: این دستگاه‌ها دارای تکنولوژی وای‌فای نبوده و ارتباطشان با زیرساخت ابری از طریق تکنولوژی‌هایی با مصرف انرژی کم است. از این تکنولوژی‌ها می‌توان به زی‌ویو[۴۲] و زیگ‌بی[۴۳] اشاره کرد. این دستگاه‌ها از طریق هاب با زیرساخت ابری ارتباط برقرار می‌کنند.

دستگاه‌های خانه هوشمند ممکن است از یک یا هر دو نوع ذکر شده پشتیبانی کنند.

- هاب: هاب دستگاهی است که به امواج بی‌سیم با برد کم مانند زی‌ویو، زیگ‌بی و وای‌فای مجهز است. دستگاه‌های اینترنت اشیاء از طریق هاب با یکدیگر و زیرساخت ابری ارتباط برقرار می‌کنند.
- زیرساخت ابری: زیرساخت ابری پردازش‌های مربوطه را روی داده‌های دریافتی انجام می‌دهد و

[۳۹] Cloud
[۴۰] Wi-Fi
[۴۱] Hub
[۴۲] Z-Wave
[۴۳] Zigbee

امکان خودکارسازی امور را با استفاده از نرم‌افزارهای اینترنت اشیاء فراهم میکند.

- نرم‌افزار تلفن همراه: نرم‌افزار مدیریتی دستگاه‌ها و هاب که به کاربر این امکان را می‌دهد که تمامی اجزا را کنترل کند.

۲-۶ سکوهای اینترنت اشیاء

سکوهای اینترنت اشیاء امکان خودکارسازی ارتباطات بین دستگاه‌های اینترنت اشیاء با یکدیگر را فراهم می‌کند. این سکوها عموما از مدل رویداد_کنش استفاده می‌کنند. برنامه‌های اینترنت اشیاء می‌توانند روی این سکوها توسعه یافته و قوانین خودکارسازی خود را پیاده کنند. یعنی زمانی که رویدادی مشخص رخ دهد، سکو دستور مربوط به آن رخداد را ارسال می‌کند. امروزه سکوهای اینترنت اشیاء زیادی وجود دارد که می‌توان به IFTTT[۴۴] [۲۷]، اسمارت‌تینگز[۴۵] [۲۸]، اپن‌هاب[۴۶] [۲۹]، زپیر[۴۷] [۳۰]، اپل هوم‌کیت[۴۸] [۳۱] و مایکروسافت پاور اتومیت[۴۹] [۳۲] اشاره کرد. این سکوها از نظر زبان برنامه‌نویسی و معماری تفاوت دارند و به طور کلی همانطور که در بخش ۲-۴ اشاره شد به دو دسته ابرمحور و هاب‌محور تقسیم می‌شوند [۳۳]. در سکوهای ابرمحور مثل اسمارت‌تینگز که رایج‌تر هستند برنامه‌های اینترنت اشیاء روی زیرساخت ابری و در سکوهای هاب‌محور مثل اپل هوم‌کیت برنامه‌ها روی هاب اجرا می‌شوند.

---

[۴۴] If This Then That
[۴۵] SmartThings
[۴۶] OpenHAB
[۴۷] Zapier
[۴۸] Apple home kit
[۴۹] Microsoft power automate

فصل ۳

کارهای پیشین

تا کنون راهکارهای بسیاری برای شناسایی فعالیت کاربر و حفظ حریم خصوصی در خانه‌های هوشمند ارائه شده است که هر یک با فرضیات و دیدگاه متفاوتی اقدام به حل مساله کرده‌اند. در این فصل هستی‌شناسی‌های مختلف، روش‌های شناسایی رفتار کاربر و راهکارهای حفظ حریم خصوصی در خانه هوشمند را بررسی و دسته‌بندی کرده و به مقایسه راهکارهای ارائه شده خواهیم پرداخت.

۳-۱ هستی‌شناسی‌های حوزه اینترنت اشیاء

در حوزه اینترنت اشیاء هستی‌شناسی‌های متعددی تا کنون تعریف شده است که آن‌ها را از جهات مختلف می‌توان دسته‌بندی نمود. در پژوهشی که توسط باجاج[۱] و همکاران [۳۴] صورت گرفته است هستی‌شناسی‌های حوزه اینترنت اشیاء به چهار دسته زیر تقسیم شده است و در هر دسته نیز هستی‌شناسی‌ها بر اساس عمومی[۲] بودن و خاص‌دامنه[۳] بودن (مانند دامنه ساختمان‌های هوشمند) تفکیک شده‌اند. در این بخش به بررسی هر یک از این دسته‌ها می‌پردازیم.

---

[۱] Bajaj
[۲] Generic
[۳] Domain specific

۳-۱-۱ هستی‌شناسی مبتنی بر حسگر

هستی‌شناسی‌هایی که در این دسته قرار می‌گیرند مفاهیمی را در رابطه با حسگرها مانند داده‌های نمایش داده شده[۴] توسط آن‌ها، قابلیت‌های حسگرها[۵] (مانند میزان دقت و گستره پوشش آن‌ها)، توسعه‌پذیری حسگرها[۶]، نحوه به اشتراک‌گذاری داده‌ها[۷] و اکتشاف حسگرها[۸] را دربر می‌گیرند. هر یک از هستی‌شناسی‌های این دسته تنها بخشی از نیازهای موجود را پوشش داده‌اند.

در این دسته از هستی‌شناسی‌ها می‌توان به SSN[۹] اشاره نمود که یک هستی‌شناسی مبتنی بر حسگر در زیردسته کاربردهای عمومی است و توسط W3C[۱۰] پیشنهاد شده است [۳۵]. هدف هستی‌شناسی SSN حل مشکل ناهمگونی[۱۱] در داده‌های نمایشی و اکتشاف حسگرهاست اما مفاهیمی که پشتیبانی می‌کند محدود است. ژو[۱۲] و همکاران [۳۶] یک هستی‌شناسی با مفهوم نوع حسگر (عادی یا پیشرفته) و قابلیت حسگر (ایستا یا پویا) معرفی کرده‌اند که برای تعداد محدودی حسگر، توصیف معنایی ارائه می‌کند. جرارد[۱۳] و همکاران [۳۷] با معرفی هستی‌شناسی به نام M3 مشکل محدودیت تعداد حسگرها را برطرف کرده‌اند و با توسعه‌ی هستی‌شناسی SSN، دامنه و مشاهدات حسگرها را پشتیبانی کرده و از آن‌ها برای استنتاج روی قواعد زمینه‌ای استفاده می‌کنند.

از آنجایی که ابزار ارتباط با حسگرها ممکن است تلفن همراه باشد که به صورت پویا جابجا می‌شود، اکتشاف حسگرها چالشی مهم است که روسومانو[۱۴] و همکاران [۳۸] در پژوهشی یک هستی‌شناسی معرفی کرده‌اند که برای شناسایی رفتار، ارتباط، عملکرد و ابرداده‌ی حسگرها استفاده می‌شود. محدودیت این راهکار، پیچیدگی زیاد و ناتوانی در توصیف مشاهدات حسگرهاست که نیلز[۱۵] و همکاران [۳۹] این مشکل را با طرح یک هستی‌شناسی مرتبط با مفاهیم SSN حل کردند.

هیرمر[۱۶] و همکاران [۴۰] هستی‌شناسی برای ثبت حسگرهای جدید به صورت پویا معرفی کرده‌اند. در این پژوهش خصیصه‌هایی مانند نوع داده‌های حسگر مشاهده شده و با نوع داده‌های حسگر که توسط تولیدکننده‌ی حسگر اعلام شده مقایسه می‌گردد تا نوع حسگر تشخیص داده شود. در پژوهش شی[۱۷] و

[۴] Sensor data description
[۵] Sensor capabilities
[۶] Sensor extensibility
[۷] Data access & sharing
[۸] Sensor discovery
[۹] Semantic Sensor Network
[۱۰] World Wide Web Consortium
[۱۱] Heterogeneity
[۱۲] Xue
[۱۳] Gyrard
[۱۴] Russomanno
[۱۵] Niles
[۱۶] Hirmer
[۱۷] Shi

همکاران [۴۱] این امر خودکار شده و با توجه به محیط حسگر، زمان ارسال داده و موقعیت آن داده‌های مشاهده شده توسط حسگر دریافت شده و ماهیت حسگر تشخیص داده می‌شود.

در زیردسته کاربردهای خاص‌دامنه می‌توان به پژوهش دنیل[۱۸] و همکاران [۴۲] اشاره نمود که برای استفاده و مدیریت لوازم خانگی هوشمند استفاده می‌شود. در پژوهش دیگری از دنیل و همکاران [۴۳] توسعه‌ای روی پژوهش قبلی انجام شد که با استانداردهای مصرف انرژی خود را تطبیق داد. پژوهشی خاص‌دامنه دیگر، پژوهش دی[۱۹] و همکاران [۴۴] است که هستی‌شناسی روسومانو و همکاران [۳۸] را برای دامنه انرژی توسعه داده است. پژوهش دیگری برای دامنه مدیریت ساختمان توسط بالاجی[۲۰] و همکاران [۴۵] انجام شده که برای تشخیص حسگرها از برچسب استفاده می‌کنند.

هستی‌شناسی دیگری توسط هاچم[۲۱] و همکاران [۴۶] معرفی شده است که در حوزه خانه‌های هوشمند برای مقابله با چالش‌هایی مانند تفاوت در وضوح حسگرها که ممکن است باعث شود که چند سرویس در نهایت فعال‌کننده‌ای را در وضعیت‌هایی متضاد با یکدیگر فعال نمایند، کاربرد دارد. این هستی‌شناسی به مفاهیمی مانند حسگرها و ویژگی‌های آن‌ها، وضوح اندازه‌گیری و خطاهای مرتبط با حسگرها و جایگاه حسگرها در خانه هوشمند به همراه همه واحدهای قابل اندازه‌گیری آن‌ها می‌پردازد.

۳-۱-۲ هستی‌شناسی مبتنی بر زمینه

هستی‌شناسی‌های این دسته با توصیف زمینه و دسته‌بندی داخلی یا خارجی تعریف می‌شوند [۴۷] و عمدتا از نوع خاص‌دامنه هستند.

در زیردسته کاربردهای عمومی، هستی‌شناسی‌های مبتنی بر زمینه برای توصیف داده‌های حسگرها به کار می‌روند [۴۸]. بالدوف[۲۲] و همکاران [۴۷] هستی‌شناسی با دسته‌بندی خارجی یا داخلی معرفی کرده‌اند که زمینه‌های خارجی، با حسگر فیزیکی و زمینه‌های داخلی، با تعاملات کاربران اندازه‌گیری می‌شوند. چن[۲۳] و همکاران [۴۹] یک هستی‌شناسی برای محیط هوشمند ارائه کرده‌اند که هر موجودیت را با استفاده از موقعیت جغرافیایی و توضیحات آن توصیف می‌کند. هابز[۲۴] و همکاران [۵۰] به کمک حسگرهای فیزیکی و مجازی تلفن همراه هستی‌شناسی مبتنی بر زمینه‌ای معرفی کرده‌اند که نتایج استنتاج در این پژوهش، از اطلاعات جی‌پی‌اس[۲۵] بسیار دقیق‌تر می‌باشد.

---

[۱۸] Daniele
[۱۹] Dey
[۲۰] Balaji
[۲۱] Hachem
[۲۲] Baldauf
[۲۳] Chen
[۲۴] Hobbs
[۲۵] GPS

در زیردسته کاربردهای خاص‌دامنه می‌توان به پژوهش اوکیو[۲۶] و همکاران [۵۱] اشاره کرد که برای توصیف معنایی فعالیت روزانه کاربر[۲۷] به کار می‌رود و از این هستی‌شناسی برای استنتاج فعالیت‌های پیچیده استفاده می‌شود. عیب این پژوهش در نظر نگرفتن فعالیت‌های گروهی مانند جلسات و مهمانی‌هاست که باعی[۲۸] و همکاران [۵۲] این مشکل را حل کرده‌اند. این پژوهش با شناسایی حسگرها و و موقعیت آن‌ها توانایی استنتاج روی فعالیت‌ها و تفکیک فعالیت‌های انفرادی و گروهی را دارد.

در پژوهشی دیگر، لی[۲۹] و همکاران [۵۳] هستی‌شناسی فعالیت دانشگاه را معرفی کردند که فعالیت افراد داخل دانشگاه را مورد بررسی قرار می‌دهد. در این پژوهش از مدل‌سازی درختی مفاهیم استفاده شده و برای هر بخش از یک زیرهستی‌شناسی[۳۰] برای تمایز با سایر بخش‌ها استفاده شده است و هر مفهوم جدید که وارد شود، موقعیت پایین‌تری در درخت مفاهیم خواهد داشت.

پژوهش دیگری که توسط چن و همکارانش [۵۴] برای شناسایی رفتار کاربر در خانه هوشمند ایجاد شده است، در زیر دسته کاربرد خاص قرار می‌گیرد. در این هستی‌شناسی بر ایجاد پروفایل کاربر ناشی از انجام فعالیت تاکید شده است و در آن پروفایل کاربر دارای دو بخش اطلاعات ایستا (مانند سن، نام و نقش کاربر) و اطلاعات پویا (مانند ترجیحات کاربر در انجام فعالیت مانند طول زمان انجام فعالیت، مکان انجام فعالیت، طریقه خاص انجام فعالیت) است [۵۵].

۳-۱-۳ هستی‌شناسی مبتنی بر مکان

هستی‌شناسی مبتنی بر مکان برای توصیف زمینه‌ی فیزیکی کاربران و دستگاه‌ها استفاده می‌شود. با اینکه مکان خود نوعی زمینه می‌باشد اما می‌توان هستی‌شناسی‌هایی که صرفاً به این مفهوم پرداخته‌اند را در دسته‌بندی جداگانه‌ای قرار داد چون بسیاری از آن‌ها را در حوزه‌هایی فراتر از اینترنت اشیاء می‌توان استفاده نمود.

در این دسته از هستی‌شناسی‌ها می‌توان به هستی‌شناسی WGS84[۳۱] که توسط برکلی[۳۲] [۵۶] ارائه شده است، اشاره نمود که در زیردسته کاربردهای عمومی قرار دارد. این هستی‌شناسی با استفاده از طول[۳۳] و عرض[۳۴] جغرافیایی، موقعیت موجودیت‌ها را توصیف کرده و مفهوم انتزاعی برای موجودیت‌های فضایی[۳۵]

[۲۶] Okeyo
[۲۷] Activity of Daily Living (ADL)
[۲۸] Bae
[۲۹] Lee
[۳۰] sub-Ontology
[۳۱] World Geodetic System version 84
[۳۲] Brickley
[۳۳] Longitude
[۳۴] Latitude
[۳۵] SpatialThings

مانند ساختمان و موجودیت‌های موجودیت‌های زمانی[۳۶] مانند مدت زمان ارائه می‌کند. هستی‌شناسی با توصیف بهتر در پژوهش فلوری[۳۷] و همکاران [۵۷] معرفی شده است که با مدل ریاضی، توصیفات مختلف مکانی دسته‌بندی می‌شود. در پژوهش دیگری، کیم[۳۸] و همکاران [۵۸] با استفاده از داده‌ی حسگرها و استنتاج روی آن‌ها موقعیت کاربران را تخمین می‌زنند.

در زیردسته کاربردهای خاص‌دامنه می‌توان به پژوهش سزاس[۳۹] و همکاران [۵۹] اشاره نمود که هستی‌شناسی مبتنی بر مکان برای دامنه داخل ساختمان و موقعیت‌یابی در آن است و از مفاهیم مختلفی از هستی‌شناسی‌های دیگر بهره می‌برد. این هستی‌شناسی قابل تعمیم برای استفاده در محیط خارج از ساختمان نیز می‌باشد.

۳-۱-۴ هستی‌شناسی مبتنی بر زمان

زمان یک زمینه موقتی است و هستی‌شناسی‌های این دسته برای نمایش این مفهوم موقتی مورد استفاده قرار می‌گیرند.

در زیردسته کاربردهای عمومی از این هستی‌شناسی می‌توان به پژوهش فیکس[۴۰] و همکاران [۶۰] اشاره نمود که بر اساس خصیصه زمان، فاصله موجودیت‌ها را تعیین می‌کند. پر استفاده‌ترین هستی‌شناسی این دسته OWL-Time است که در آن مفاهیمی مانند زمان و تاریخ بر اساس موقعیت جغرافیایی تعریف شده‌اند و توسط هابز و همکاران [۵۰] معرفی شده‌اند.

پوستجوفستکی[۴۱] و همکاران [۶۱] هستی‌شناسی مبتنی بر زمان در زیردسته کاربردهای خاص‌دامنه تعریف کرداند که بر پایه مدت زمان و رویداد است از پردازش زبان طبیعی بهره می‌برد. دی و همکاران [۶۲] این پژوهش را توسعه داده و برای حسگرهای انرژی کاربرد دارد. در پژوهش دیگری ژانگ[۴۲] و همکاران [۶۳] بر اساس فرهنگ و تاریخ، رویدادها را تشخیص داده و از تقویم چینی برای تشخیص زمان‌های مهم و موقتی استفاده کرده است.

---

[۳۶] TemporalThings
[۳۷] Flury
[۳۸] Kim
[۳۹] Szász
[۴۰] Fikes
[۴۱] Pustejovsky
[۴۲] Zhang

۳-۱-۵ جمع‌بندی

در این بخش پژوهش‌های مربوط به انواع هستی‌شناسی در حوزه اینترنت اشیاء را بر اساس دسته‌بندی باجاج و همکاران [۳۴] بررسی کردیم. این دسته‌بندی بر اساس دامنه و کاربرد هر یک از هستی‌شناسی‌های حوزه اینترنت اشیاء ارائه شده است. جمع‌بندی کلی این هستی‌شناسی‌ها در جدول ۳-۱ قابل مشاهده است.

جدول ۳-۱: جمع‌بندی کلی هستی‌شناسی‌های حوزه اینترنت اشیاء

پژوهش	دسته‌بندی	کاربرد	ویژگی‌های راهکار
W3C [۳۵]	مبتنی بر حسگر	عمومی	حل ناهمگونی داده‌های نمایشی
ژو و همکاران [۳۶]	مبتنی بر حسگر	عمومی	توصیف معنایی تعداد محدودی حسگر
جرارد و همکاران [۳۷]	مبتنی بر حسگر	عمومی	توسعه SSN و حل مشکل محدودیت تعداد حسگر
روسومانو و همکاران [۳۸]	مبتنی بر حسگر	عمومی	شناسایی حسگرها در محیط پویا و ناتوان در توصیف مشاهدات حسگرها
نیلز و همکاران [۳۹]	مبتنی بر حسگر	عمومی	حل مشکل ناتوانی در توصیف مشاهدات حسگرها با استفاده از مفاهیم SSN
هیرمر و همکاران [۴۰]	مبتنی بر حسگر	عمومی	شناسایی حسگرها در محیط پویا با استفاده از نوع داده‌ی حسگرها
شی و همکاران [۴۱]	مبتنی بر حسگر	عمومی	خودکارسازی شناسایی حسگرها با استفاده از فواصل زمانی داده‌های ارسالی
دنیل و همکاران [۴۲]	مبتنی بر حسگر	خاص‌دامنه	مدیریت لوازم خانگی هوشمند
دنیل و همکاران [۴۳]	مبتنی بر حسگر	خاص‌دامنه	مدیریت لوازم خانگی هوشمند منطبق با استانداردهای انرژی
دی و همکاران [۴۲]	مبتنی بر حسگر	خاص‌دامنه	توسعه‌ی [۳۸] برای دامنه انرژی
بالاجی و همکاران [۴۳]	مبتنی بر حسگر	خاص‌دامنه	مدیریت ساختمان با برچسب‌گذاری روی حسگرها
هاچم و همکاران [۴۶]	مبتنی بر حسگر	خاص‌دامنه	مدیریت عملکرد حسگرها
بالدوف و همکاران [۴۷]	مبتنی بر زمینه	عمومی	زمینه حسگر فیزیکی و تعاملات کاربران
چن و همکاران [۴۹]	مبتنی بر زمینه	عمومی	استفاده از محیط جغرافیایی موجودیت‌ها
هابز و همکاران [۵۰]	مبتنی بر زمینه	عمومی	استفاده از حسگرهای تلفن همراه و ارائه موفعیت مکانی دقیق‌تر از جی‌پی‌اس
اوکیو و همکاران [۵۱]	مبتنی بر زمینه	خاص‌دامنه	توصیف معنایی فعالیت روزانه کاربر (بدون فعالیت‌های گروهی)
باعی و همکاران [۵۲]	مبتنی بر زمینه	خاص‌دامنه	توصیف فعالیت‌های روزانه کاربر با دسته‌بندی انفرادی و گروهی
لی و همکاران [۵۳]	مبتنی بر زمینه	خاص‌دامنه	هستی‌شناسی فعالیت‌های دانشگاه
چن و همکارانش [۵۴]	مبتنی بر زمینه	خاص‌دامنه	ایجاد پروفایل کاربر در خانه هوشمند با اطلاعات ایستا و پویا
برکلی [۵۶]	مبتنی بر مکان	عمومی	توصیف طول و عرض جغرافیایی برای موقعیت موجودیت‌ها
فلوری و همکاران [۵۷]	مبتنی بر مکان	عمومی	دسته‌بندی توصیفات مکانی با مدل ریاضی
کیم و همکاران [۵۸]	مبتنی بر مکان	عمومی	تخمین موقعیت کاربر با استفاده از داده‌های حسگرها
سزاس و همکاران [۵۹]	مبتنی بر مکان	خاص‌دامنه	موقعیت‌یابی داخل ساختمان
فیکس و همکاران [۶۰]	مبتنی بر زمان	عمومی	تعیین فاصله با استفاده از فواصل زمانی داده‌ها
هابز و همکاران [۵۰]	مبتنی بر زمان	عمومی	تعریف خصیصه‌های زمان با استفاده از موقعیت جغرافیایی
پوستجوفستکی و همکاران [۶۱]	مبتنی بر زمان	خاص‌دامنه	تعریف بر اساس رویداد و مدت زمان با استفاده از پردازش زبان طبیعی
دی و همکاران [۶۲]	مبتنی بر زمان	خاص‌دامنه	توسعه [۶۱] و استفاده در دامنه انرژی
ژانگ و همکاران [۶۳]	مبتنی بر زمان	خاص‌دامنه	استفاده از فرهنگ و تاریخ و تقویم چینی برای توصیف موقت رویدادهای زمانی

۳-۲ تشخیص فعالیت‌های کاربران در خانه های هوشمند

فعالیت کاربر، به هرگونه اقدام، رفتار و یا حرکت از سمت انسان گفته می‌شود که این فعالیت‌ها شامل طیف وسیعی از اقداماتی است که انسان به عنوان بخشی از کارهای روزمره یا وظایف خود انجام می‌دهد. تشخیص فعالیت کاربر[۴۳] اقدامی خودکار است که کار شناسایی و دسته‌بندی فعالیت‌های کاربر را با استفاده از اطلاعات دریافت شده از حسگرها انجام می‌دهد [۶۴]. به طور مثال روشن کردن لامپ خانه یک نمونه فعالیت از جانب کاربر است که حسگر تشخیص نور، افزایش نور را تشخیص می‌دهد و به صورت خودکار می‌توان متوجه روشن شدن لامپ توسط کاربر شد.

راهکارهایی که تا کنون برای مدل‌سازی و تشخیص فعالیت‌های کاربران در خانه‌های هوشمند ارائه شده است به طور کلی در سه دسته داده محور، دانش محور و ترکیبی قرار دارند که مزایا و معایب هر یک در بخش ۲-۲ شرح داده شد. در این بخش پژوهش‌های انجام شده در این زمینه را بررسی می‌کنیم تا راهکارهای مختلف مدل‌سازی فعالیت‌های کاربران را بدانیم و در راهکار پیشنهادی از مدل‌سازی مناسب استفاده کنیم.

۳-۲-۱ راهکارهای داده‌محور

مدت زیادی است که در بسیاری از پژوهش‌های مدل کردن فعالیت‌های کاربران در خانه‌های هوشمند، تکنیک‌های یادگیری ماشین استفاده می‌گردد. این راهکارها عموما از مدل‌های آماری و احتمالاتی مثل دسته‌بند بیز ساده[۴۴] [۶۵]، شبکه‌های بیزین[۴۵] [۱۰، ۶۶]، مدل پنهان مارکوف[۴۶] [۶۷، ۶۸، ۶۹]، خوشه‌بندی[۴۷] سلسله‌مراتبی [۷۰]، فرایندهای تصمیم مارکوف تا حدودی مشاهده‌پذیر[۴۸] [۷۱] و مدل پنهان مارکوف جفت‌شده[۴۹] [۷۲] استفاده می‌کنند.

راهکارهایی مبتنی بر تکنیک‌های دسته‌بندی مثل استفاده از نزدیک‌ترین همسایه[۵۰] [۷۳]، استفاده از ماشین بردار پشتیبان[۵۱] [۷۴]، استفاده از درخت‌های تصمیم[۵۲] [۷۵]، میدان تصادفی شرطی سلسله‌مراتبی [۷۶] و دسته‌بندهای استفاده از فراسطح[۵۳] که از ترکیب نتایج چندین دسته‌بند پایه استفاده می‌کنند [۷۷]

[۴۳] Human activity recognition (HAR)
[۴۴] Naive bayes
[۴۵] Bayesian networks
[۴۶] Hidden markov model
[۴۷] Clustering
[۴۸] Partially observable
[۴۹] Coupled hidden markov model
[۵۰] Nearest neighbor
[۵۱] Support vector machine
[۵۲] Decision tree
[۵۳] Meta-level

نیز ارائه شده اند که توالی‌ای از مشاهدات حسگرها را به نزدیک‌ترین فعالیت انتساب می‌دهند.

راهکارهای دیگری نیز وجود دارند که از تکنیک‌های داده‌کاوی [۷۸]، استفاده از یادگیری استقرایی[۵۴] [۷۹، ۸۰] و استفاده از شبکه عصبی[۵۵][۸۱] بهره برده‌اند.

۳-۲-۲ راهکارهای دانش‌محور

راهکارهای دانش‌محور خود در دو دسته قرار دارند. دسته اول،راهکارهایی هستند که از منابع موجودی که مثل اسناد وب در دسترس عموم هستند استفاده می‌کنند [۸۲، ۸۳، ۸۴]. این راهکارها با استفاده از تکنیک‌های بازیابی اطلاعات تعاریف فعالیت‌ها را به دست می‌آورند و سپس با استخراج روابط آن‌ها فعالیت‌ها را مدل می‌کنند. دسته دوم، راهکارهایی هستند که یک فرد با دانش خبره، مدل فعالیت‌ها را به صورت دستی وارد می‌کند [۸۵، ۸۶، ۸۷، ۸۸].

راهکارهای دانش‌محور با استفاده از ابزارهای نمایش دانش برای مدل کردن فعالیت‌ها و تحلیل استفاده از استدلال منطقی[۵۶]، کار می‌کنند که هستی‌شناسی به دلیل سادگی و انعطاف بالا در ارائه فعالیت‌ها و روابطشان، استفاده بیشتری نسبت به سایر روش‌های نمایش دانش دارد [۸]. برخی از راهکارهای مبتنی بر هستی‌شناسی برای شناسایی فعالیت‌های عادی و روزمره، از عوامل زمینه‌ای بهره برده‌اند [۸۸، ۸۹، ۹۰] اما همبستگی زمانی را در نظر نگرفته‌اند.

در راهکار پیشنهادی توسط ریبونی[۵۷] و همکاران [۹۱] ویژگی‌های زمانی مثل استفاده‌ی اخیر در تعریف فعالیت‌ها آورده شده است. برخی دیگر از فعالیت‌های مبتنی بر هستی‌شناسی تمرکز روی مدل‌سازی فعالیت‌های کاربران، مستقل از شیوه‌ی انجام فعالیت توسط یک کاربر مشخص دارند [۹۲، ۹۳، ۹۴، ۹۵]. در نظر نگرفتن شیوه‌ی انجام فعالیت‌ها توسط هر کاربر باعث می‌شود تا امکان توسعه برنامه‌های مختلف با توجه به ترجیحات هر کاربر وجود نداشته باشد.

راهکار پیشنهادی توسط چن و همکاران [۸۸] فعالیت‌های کاربران را در دو سطح درشت‌دانه[۵۸] و ریزدانه[۵۹] به صورت انتزاعی مدل می‌کند. در مدل‌های درشت‌دانه فعالیت‌های کاربران با تعدادی ویژگی توصیف می‌شوند که این ویژگی‌ها نوع موجودیت مدنظر برای انجام یک فعالیت را توصیف می‌کنند. ولیکن در مدل‌های ریزدانه شیوه‌ی انجام فعالیت‌ها توسط هر کاربر را در نظر گرفته و توصیف می‌شود. این پژوهش

[۵۴] Inductive learning
[۵۵] Neural network
[۵۶] Logical reasoning
[۵۷] Riboni
[۵۸] Coarse-grained
[۵۹] Fine-grained

با استفاده از استنتاج روابط شمول[۶۰] و استفاده از منطق توصیفی[۶۱] الگوریتمی ارائه می‌دهد که فعالیت‌ها به صورت تدریجی شناسایی می‌شوند. این راهکار ابتدا از فعالیت‌های درشت‌دانه شروع کرده و خود را به فعالیت‌های ریزدانه می‌رساند و به طور کلی شناسایی فعالیت در هر دو سطح درشت‌دانه و ریزدانه انجام می‌گردد.

راهکار دیگری توسط اکیو و همکاران [۹۴] مطرح شده است که با استفاده از توالی زمانی بین چند فعالیت، یک فعالیت مرکب و همچنین موجودیت‌های درگیر در آن‌ها را در نظر می‌گیرد. مدیتسکوس[۶۲] و همکاران [۹۵] در پژوهش خود، علاوه بر فعالیت‌های مرکب، سلسله‌مراتب فعالیت‌ها را نیز در نظر گرفته‌اند.

۳-۲-۳ راهکارهای ترکیبی

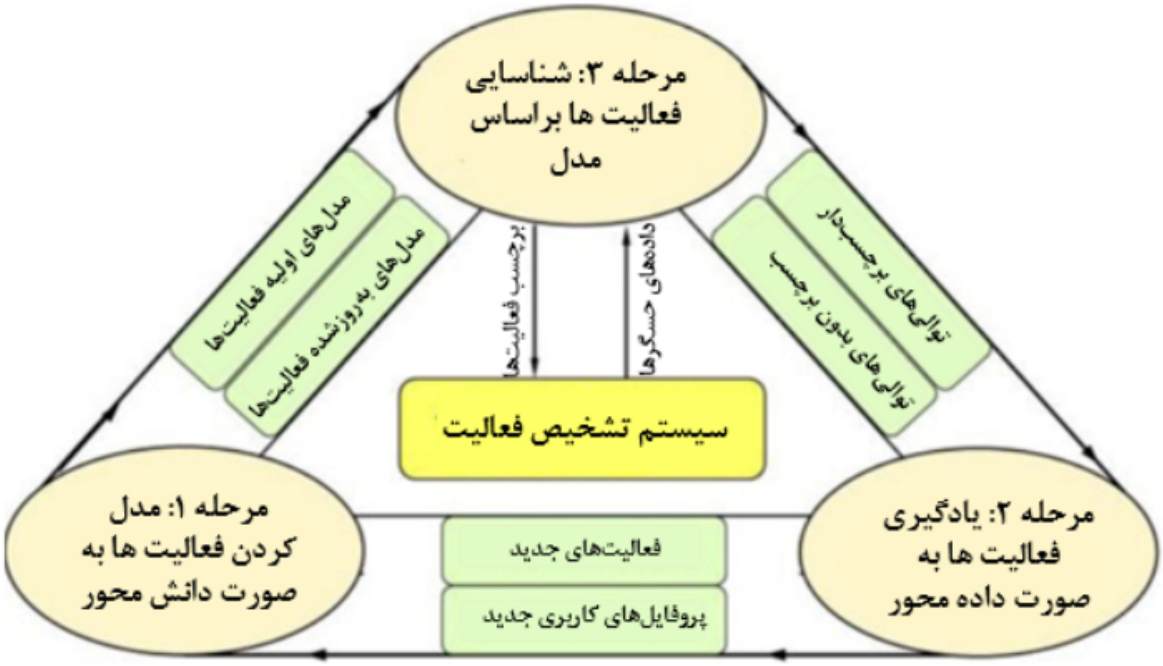
همانطور که در بخش ۲-۲ شرح داده شد هر یک از روش‌های داده‌محور و دانش‌محور معایبی دارند. یک سیستم تشخیص فعالیت باید تمامی فعالیت‌ها را با دقت بالا مدل کند. زیرا در صورت وجود فعالیتی مدل نشده، سیستم قادر به شناسایی دقیق آن فعالیت نیست. از طرفی در یک خانه هوشمند فعالیت‌های قابل انجام زیادی امکان‌پذیر است و وارد کردن تمامی آن‌ها به صورت دستی توسط فرد خبره عملا ناممکن است. همچنین دانش فرد خبره پویا نیست و تغییرات فعالیت‌ها و رفتارهای کلی کاربران را در دسترس ندارد. از طرف دیگر برای جلوگیری از شروع سرد در روند یادگیری روش‌های داده‌محور، به دانش فرد خبره نیاز است. به منظور فائق آمدن بر معایب این دو روش، راهکارهای ترکیبی ارائه شده‌اند.در جدول ۳-۲ می‌توان مقایسه‌ی کلی روش‌های ترکیبی را مشاهده کرد.

ریبونی و همکاران [۹۶] برای اولین بار از ترکیب روش‌های یادگیری ماشین (داده‌محور) و هستی‌شناسی (دانش‌محور) استفاده کردند. در این پژوهش با استفاده از روش‌های آماری تعدادی فعالیت به عنوان فعالیت احتمالی انجام شده انتخاب می‌شوند و سپس با استفاده از هستی‌شناسی، فعالیتی با احتمال وقوع بیشتر انتخاب می‌گردد.

چن و همکاران [۹۷، ۹۸] روش ترکیبی مطرح کردند که از یک فرایند سه مرحله‌ای تکرارشونده استفاده می‌کند. همانطور که در شکل ۳-۱ مشاهده می‌شود، در مرحله اول فرد خبره به صورت دستی دانش خود را برای ایجاد مدل‌های اولیه می‌دهد. در مرحله دوم تشخیص فعالیت‌ها با استفاده از دانش ارائه شده در مرحله اول انجام می‌گردد اما اگر فعالیتی مدل نشده باشد آن را شناسایی نمی‌کند و خروجی این مرحله

[۶۰] Subsumption reasoning
[۶۱] Description logic
[۶۲] Meditskos

سلسله داده‌هایی است که توسط حسگرها ارسال شده اما فعالیت مربوط به آن‌ها شناسایی نشده‌اند. در مرحله سوم با استفاده از روش‌های مبتنی بر داده‌کاوی، خروجی مرحله دوم را تحلیل کرده و پروفایل شخصی هر کاربر و همچنین فعالیت‌های جدید را یاد می‌گیرد. حال بر اساس شباهت سلسله داده‌های دریافتی از حسگرها، فعالیت‌ها گروه‌بندی شده و چنانچه اعضای یک گروه از تعداد مشخصی بیشتر شود، آن فعالیت‌های شناسایی شده به بخش هستی‌شناسی سیستم اضافه می‌شوند. برچسب‌گذاری هر فعالیت و محل انجامش در سلسله فعالیت‌ها به صورت دستی انجام می‌شود.



شکل ۳-۱: فرایند سه مرحله‌ای تکرارشونده‌ی مدل کردن فعالیت‌ها در راهکار چن و همکاران [۸۴]

عبدالصکور[۶۳] و همکاران [۹۹] برای غلبه بر مشکل ناکامل بودن و به‌روز نبودن دانش فرد خبره، از ترکیب هستی‌شناسی و دسته‌بندهای بیز ساده، ماشین بردار پشتیبان و شبکه عصبی پرسپترون چند لایه[۶۴] استفاده کرده‌اند. از آنجایی که به دلیل خرابی حسگرها و یا تغییر شیوه انجام فعالیت‌ها توسط کاربران سیستم دچار خطا و عدم قطعیت می‌شود، رودریگز و همکاران [۸] یک منطق فازی[۶۵] برای نمایش فعالیت‌ها ارائه کردند که امکان مدل‌سازی دانش غیرقطعی و مبهم را دارد. به این صورت که بر خلاف منطق دودویی که در آن هر عبارت به غلط یا درست قابل ارزیابی است، از منطق فازی[۶۶] استفاده می‌کند که میزان درستی هر عبارت مقداری بین صفر تا یک است.

گایاتری[۶۷] و همکاران [۱۰۰] از ترکیب هستی‌شناسی و شبکه منطق مارکوف[۶۸] استفاده کرده‌اند. به

[۶۳] A. Sukor
[۶۴] Multi-layer perceptron neural network
[۶۵] Fuzzy logic
[۶۶] Fuzzy logic
[۶۷] Gayathri
[۶۸] Markov logic network

این صورت که تی‌باکس به منطق مرتبه اول[۶۹] تبدیل می‌شود و با استفاده از نمونه‌های موجود در ای‌باکس، وزن‌ها محاسبه می‌شوند تا به فعالیت‌ها در شبکه منطق مارکوف، وزنی اختصاص یابد و توالی احتمالاتی فعالیت‌ها را داشته باشند.

راهکار دیگری توسط ریبونی و همکاران [۱۰۱] مطرح شد که با استفاده از استنتاج روی هستی‌شناسی، همبستگی معنایی بین فعالیت‌ها و نتایج رخداد هر یک استخراج می‌شود، سپس با استفاده از اطلاعات به دست آمده، فعالیت کاندید بعدی را شناسایی می‌کند. برخی فعالیت‌ها الگوهای مشابهی دارند و سیستم‌های مطرح شده در تشخیص دقیق این فعالیت‌ها ممکن است با مشکل مواجه شوند. در راهکار دیگر که توسط بتینی[۷۰] و همکاران [۱۰۲] مطرح شده است، شناسایی فعالیت‌ها به کمک حسگرهای تلفن همراه انجام می‌شود و برای این منظور از یادگیری نیمه‌نظارتی[۷۱] و استنتاج مبتنی بر هستی‌شناسی و همچنین اطلاعات زمینه‌ای استفاده نی‌شود. در این راهکار داده‌های دریافت شده از حسگرها که با اطلاعات زمینه‌ای نظیر مکان کاربر، طبق روابط معنایی تعریف شده در هستی‌شناسی همخوانی ندارند حذف شده و در گام آخر، اگر میزان اطمینان فعالیت شناسایی شده از حد مشخصی کمتر باشد، از کاربر استعلام فعالیت کنونی را گرفته و برچسب‌گذاری انجام می‌گردد. با انجام این فرایند، مدل همواره در حال یادگیری و به‌روز شدن است.

۳-۲-۴ جمع‌بندی

در این بخش، پژوهش‌های انجام شده در حوزه مدل‌سازی و تشخیص فعالیت‌های کاربران در خانه‌های هوشمند بررسی و از نظر نوع رویکرد در دریافت اطلاعات مورد نیاز دسته‌بندی شدند. مقایسه‌ی کلی پژوهش‌های مطرح شده در جدول ۳-۲ قابل مشاهده است.

۳-۳ حفظ حریم خصوصی مبتنی بر سکوی نامعتمد

تاکنون پژوهش‌های زیادی در جهت حفظ حریم خصوصی کاربر در حوزه‌های مختلف صورت پذیرفته است. مانند تحقیقاتی که در حوزه حفظ حریم خصوصی کاربر در داده کاوی[۷۲]، انتشار اطلاعات، واکشی اطلاعات، و شبکه‌های ناامن صورت گرفته است. از منظر سکوهای اینترنت اشیاء نیز حفظ حریم خصوصی کاربر در ابعاد مختلف بررسی شده است مانند امنیت در محل ذخیره سازی داده‌ها، امنیت در واکشی،

[۶۹] First order logic
[۷۰] Bettini
[۷۱] Semi-supervised learning
[۷۲] Data mining

جدول ۳-۲: مقایسه کلی روش‌های ترکیبی

پژوهش	راهکار	چالش‌ها							
		محدودیت‌های راهکارهای دانش‌محور			محدودیت‌های راهکارهای داده‌محور			پیچیدگی‌های شناسایی فعالیت	
		ناکامل بودن دانش فرد خبره	عدم به روز شدن خودکار	عدم توانایی در مواجهه با عدم قطعیت	شروع سرد	عدم امکان استفاده مجدد	ناکامل بودن مدل ایجاد شده توسط روش‌های یادگیری	امکان حضور چند کاربر	شناسایی فعالیت‌های همزمان
چن و همکاران [۹۷، ۹۸]	یادگیری مبتنی بر داده‌کاوی	✓	✓	×	✓	✓	✓	×	×
عبدالصکور و همکاران [۹۹]	ماشین بردار پشتیبان و شبکه عصبی پرسپترون	✓	✓	×	✓	✓	✓	×	×
رودریگز و همکاران [۸]	منطق فازی	×	×	✓	✓	✓	✓	✓	✓
گایاتری و همکاران [۱۰۰]	شبکه منطق مارکوف	×	×	✓	✓	✓	✓	×	✓
ریبونی و همکاران [۱۰۱]	روش‌های آماری و شبکه منطق مارکوف	×	×	×	✓	✓	✓	×	✓

اعتبارسنجی داده‌هایی که در سکو ذخیره می‌گردد و حفظ محرمانگی داده‌های کاربران در سرویس‌دهنده و یا سکوی نامعتمد [۱۰۳]. این موضوع بیانگر جنبه‌های امنیتی گسترده‌ای است که کاربر را با تهدید مواجه می‌نماید.

راهکارهایی که در هر یک از حوزه‌های مورد اشاره ارائه شده است بعضاً قابل تعمیم به حوزه‌های دیگر هم هستند. برای مثال در حوزه داده کاوی (که بر ذخیره داده‌های کاربران در طول زمان بر روی سکوهای ابری و انجام استنتاج‌های مربوطه و کلاسه‌بندی‌های مورد نیاز، اشاره دارد) یکی از تکنیک‌های حفظ امنیت کاربران، جلوگیری از استنتاج‌های غیر ضروری با استفاده از آشفته‌سازی داده‌های ذخیره شده در سکو است. همین راهبرد در برخی از راهکارهای حفظ حریم خصوصی کاربر در انتشار داده، به کار رفته است. ضمنا باید به این نکته نیز توجه داشت که پیاده‌سازی همه راهکارها، در اختیار کاربر نیست و در برخی راهکارها مستلزم همکاری سرویس‌دهنده، سکو و یا ارائه دهنده خدمات شبکه است. به طور مثال استفاده از رمزنگاری برای حفظ حریم خصوصی نیازمند همکاری بخش‌های مختلفی در زیرساخت اینترنت اشیاء است ولیکن روشی مانند آشفته‌سازی داده‌ها می‌تواند توسط عوامل تحت اختیار کاربر انجام پذیرد.

بسیاری از پژوهش‌های حفظ حریم خصوصی، با مدل تهدید سکوی نامعتمد انجام شده است که بر اساس راهکار، به دسته‌های مبتنی بر رمزنگاری، مبتنی بر کمینه‌سازی، مبتنی بر آشفته‌سازی داده و مبتنی بر تولید رویداد جعلی تقسیم می‌شوند. در این مدل تهدید فرض می‌شود که داده‌های حساس همواره در معرض خطر قرار دارند همانطور که تعداد بسیاری حمله با استفاده از این داده‌های حساس انجام شده است [۱۰۴، ۱۰۵]. به دلیل متمرکز بودن سکوها و دسترسی به تمامی داده‌های حساس توسط سکوها، مهاجمین توجه ویژه‌ای به آن‌ها می‌کنند. ضمن این که سکوها با داشتن دسترسی به تمامی اطلاعات خانه‌هوشمند و جمع‌آوری داده‌های محرمانه کاربر (که بعضا به آن نیاز ندارند)، می‌توانند از آن‌ها برای اهدافی مانند تبلیغات یا فروش به شرکت‌های دیگر استفاده کنند [۱۰۶، ۱۰۷]. راهکارهای ارائه شده در این حوزه را می‌توان به دسته‌های مختلفی تقسیم نمود که در ادامه به معرفی راهکارهای ارائه شده در هر دسته و تحلیل آن‌ها پرداخته شده است.

۳-۳-۱ راهکارهای مبتنی بر رمزنگاری

یکی از روش‌های حفظ حریم خصوصی در برابر سکوی نامعتمد، استفاده از یک یا چند روش رمزنگاری روی داده‌هاست به نحوی که تمامی داده‌ها از دید سکو پنهان شود.

شوتلر[۷۳] و همکاران [۱۰۸] سکویی دو بخشی به نام والنات[۷۴] تعریف کرده‌اند که هیچ کدام دسترسی به داده‌های حساس بخش دیگر ندارند. در این روش از تلفیق محاسبه دو جانبه امن و محیط اجرای امن به ترتیب برای حفظ محرمانگی و صحت استفاده می‌شود. این پژوهش فقط مدل ارتباطی رویداد_کنش را در نظر گرفته است و مدل ارتباطی رویداد_محاسبه_کنش را در نظر نمی‌گیرد.

در پژوهش دیگری، زاوالیشن[۷۵] و همکاران [۱۰۹] از طریق معرفی یک سکوی معتمد با معماری جدید به نام پاترآی‌اتی[۷۶]، سعی بر حفظ حریم خصوصی کاربر را دارند. این پژوهش مانند والنات از محیط اجرای امن استفاده کرده است. همچنین این پژوهش یک لایه محافظتی مبتنی بر خط مشی ارائه کرده است که کنترل جریان‌های داده از دستگاه‌ها تا سکو و حذف جریان‌های نامطلوب از دید کاربر را بر عهده دارد.

در پژوهش دیگری، چیانگ[۷۷] و همکاران [۱۱۰] معتقدند که سکو به کمک داده‌های دریافتی، برای هر کاربر پروفایل اختصاصی درست می‌کند. داده‌های ارسال شده توسط هر نرم‌افزار، استفاده یا عدم استفاده از دستگاهی خاص مانند دستگاه اندازه‌گیری قند خون و حتی عدم دریافت داده مورد انتظار از یک برنامه در زمان مشخص، اطلاعات حساسی هستند که سکو به آن‌ها دسترسی دارد. برای مبهم‌سازی اطلاعات در این پژوهش دو راهکار OTAP[۷۸] و ATAP[۷۹] ارائه شده است. OTAP اطلاعات وقوع یا عدم وقوع رخدادها را از سکو پنهان می‌کند. با استفاده از رمزنگاری انتها به انتها[۸۰] بین رویداد و کنش، اطلاعات را از سکو پنهان می‌کند. ATAP علاوه بر انجام راهکارهایی که OTAP ارائه می‌دهد، اطلاعات مالکیت را نیز پنهان می‌کند. با استفاده رمزنگاری متقارن بین سرویس رویداد و کنش، سعی در مخفی نگه داشتن پروفایل کاربران دارد و داده‌ای که در اختیار سکو قرار می‌گیرد تنها به سرویس کنش تحویل داده می‌شود. این پژوهش نیز مانند پژوهش ژو[۸۱] و همکاران [۱۱۱]، دستگاه‌هایی که مستقیما و بدون واسط با سکو در ارتباط هستند را در نظر نگرفته است. در این پژوهش تنها الگوی رویداد_کنش در نظر گرفته شده است در صورتی که ممکن است الگوی رویداد_محاسبه_کنش مورد استفاده باشد.

در پژوهش دیگری، چن و همکاران [۱۱۲] eTAP[۸۲] را معرفی کرده‌اند. سکوی معرفی شده محاسبات مورد نیاز را روی داده‌های رمز شده انجام داده و قابلیت استنتاج از نتیجه‌ی محاسبات را ندارد. در مدل تهدید این پژوهش مهاجم ممکن است فعال باشد. ضعف این پژوهش عدم پنهان‌سازی وقوع یا عدم وقوع رخدادهاست که علی‌رغم رمزنگاری، توسط مهاجم قابل تشخیص است.

---

[۷۳] Schoettler
[۷۴] Walnut
[۷۵] Zavalyshyn
[۷۶] PatrIoT
[۷۷] Chiang
[۷۸] Obfuscated Trigger-Action Platform
[۷۹] Anonymous Trigger-Action Platform
[۸۰] End-to-End
[۸۱] Xu
[۸۲] Encrypted Trigger Action Platform

۳-۳-۲ راهکارهای مبتنی بر کمینه‌سازی

در روش‌های ارائه شده مبتنی بر کمینه‌سازی، به طور کلی در خروجی سیگنال‌های سری زمانی حسگرهای یک خانه هوشمند، بخش مربوط به فعالیت‌های حساس کاربر حذف می‌گردد و به جای آشفته سازی، خروجی برش می‌خورد. برای مثال در یکی از راهکارهایی که در این دسته قرار می‌گیرد انتشار جریان زمینه کاربر برای حفظ حریم خصوصی وی به انتخاب کاربر منقطع می‌گردد و در هر زمینه جدید، در مورد انتشار و یا عدم انتشار اطلاعات تصمیم گیری می‌شود [۱۱۳]. به عنوان مثالی در این حوزه، فرض کنید که کاربر مایل باشد که از سرویس ساکت شدن زنگ موبایل خود هنگامی که در جلسه کاری است استفاده کند اما در لحظات دیگر تمایلی نداشته باشد که سرویس دهنده از شرایط محلی که وی در آن حضور دارد مطلع شود و مثلاً متوجه نشود که کاربر در خانه است، در حال رانندگی کردن است و یا در حال قدم زدن است.

روش کمینه‌سازی شناسایی فعالیت‌های حساس کاربر را برای مهاجم دشوار می‌نماید اما همچنان فاصله‌های زمانی انجام فعالیت‌های حساس کاربر و در مواردی حتی شناسایی خود فعالیت‌ها را برای مهاجم میسر می‌نماید. مثال‌های گوناگونی در این حوزه وجود دارند که چگونگی نشت اطلاعات محرمانه از روی داده‌های غیر محرمانه را به کمک روش‌های مهندسی اجتماعی و یا مدل‌سازی رفتار کاربر در طول زمان توسط مهاجم (به کمک مدل زنجیره مارکوف) را نشان می‌دهند [۱۱۳]. در ضمن با حذف بخشی از سری زمانی مربوط به سیگنال‌ها، بخشی از اطلاعات مربوط به فعالیت‌های غیرحساس کاربر نیز ممکن است حذف شود و به این ترتیب از کارایی داده‌های ارسالی به سکوهای اینترنت اشیاء کاسته می‌شود. همچنین تکنیک‌های کمینه‌سازی که بخواهند تحلیل کاملی را روی داده قبل از ارسال آن انجام دهند ممکن است با مشکلات پردازشی و تجربه بد کاربری روبرو شوند. مشکل دیگر روش‌های کمینه‌سازی این است که برای دریافت برخی سرویس‌ها (مثلا اتوماسیون امور با استفاده از برنامه‌های اینترنت اشیاء) لازم است که داده‌هایی به سکو الزاما ارسال شود که از دید کاربر بخشی از آن‌ها حساس و محرمانه است. لذا یک تعارض بین دریافت سرویس از سکوی اینترنت اشیاء و حفظ حریم خصوصی پیش می‌آید که این دسته از روش‌ها قادر به رفع این تعارض نیستند و باید از روش‌های دیگری برای حفظ حریم خصوصی استفاده کرد.

ژو و همکاران [۱۱۱] ارتباط بین سکوی اسمارت‌تینگز و سکوهای شخص ثالث مانند ایفت را بررسی کرده‌اند و به این نتیجه رسیده‌اند که سکوی شخص ثالث به اطلاعات زیادی دسترسی دارد که به نوعی نقض حریم خصوصی کاربر محسوب می‌شود. در این پژوهش با استفاده از ماژولی به نام F&F[۸۳] برای مبهم‌سازی الگوی اطلاعات دریافتی، اطلاعات جعلی ایجاد می‌شود. همچنین در مواردی که سکو برای کنش، نیاز به اطلاعات دقیق ندارد، داده‌های تقریبی و دستکاری شده برای سکو ارسال می‌گردد. مهمترین چالش این پژوهش این است که سکوی اسمارت‌تینگز مورد اعتماد فرض شده و فقط سکوی ثالث نامعتمد

[۸۳] Filter & Fuzz

است.

در پژوهشی دیگر، چی[۸۴] و همکاران [۱۱۴] سیستم کنترل جریان داده‌ای به نام پی‌فایروال[۸۵] معرفی کرده‌اند. این ابزار ابتدا کد برنامه‌ی اینترنت اشیاء در هر دستگاه را بررسی کرده و سپس داده‌ی مورد نیاز برای ارسال رویدادهای هر یک را استخراج می‌کند. با این کار تنها داده‌های مورد نیاز به سکو ارسال شده و از ارسال داده‌های اضافی جلوگیری می‌شود. مزیت این ابزار آن است که نیازی به تغییر سکو، هاب و یا دستگاه‌ها نیست و ابزار پی‌ـ‌فایروال بین هاب و سکو قرار می‌گیرد. همانطور که در بخش ۲-۴ گفته شد برخی دستگاه‌ها مستقیما با سکو در ارتباط هستند و این پژوهش این دسته از دستگاه‌ها را در نظر نگرفته است. همچنین ممکن است کد برنامه‌ی اینترنت اشیاء در تمامی دستگاه‌ها در دسترس نباشد و پی‌فایروال قادر به استخراج اطلاعات مورد نیاز نباشد.

در پژوهشی دیگر، چن و همکاران [۱۱۵] با ارائه مین‌تپ[۸۶] کارایی روش کمینه‌سازی را افزایش دادند. در این پژوهش کاربر قواعد را تعریف کرده و سپس با پردازش روی این قواعد، اطلاعات کمینه‌سازی استخراج شده و به همراه قواعد به سکو ارسال می‌شود. هنگام رخداد رویداد، خصیصه‌های آن به همراه کمینه‌ساز به سرویس‌دهنده ارسال شده و اطلاعات اضافه از رویداد حذف شده و سپس به سکو ارسال می‌گردد. این راهکار کارایی بالا و سربار قابل قبولی دارد اما نیاز به اعمال تغییرات سمت سرویس‌دهنده است.

راهکار دیگری که مبتنی بر کمینه‌سازی ارائه شده است، استفاده از نگاشت[۸۷] است که در این روش‌ها به کمک اعمال محدودیت‌ها بر روی سری‌های زمانی، سری‌های زمانی جدیدی از خروجی حسگرها ایجاد می‌گردد [۱۱۶]. و غالباً اطلاعات سری زمانی اولیه حسگرها به فضای حالتی با ابعاد کوچکتر نگاشت می‌گردد تا از میزان اطلاعاتی که به همراه دارد کاسته شود. در این روش به جای ارسال سری زمانی با ابعاد بالا (مثلا یک حسگر مربوط به سیستم تهویه مطبوع ممکن است ویژگی‌هایی مانند وضعیت عملکرد فعلی، شدت فن، و دمای محیط را ارسال نماید)، فقط بخشی از ویژگی‌های اصلی از خروجی حسگرها استخراج شده و ارسال می‌شود به نحوی که داده‌ها فقط برای استنتاج فعالیت‌های غیرحساس مفید واقع شود. در یکی از راهکارهایی که در این دسته پیشنهاد شده است، داده‌های کاربر قبل از ارسال، پیش پردازش می‌گردد تا فقط داده‌هایی که برای دریافت سرویس‌های داده کاوی خاص از سکو مورد نیاز است، ارسال شود. بدین ترتیب برای مثال اگر تجهیزی برای شناسایی حرکت در خانه استفاده می‌گردد، برای استنتاج این موضوع که آیا شخصی در خانه حضور دارد یا خیر نمی‌تواند مورداستفاده قرار گیرد. در این راهکار از یک ماژول استخراج کننده ویژگی به کمک شبکه عصبی استفاده شده است تا با کمک تکنیک‌های کاهش

[۸۴] Chi
[۸۵] PFirewall
[۸۶] minTAP
[۸۷] Mapping

ابعاد، ساختار داده اصلی حفظ شود و فقط آنچه لازم نیست حذف گردد و این امر با کم کردن فاصله معنایی ویژگی‌های مربوط به هم و افزایش فاصله معنایی ویژگی‌های غیرمرتبط با هم صورت می‌گیرد. مثلا جنسیت کاربر توسط سکو شناسایی می‌شود اما تصویر وی شناسایی نمی‌گردد [۱۱۶].

قسمت دشوار استفاده از راهکار نگاشت، انتخاب بهترین مجموعه کوچک ویژگی به خصوص در روش‌های متکی بر یادگیری ماشین است که بتواند یک توازن منطقی بین کارایی داده‌های ارسالی و حفظ حریم خصوصی ایجاد نماید. در راهکار دیگری که در این دسته ارائه شده است داده‌های مربوط به سری زمانی در هر مقطع از زمان طی یک فرآیند آماری به صورت تصادفی نگاشت می‌گردد و این نگاشت به نحوی صورت می‌گیرد که کاربردپذیری خروجی را به صورت حداکثری نگاه دارد. هدف از انجام این نگاشت این است که نمونه‌های داده سری زمانی در طول زمان مستقل از هم بشوند و همبستگی خود را از دید مهاجم بیرونی مخفی نمایند [۱۱۷]. چالشی که در این دسته از راهکارها وجود دارد این است که ممکن است برخی از سرویس‌های سکوها، با سیگنال‌هایی که ویژگی‌های اصلی آن‌ها تغییر یافته است نتوانند کار کنند. در ضمن در این روش اصلا نمی‌توان تضمین نمود که هر ویژگی در مجموعه‌ای که انتخاب شده است تنها حاوی اطلاعات یک فعالیت غیرحساس باشد و هیچ اطلاعی از یک فعالیت حساس را به همراه نداشته باشد.

۳-۳-۳ راهکارهای مبتنی بر آشفته‌سازی

در این روش هدف ارسال داده تغییریافته است به شکلی که کاربر بتواند سرویس‌های دلخواه خود را از سکو یا سرویس دهنده دریافت نماید و در عین حال عدم قطعیت داده تضمین گردد [۱۱۸]. منظور از عدم قطعیت داده این است که از روی داده‌های آشفته، ساخت اصل داده‌ها میسر نباشد.

تنوع راه‌حل‌ها و چالش‌ها در این دسته از راهکارها زیاد است و بررسی‌های متعددی در این خصوص صورت گرفته است برای مثال مطالعاتی در خصوص اعمال انواع مختلف نویز به داده‌های سری زمانی حسگرها، مانند اعمال نویز تصادفی با الگوریتم‌های مختلف [۱۱۹، ۱۲۰]، اعمال نویزی که با سری زمانی اصلی همبستگی دارد [۱۲۱]، اعمال نویز در کنار فشرده‌سازی سیگنال اولیه [۱۱۹]، و یا اعمال نویز با توجه به حفظ فاصله اقلیدسی سیگنال‌های حسگرهای مشابه از کاربران مختلف [۱۲۰] صورت گرفته است و مزایا و معایب هر روش پیشنهادی با توجه به معیارهایی که برای اندازه‌گیری حفظ حریم خصوصی کاربر و کاربردپذیری سیگنال تعریف شده اندازه‌گیری شده است.

در کنار این راهکارها، روش‌هاي متعددی نیز برای بازسازی داده‌های اولیه از روی داده‌های آشفته‌سازی‌شده پیشنهاد شده است که نقاط ضعف استفاده از این راهکارها را به چالش کشیده است [۱۲۰، ۱۲۲، ۱۲۳]. به

طور خلاصه در مورد این راه‌حل‌ها می‌بایست گفت که میزان نویز اعمال شده به سیگنال اولیه، قابلیت کاربرد پذیری سیگنال (منظور امکان استفاده کاربر از سرویس‌های دلخواه) را به شدت تحت تاثیر قرار می‌دهد و حتی ممکن است اصل داده‌ها را تخریب نماید. بسیاری از راهکارهای این دسته شناسایی فعالیت‌های حساس کاربر را برای مهاجم با دشواری مواجه می‌کنند اما عدم قطعیت داده را تضمین نمی‌کنند.

راهکار دیگری برای آشفته‌سازی داده، استفاده از جایگزینی[۸۸] است. در این روش، بخش حساس سری‌های زمانی با داده‌های غیرحساس جایگزین می‌گردد. برای مثال در یکی از رویکردهایی که در این دسته ارائه شده است [۱۲۴] به کمک تکنیک‌های یادگیری ماشین، رفتارهای حساس کاربر با رفتارهای غیرحساس جایگزین می‌گردد. در این راهکار فعالیت‌های کاربر به دو گروه سفید (غیر حساس) و سیاه (حساس) تقسیم می‌گردد. منظور از رفتارهای حساس کاربر رفتارهایی هستند که برخی از حالات روانشناسی کاربر مانند وجود استرس از آن‌ها قابل استخراج است. شبکه پویای بیزین پیشنهادی در این روش، آموزش می‌بینند و در مرحله اول برای شناسایی بلاک‌های حساس کاربر در سیگنال‌های خروجی استفاده می‌گردد. سپس در مرحله بعد برای انتخاب بخش‌هایی که می‌بایست جایگزین بلاک‌های حساس گردد، استفاده می‌گردد. البته با حفظ این محدودیت که طول بلاک‌هایی که حذف می‌گردد با طول بلاک‌هایی که جایگزین می‌گردد مساوی باشد. در این راهکار توسط ماژولی سیگنال خام ورودی به سیگنالی مبتنی بر ویژگی‌های رفتاری تبدیل می‌شود و با حذف رفتارهای حساس تعدادی حفره باقی می‌مانند که می‌بایست توسط رفتارهای غیرحساس پر شوند و جایگزینی به نحوی صورت می‌گیرد که به تداوم جریان سیگنال لطمه وارد نشود. برای جایگزینی رفتارها از یک پایگاه داده نگاشت رفتاری با طول‌های زمانی مختلف استفاده می‌گردد. این راهکار به صورت برون‌ـخط عمل می‌کند یعنی قبل از تصمیم‌گیری برای جایگزینی، اطلاعات سری‌زمانی فعالیت کاربر می‌بایست در دسترس باشد.

به عنوان مثالی دیگر از راهکارهای این دسته به راهکار ارائه شده در پژوهش آقای ملک زاده [۱۲۵] می‌توان اشاره کرد که در آن یک شبکه عصبی آموزش دیده به نام خودرمزگذار[۸۹] جایگزین‌کننده، عملی مانند رفع نویز انجام می‌دهد و الگوریتم جایگزینی فعالیت‌های حساس با فعالیت‌های غیرحساس را انجام می‌دهد به نحوی که امکان تشخیص و استنتاج فعالیت‌های حساس از بین می‌رود. داده‌ها در این راهکار به سه دسته‌ی سیاه (حساس)، خاکستری (غیر حساس است و استنتاج این رفتار برای کاربر مهم نیست) و سفید (غیر حساس است و استنتاج این رفتار برای کاربر مهم و کاربردپذیر است) تقسیم می‌شوند. در هر سری زمانی و بازه‌ی دلخواه، الگوریتم مطرح شده با در نظر گرفتن کاربردپذیری از بعد دریافت سرویس‌های مد نظر کاربر از سکو، داده‌هایی از لیست سیاه را با خاکستری جایگزین می‌کند. در این راهکار اثبات می‌شود که مهاجم بدون داشتن داده‌های مورد استفاده برای تعلیم لیست خاکستری، امکان تشخیص داده‌های حساس

[۸۸] Replacement
[۸۹] AutoEncoder

را ندارد و از کاربردپذیری داده‌های ارسالی به سکوهای اینترنت اشیاء برای دریافت سرویس‌های دلخواه کاسته نمی‌شود.

۳-۳-۴ راهکارهای مبتنی بر تولید رویدادهای جعلی

در پژوهش‌های مبتنی بر تولید رویداد جعلی، برای حفظ حریم خصوصی کاربر، به جای حذف، تغییر یا پنهان‌سازی داده، از اضافه کردن داده‌های جعلی برای گمراه کردن سکو استفاده می‌شود. تولید و ارسال داده‌های جعلی در این پژوهش‌ها به نحوی است که سکو امکان تمایز این داده‌ها با داده‌های واقعی را نداشته باشد و از طرفی کنش‌های سکو در جواب رویدادهای جعلی اعمال نشود.

در پژوهش ژو و همکاران [۱۱۱] الگوهای آماری باعث نقض حریم خصوصی می‌شود و برای عدم استنتاج سکو، تعدادی رویداد جعلی با برچسب مشخص ارسال می‌شود و از آن‌جایی که سکو قادر به تمایز میان رویدادهای جعلی و حقیقی نیست، کنش‌های مرتبط را ایجاد می‌کند. در آخر، با توجه به برچسب هر رویداد و وابستگی کنش‌ها به رویدادها، کنش‌های ناشی از رویدادهای جعلی تشخیص داده شده و کنار گذاشته می‌شوند.

در پژوهش چیانگ و همکاران [۱۱۰]، محرمانگی داده‌های ارسالی به سکو تضمین شده است اما در بسیاری از موارد، وقوع یا عدم وقوع یک رویداد، یک داده‌ی حساس به شمار می‌آید. برای پنهان کردن این اطلاعات، از رویکرد تولید رویداد جعلی استفاده شده است که به صورت متناوب، در بازه‌های زمانی مشخص، هر دستگاه اینترنت اشیاء، داده‌های جعلی و حقیقی را با هم به سکو ارسال می‌کند.

اقوامی‌پناه و امینی[۹۰] [۱۲۶] برای فریب مهاجمی که با استفاده از شبکه عصبی LSTM[۹۱] اقدام به استنتاج فعالیت‌های کاربر می‌کند، راهکاری بر اساس تولید رویدادهای جعلی ارائه کرده‌اند[۹۲]. در این پژوهش با استفاده از نمونه خصمانه[۹۳] و تولید رویدادهای جعلی در بازه‌های زمانی مشخص، جلوی استنتاج مهاجم گرفته می‌شود. همچنین، کنش‌هایی که در نتیجه‌ی رویدادهای جعلی دریافت می‌شود کنار گذاشته می‌شود تا کاربردپذیری خانه هوشمند کاهش پیدا نکند. هرچند این پژوهش در ابتدا فرض را بر استفاده‌ی سکو از شبکه عصبی LSTM برای استنتاج فعالیت‌های کاربر قرار داده است؛ ولیکن در ادامه با بررسی قابلیت انتقال‌پذیری، امکان موثر بودن نمونه‌های خصمانه تولید شده در به اشتباه انداختن دسته‌بندهای مبتنی بر روش‌های دیگر را با انجام آزمایش‌های عملی به اثبات رسانده است. یکی از ضعف‌های این پژوهش این

[۹۰] آزمایشگاه امنیت داده و شبکه، دانشگاه صنعتی شریف
[۹۱] Long-Short Term Memory
[۹۲] مقاله پژوهش تا زمان نگارش این مطلب منتشر نشده است.
[۹۳] Adversarial example

است که در صورت هوشمندی بیشتر سکوی مهاجم و تجهیز آن به برخی روش‌های شناسایی ناهنجاری، امکان تشخیص و شناسایی رویدادهای جعلی تولید شده با این روش و حذف آن‌ها وجود خواهد داشت. ضعف دیگر، وجود روش‌های دفاعی [۱۲۷] در مقابل نمونه‌های خصمانه است که موجب ناکارآمد شدن این راهکار در برابر مهاجم هوشمند می‌شود.

۳-۳-۵ جمع‌بندی

در این بخش پژوهش‌های حفظ حریم خصوصی مبتنی بر سکوی نامعتمد بررسی شد که اکثرا نیاز به تغییر سکو و دستگاه‌ها دارند که سبب مشکل در استفاده شده است. در جدول ۳-۳ می‌توان مقایسه کلی این روش‌ها را مشاهده کرد.

جدول ۳-۳: مقایسه کلی روش‌های حفظ حریم خصوصی مبتنی بر سکوی نامعتمد

پژوهش	راهکار	پشتیبانی از محاسبه	عدم تغییر سرویس‌ها	عدم تغییر سکو	سربار قابل قبول	حفظ کاربردپذیری	حفاظت در برابر مهاجم هوشمند
ژو و همکاران (F&F) [۱۱۱]	کمینه‌سازی، آشفته‌سازی و تولید رویدادهای جعلی	✓	✓	✗	✓	✓	✗
چی و همکاران (PFirewall) [۱۱۴]	کمینه‌سازی	✓	✓	✓	✓	✗	✗
چیانگ و همکاران (OTAP) [۱۱۰]	رمزنگاری و تولید رویدادهای جعلی	✗	✗	✓	✗	✓	✓
چیانگ و همکاران (ATAP) [۱۱۰]	رمزنگاری و تولید رویدادهای جعلی	✗	✗	✗	✓	✓	✓
چن و همکاران (ETAP) [۱۱۲]	رمزنگاری	✓	✗	✗	✗	✓	✓
چن و همکاران (minTAP) [۱۱۵]	کمینه‌سازی	✓	✗	✓	✓	✗	✗
زاوالیشن و همکاران (PatrIoT) [۱۰۹]	رمزنگاری	✓	✓	✗	✗	✓	✓
شوتلر و همکاران (Walnut) [۱۰۸]	رمزنگاری	✓	✗	✗	✗	✓	✓
ملک‌زاده و همکاران [۱۲۵]	آشفته‌سازی	✓	✓	✓	✗	✗	✗
اقوامی‌پناه و امینی [۱۲۶]	تولید رویدادهای جعلی	✓	✓	✓	✗	✓	✗

فصل ۴

راهکار پیشنهادی

همانطور که در فصل ۳ مشاهده کردیم، پژوهش‌های زیادی در زمینه حفظ حریم خصوصی در خانه هوشمند انجام شده است اما هر یک نواقصی داشتند. راهکار پیشنهادی این پژوهش، استفاده از تولید رویداد جعلی بر پایه هستی‌شناسی مبتنی بر زمینه به صورت خاص‌دامنه برای خانه‌های هوشمند است تا سکوی نامعتمد، امکان تمایز بین داده‌های حساس کاربر و داده‌های جعلی را نداشته باشد. این رویدادهای جعلی با استفاده از هستی‌شناسی که با استفاده از روش دانش‌محور توسط فرد خبره به برنامه داده شده است، تولید می‌شوند.

۴-۱ توصیف اجمالی

راهکار ارائه شده با توجه به دسته‌بندی بخش ۳-۱ یک هستی‌شناسی مبتنی بر زمینه ارائه می‌دهد. هستی‌شناسی در این راهکار شامل نقشه‌ی خانه هوشمند، موقعیت موجودیت‌ها (اشیاء و انسان‌ها) و حالت آن‌ها، فعالیت‌های کاربر و شروط و نتایج هر فعالیت و همچنین فعالیت‌های احتمالی آتی بعد از هر فعالیت می‌باشد. بنابراین، هستی‌شناسی در این راهکار یک هستی‌شناسی مبتنی بر زمینه بوده که به صورت خاص‌دامنه برای خانه‌های هوشمند طراحی شده است.

مدل‌سازی اطلاعات در این راهکار با توجه به دسته‌بندی در بخش ۳-۲، با استفاده از روش دانش‌محور می‌باشد زیرا فرد خبره با دانش خود اطلاعات کامل خانه، موجودیت‌ها و فعالیت کاربران را داشته و به برنامه می‌دهد. در این روش شروع سرد نخواهیم داشت و داده‌ها قابلیت استفاده مجدد دارند زیرا در صورت نیاز به استفاده مجدد، تنها بخش کمی از هستی‌شناسی نیاز به تغییر دارد. برای مثال در صورت تغییر منزل توسط کاربران، تنها بخش نقشه خانه و موقعیت موجودیت‌ها در هستی‌شناسی عوض شده و سوابق فعالیت‌های کاربر در هستی‌شناسی ثابت می‌ماند. توجه شود که این راهکار دچار مشکل ناکامل بودن دانش فرد خبره

نمی‌شود چرا که برای تولید فعالیت‌های جعلی، نیازی به آگاهی از تمامی فعالیت‌های کاربر نداریم و تنها با داشتن فعالیت‌های روزانه و پرتکرار کاربران، امکان تولید سلسله فعالیت جعلی برای برنامه فراهم است.

در این پژوهش برای حفظ حریم خصوصی در برابر سکوی نامعتمد، با توجه به دسته‌بندی در بخش ۳-۳، از راهکار مبتنی بر تولید رویداد جعلی استفاده شده است. با استفاده از هستی‌شناسی مبتنی بر زمینه که شامل سلسله فعالیت‌های کاربران و توالی احتمالی هر دو فعالیت است، برای گمراه کردن سکوی نامعتمد در این پژوهش سلسله فعالیت جعلی مبتنی بر هستی‌شناسی تولید شده که از دید سکو قابل تمایز نباشد. توجه شود که مانند پژوهش ژو و همکاران [۱۱۱]، رویدادهای جعلی با برچسب ارسال می‌شود و کنش‌های مربوط به رویدادهای جعلی کنار گذاشته می‌شوند.

۴-۲ مدل تهدید

در این پژوهش، سکوی اینترنت اشیاء صادق ولی کنجکاو فرض شده است و دسترسی به تمامی رویدادهای ارسال شده از حسگرهای خانه هوشمند را داشته و به طور کامل آن‌ها را به همراه زمان دریافت داده از حسگر، ذخیره می‌کند. این ذخیره‌سازی برای آن است که در صورت دریافت سلسله فعالیت جدید کاربران، داده‌های ذخیره شده را با داده‌ی دریافت شده مقایسه کرده و در صورت شباهت بیش از حد، مشکوک به جعلی بودن سلسله فعالیت دریافتی شود.

همچنین سکو به هستی‌شناسی کامل خانه دسترسی داشته و از موقعیت و حالات موجودیت‌ها، روابط و همبستگی‌های بین آن‌ها (ناشی از انواع کانال‌های ارتباط دهنده)، متغیرهای محیطی و احتمال سلسله فعالیت‌های مختلف کاربران برای تشخیص واقعی و یا جعلی بودن رویدادها استفاده می‌کند.

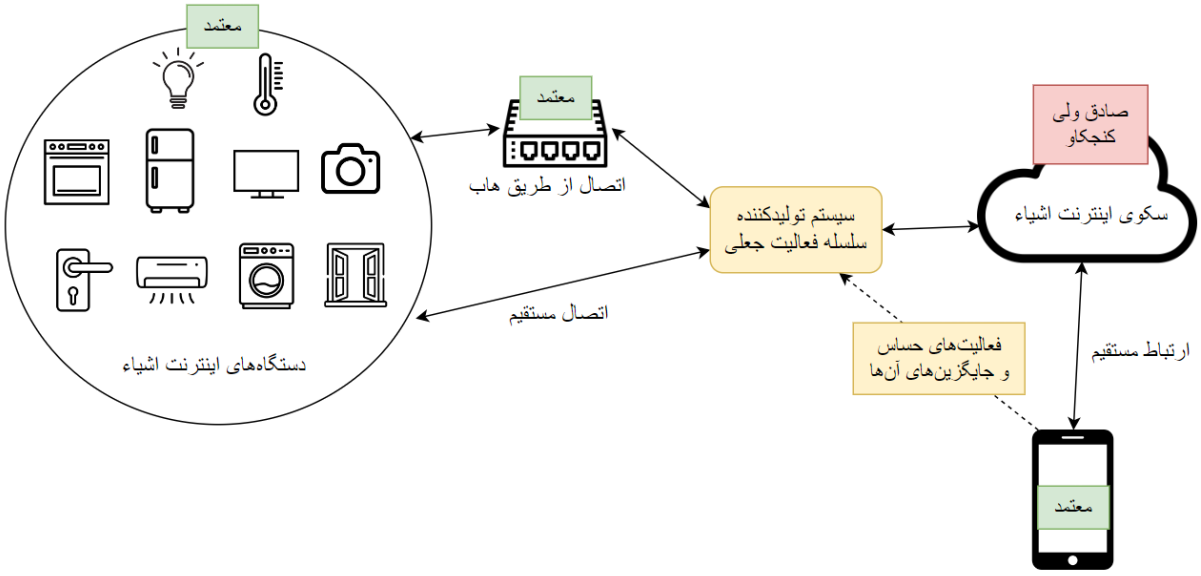
برای آن که سکو مشکوک به جعلی بودن یک سلسله فعالیت نشود، نیاز است تا تولید آن سلسله فعالیت با توجه به هستی‌شناسی خانه انجام شده و برای یکسان نبودن هر دو سلسله از فعالیت‌های جعلی، ریز فعالیت‌های انجام شده و همچنین فواصل زمانی بین انجام فعالیت‌ها، به صورت تصادفی انتخاب شده تا سکو با مقایسه‌ی سلسله فعالیت با داده‌های قبلی، متوجه شباهت زیاد بین آن‌ها نشود.

۴-۳ راهکار پیشنهادی

راهکار ارائه شده در این پژوهش، به حفظ حریم خصوصی کاربران در برابر سکوی نامعتمد بر اساس ترجیحات حریم خصوصی کاربر می‌پردازد. در ادامه به بررسی معماری این راهکار خواهیم پرداخت.

۴-۳-۱ معماری کلان

ابزار مبتنی بر راهکار پیشنهادی این پژوهش مطابق شکل ۴-۱، بین هاب و سکوی اینترنت اشیاء قرار دارد که هر کاربر، فعالیت‌های حساس خود و همچنین فعالیت‌های جعلی جایگزین را اعلام کرده است تا در صورت انجام آن فعالیت‌های حساس، سلسله فعالیت جعلی مبتنی بر نیاز کاربر تولید شود. هر فعالیت موجود در سلسله فعالیت جعلی تولید شده توسط ابزار مبتنی بر راهکار پیشنهادی، در زمان مشخص، رویداد مورد نیاز را از جانب حسگر مربوطه به سکو ارسال می‌کند.



شکل ۴-۱: محل استقرار سیستم تولیدکننده سلسله فعالیت جعلی

از طرفی، زمانی که سکو دستور کنش‌های مورد نیاز برای رویدادهای دریافتی را ارسال می‌کند؛ ابزار مبتنی بر راهکار پیشنهادی، تمامی کنش‌ها را بررسی کرده و هر کنش که در جواب رویدادی با برچسب جعلی آمده باشد را کنار می‌گذارد تا کارایی خانه هوشمند کاهش پیدا نکند.

کاربر برای تولید سلسله فعالیت جعلی دلخواهش، یکی از دو نوع ورودی مبتنی بر فعالیت و مبتنی بر نتیجه را به عنوان ورودی به ابزار مبتنی بر راهکار پیشنهادی می‌دهد. حال ابزار مبتنی بر راهکار پیشنهادی، یک سلسله فعالیت جعلی تولید می‌کند که فعالیت جعلی مدنظر کاربر در آن وجود داشته باشد.

از طرفی اگر درخواست کاربر، تولید سلسله فعالیت جعلی بر اساس نتیجه باشد، ابزار مبتنی بر راهکار پیشنهادی در ابتدا تمامی فعالیت‌های منتهی به آن نتیجه را پیدا کرده و سپس بر اساس یکی از آن فعالیت‌ها یک سلسله فعالیت جعلی تولید می‌کند که آن فعالیت را شامل باشد و نتیجه‌ی مدنظر کاربر از فعالیت مذکور گرفته شود. به عنوان مثال زمانی که ترجیح کاربر تولید سلسله فعالیت جعلی مبتنی بر نتیجه‌ی «افزایش نور محیط» است؛ نرم‌افزار ابتدا فعالیت‌های «باز کردن پنجره» (به شرط روز بودن) و «روشن کردن تلویزیون» پیدا کرده، سپس، به صورت تصادفی، یکی از آن‌ها را انتخاب کرده و بر اساس آن فعالیت، یک سلسله

فعالیت جعلی شامل فعالیت انتخاب شده تولید می‌کند.

خروجی هر بار اجرای ابزار مبتنی بر راهکار پیشنهادی، یک سلسله فعالیت جعلی بوده که هر یک در زمان خاصی باید از جانب حسگر مربوطه به سکو ارسال شوند. به طور مثال یک سلسله فعالیت جعلی شامل «باز کردن درب پذیرایی در زمان ۱۰:۴۰:۰۱:۰۰» و «روشن کردن کولر در زمان ۱۰:۴۰:۰۲:۵۰» است که بدین معناست که در زمان ۱۰:۴۰:۰۱:۰۰، کاربر درب پذیرایی را باز کرده و در زمان ۱۰:۴۰:۰۲:۵۰، کولر را روشن می‌کند. ابزار مبتنی بر راهکار پیشنهادی برای این سلسله فعالیت جعلی، در زمان ۱۰:۴۰:۰۱:۰۰، رویداد باز شدن درب پذیرایی را از جانب حسگر درب پذیرایی به سکو ارسال می‌کند؛ سپس رویداد روشن شدن کولر را در زمان ۱۰:۴۰:۰۲:۵۰، یعنی یک و نیم ثانیه پس از باز شدن درب پذیرایی، از جانب کولر به سکو ارسال می‌کند.

۴-۴ مدل‌سازی

برای مدل‌سازی جامع و کامل در برنامه‌ی این پژوهش، نیاز به هستی‌شناسی مبتنی بر زمینه به صورت خاص‌دامنه برای خانه هوشمند داریم و سپس با استفاده از این هستی‌شناسی اقدام به تولید سلسله فعالیت جعلی می‌نماییم.

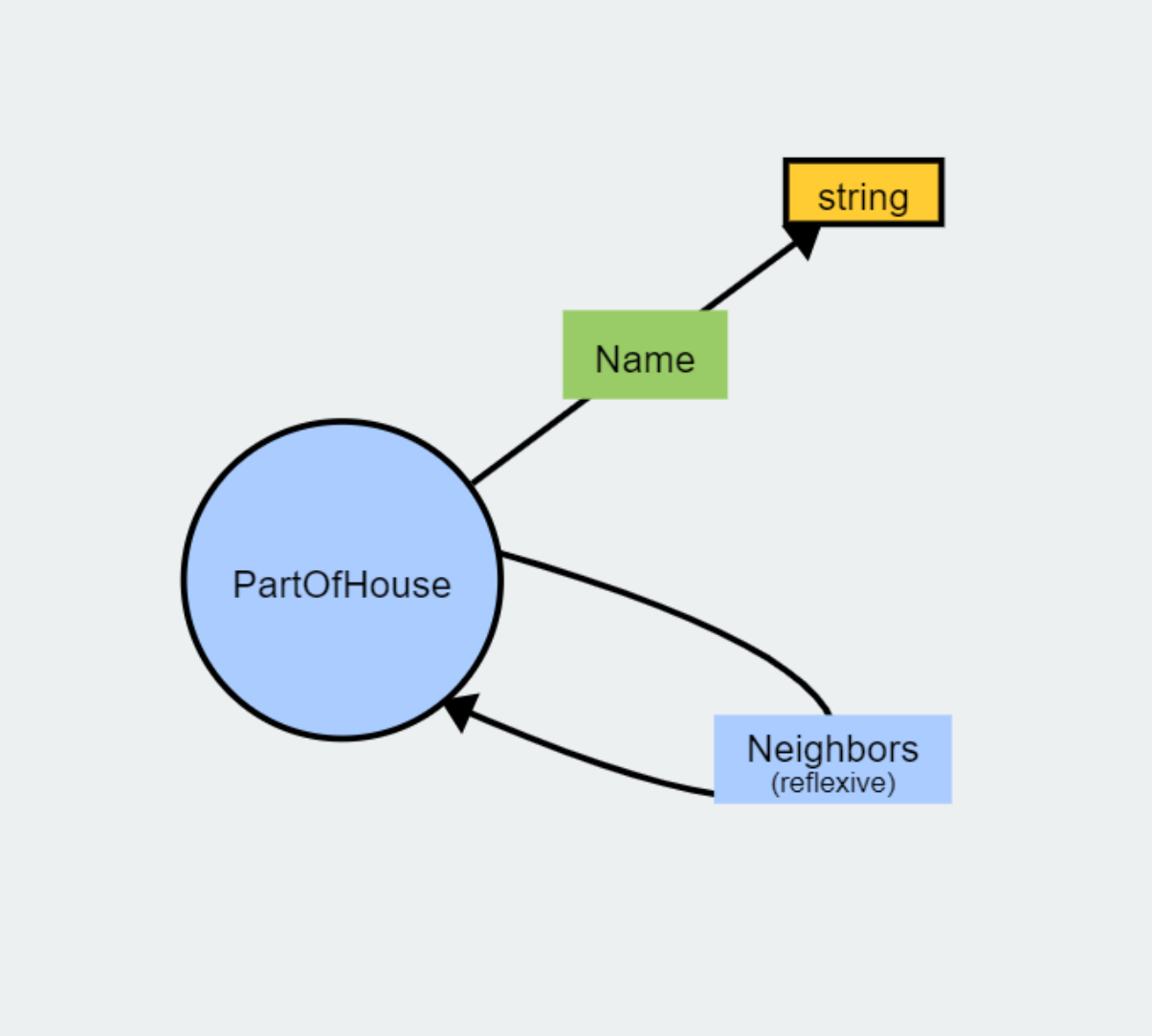
۴-۴-۱ هستی‌شناسی

در این پژوهش، هستی‌شناسی که با دانش فرد خبره به صورت دانش‌محور داده شده است؛ در چند بخش مختلف قرار می‌گیرد که در ادامه به توضیحات هر بخش خواهیم پرداخت.

- خانه هوشمند: همانطور که در شکل ۴-۲ قابل مشاهده است، نقشه‌ی خانه هوشمند توسط فرد خبره تعریف شده است. هر بخش شامل نام و لیست بخش‌هایی از خانه هوشمند است که به طور مستقیم به بخش مورد نظر متصل هستند (به عنوان مثال اتاق خواب، به طور مستقیم به راهرو و بالکن متصل است). استفاده از نقشه‌ی خانه در این پژوهش طوری تعریف شده است که با تغییر نقشه‌ی خانه، نیازی به تغییر دیگری در برنامه نخواهیم داشت و عملکرد برنامه تحت تاثیر قرار نخواهد گرفت.

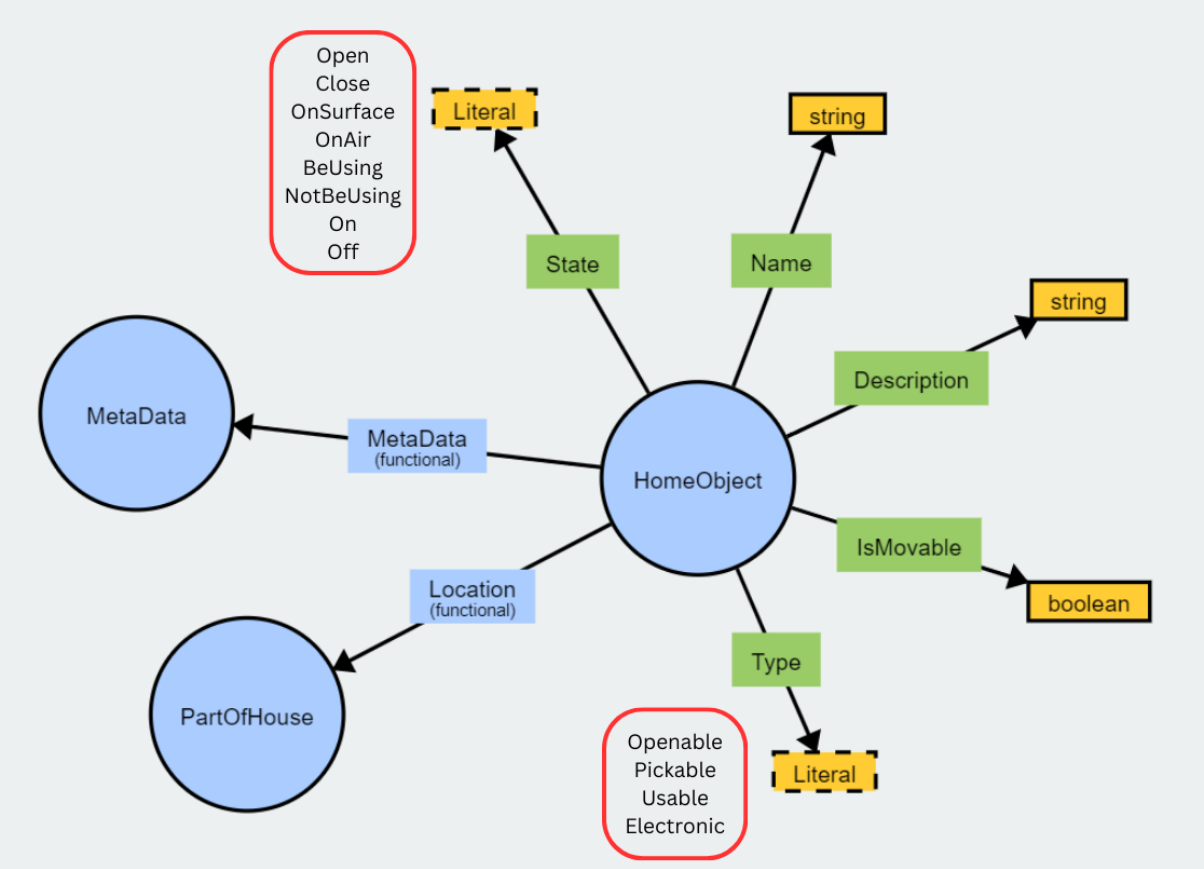
  هستی‌شناسی تعریف شده برای خانه هوشمند، در شکل ۴-۳ قابل مشاهده است. در این هستی‌شناسی هر بخش خانه هوشمند مانند آشپزخانه، راهرو، پذیرایی، بالکن، دستشویی و اتاق‌خواب‌ها و همچنین همسایگان هر کدام، توسط فرد خبره مشخص شده اند.

شکل ۴-۲: نمونه نقشه خانه هوشمند



شکل ۴-۳: هستی‌شناسی بخش‌های خانه هوشمند

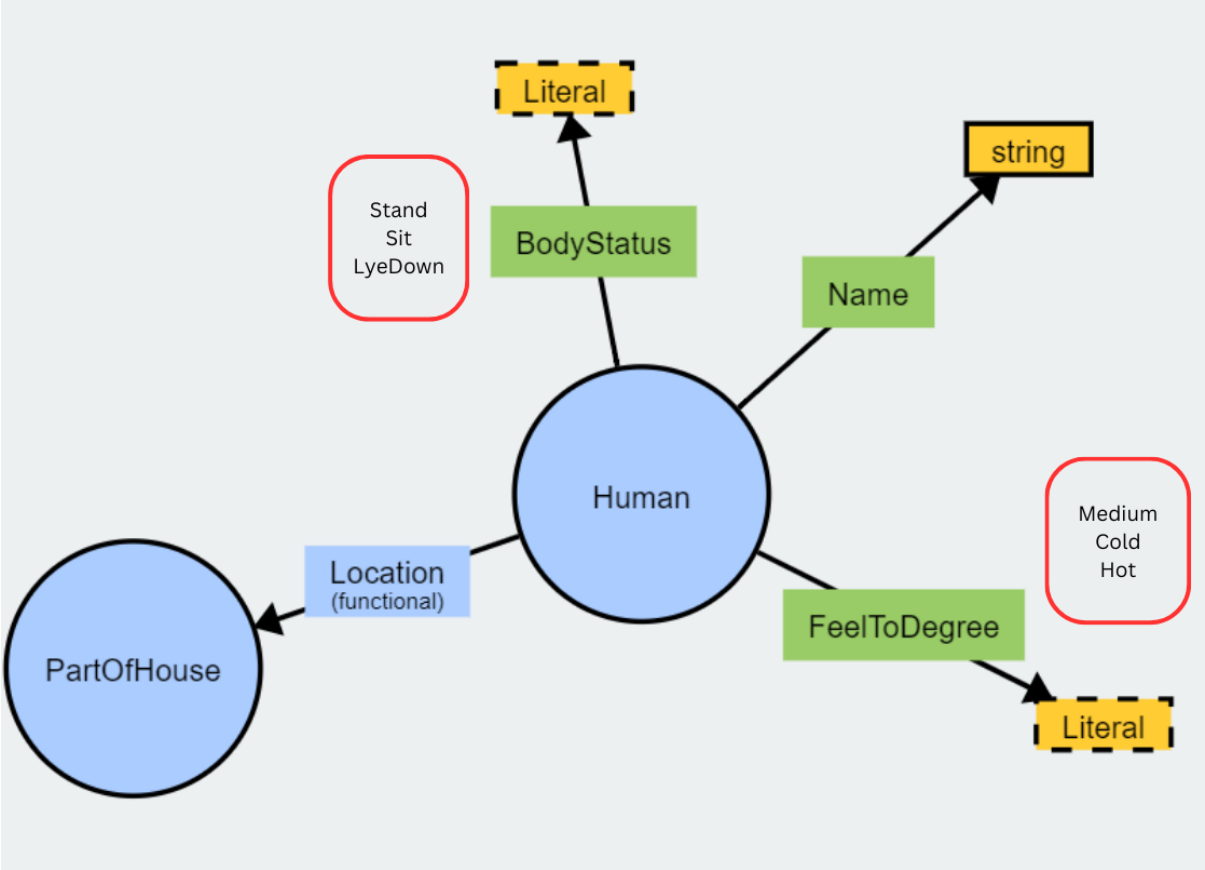
- موجودیت‌ها: در هستی‌شناسی که توسط فرد خبره به برنامه داده می‌شود، موقعیت و حالت اولیه برای تمامی موجودیت‌ها (انسان‌ها و اشیاء که حسگر دارند) فراهم می‌شود. در صورتی که در سلسله فعالیت‌های جعلی، موقعیت یا حالت یک موجودیت تغییر کند، این تغییر در پایگاه داده ذخیره می‌شود تا سلسله فعالیت‌های بعدی بر اساس اطلاعات جدید موجودیت‌ها باشد تا از دید سکوی نامعتمد، تناقضی در داده‌ها رخ ندهد اما کنش‌های مربوط به تمامی رویدادهایی که با برچسب جعلی به سکو ارسال شده‌اند، کنار گذاشته می‌شوند تا بهره‌وری خانه هوشمند کاهش پیدا نکند. هستی‌شناسی اشیاء خانه هوشمند در شکل ۴-۴ و هستی‌شناسی افراد حاضر در خانه در شکل ۴-۵ قابل مشاهده است.



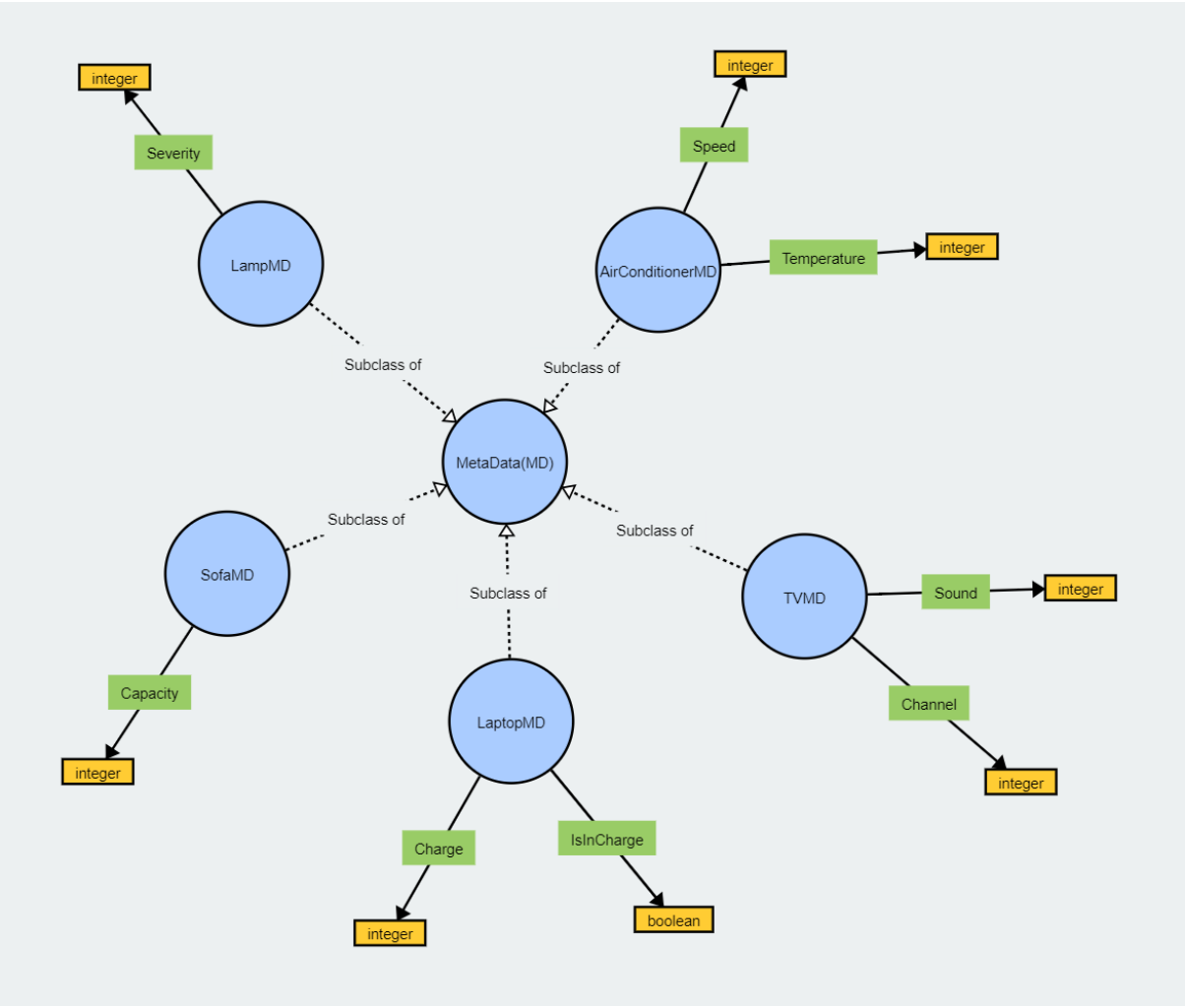
شکل ۴-۴: هستی‌شناسی اشیاء خانه هوشمند

همچنین هستی‌شناسی فرادادههای اشیاء خانه هوشمند در شکل ۴-۶ قابل مشاهده است. هر یک از اشیاء خانه هوشمند حداکثر یکی از این فرادادهها را در هستی‌شناسی خود دارند تا با استفاده از آن‌ها شروط هر یک از اشیاء به طور دقیق‌تر بررسی شود و تغییر حالت آن به طور دقیق‌تر اعمال شود.

- شرایط محیطی: برای دخیل کردن شرایط محیطی مانند میزان نور، دما، صدا و ... در تولید سلسله فعالیت‌های جعلی، مقدار اولیه برای اجزای محیط در دانش اولیه‌ی برنامه قرار دارد. توجه شود که تاریخ و ساعت به عنوان یک متغیر محیطی در تولید سلسله فعالیت جعلی دخیل بوده اما در هستی‌شناسی اولیه قرار ندارد چرا که متغیر زمان به صورت خودکار مقدار می‌گیرد. استفاده از این

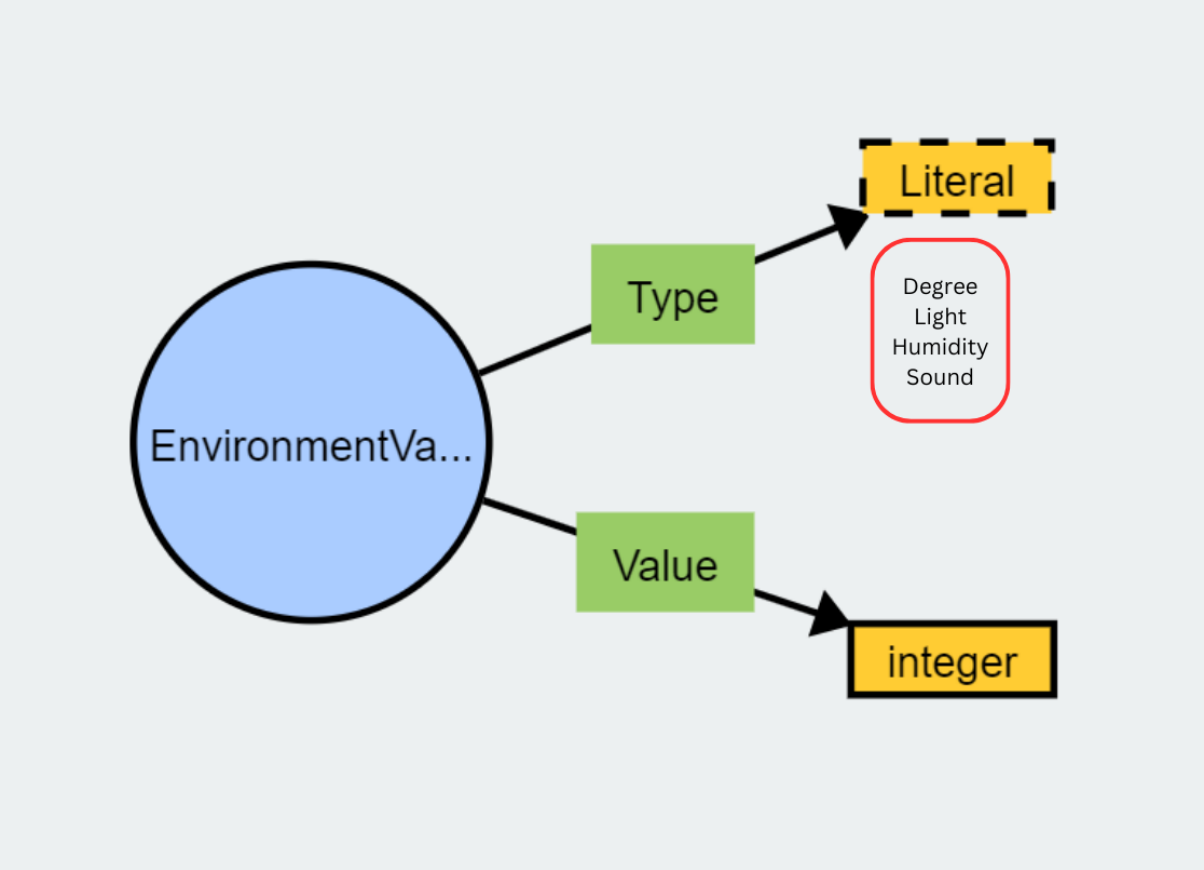


شکل ۴-۵: هستی‌شناسی افراد حاضر در خانه هوشمند



شکل ۴-۶: هستی‌شناسی فراداده‌های اشیاء خانه هوشمند

اجزای محیط کمک به واقعی به نظر رسیدن سلسله فعالیت‌های جعلی خواهد کرد. به عنوان مثال زمانی که پنجره باز می‌شود، حسگر سنجش نور محیط، در ساعاتی که خورشید در آسمان است، رویداد افزایش نور محیط را ارسال می‌کند. پس زمانی که به صورت جعلی ارسال رویداد باز شدن پنجره را به حسگر مربوطه ارسال می‌کنیم، برنامه با توجه به ساعت شبانه روز تصمیم به ارسال یا عدم ارسال رویداد افزایش نور محیط توسط حسگر سنجش نور می‌گیرد. هستی‌شناسی مربوط به اجزای محیط در شکل ۴-۷ قابل مشاهده است.

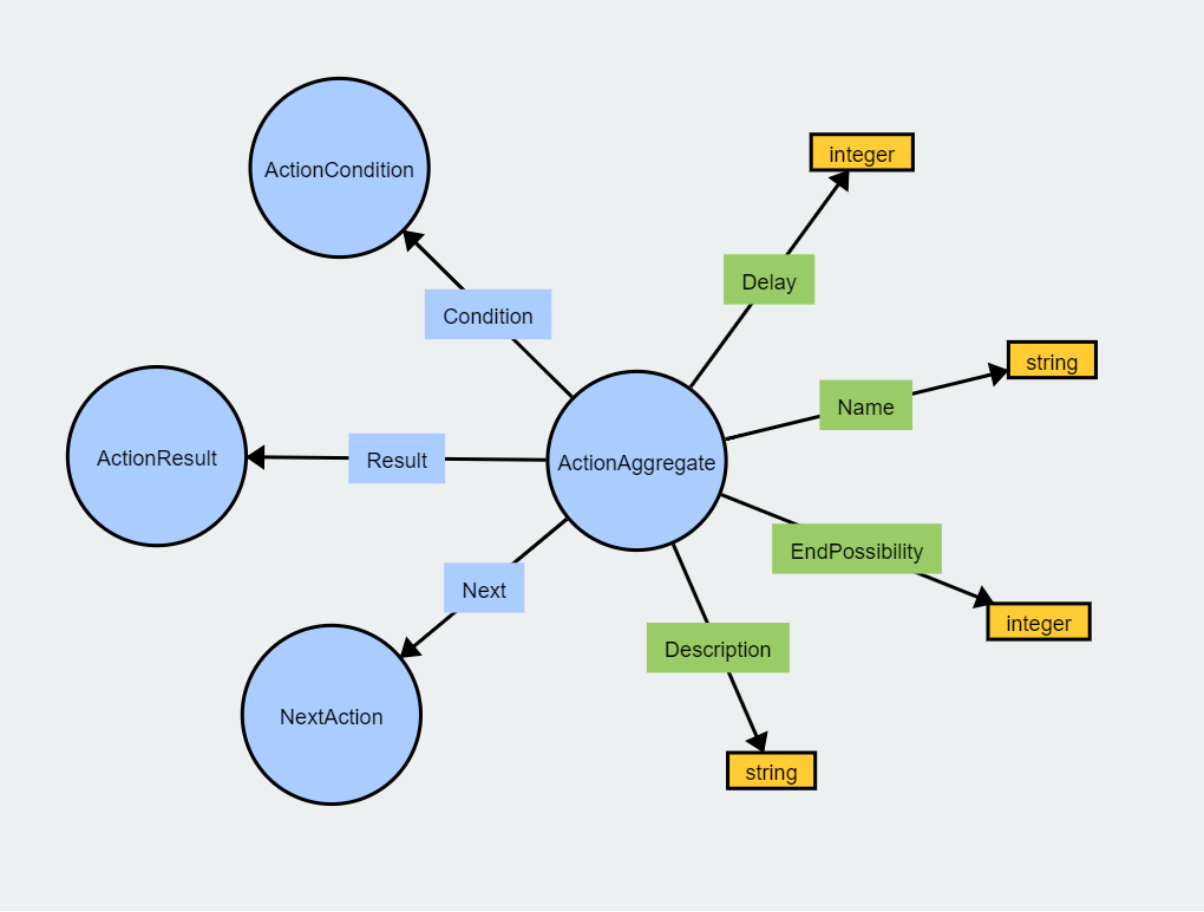


شکل ۴-۷: هستی‌شناسی اجزای محیط خانه هوشمند

- فعالیت: هر فعالیت کاربر به صورت سابقه‌ی مختص به آن کاربر در هستی‌شناسی اولیه توسط فرد خبره داده می‌شود که فعالیت‌ها شامل شروط مورد نیاز برای انجام به صورت جعلی (به عنوان مثال حضور در پذیرایی برای «روشن کردن تلویزیون»)، نتایج اعمال شده پس از ارسال رویداد جعلی یک فعالیت به سکو (به عنوان مثال کاهش دمای محیط پس از «روشن کردن کولر») و فعالیت‌های احتمالی انجام شده توسط کاربر پس از انجام یک فعالیت مشخص (به عنوان مثال پس از «ورود به پذیرایی»، «روشن کردن کولر» با احتمال ۷۰% و «نشستن روی مبل» با احتمال ۳۰% انجام می‌شود) است که نحوه مدل‌سازی احتمال توالی فعالیت‌ها را در بخش ۴-۴-۲ توضیح خواهیم داد. توجه شود که هر فعالیت با احتمالی مشخص با توجه به سوابق کاربر، می‌تواند آخرین فعالیت یک

سلسله فعالیت جعلی باشد و لزوما نیاز به ادامه در تولید سلسله فعالیت جعلی با توجه به فعالیت احتمالی بعدی نداریم. هستی‌شناسی کلی فعالیت‌ها در شکل ۴-۸ قابل مشاهده است.

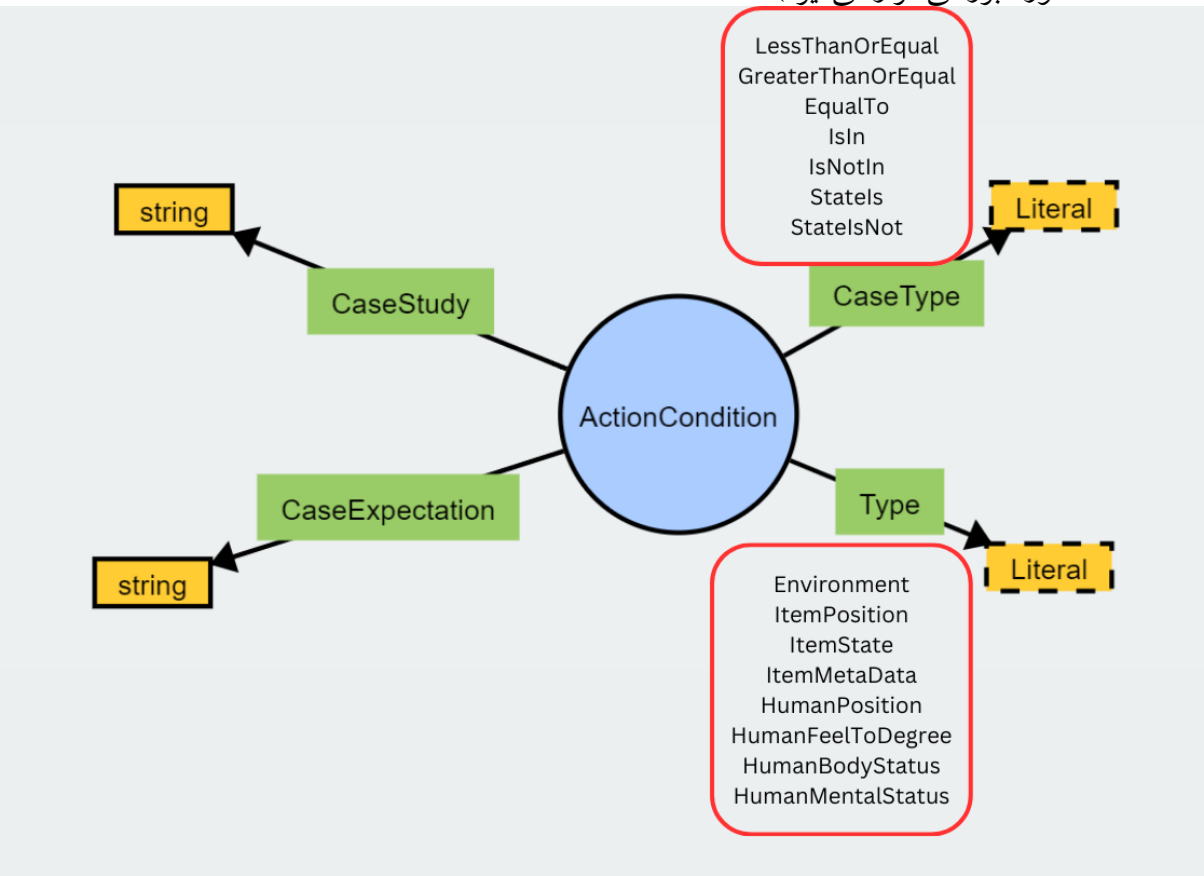
به عنوان مثال در فعالیت «روشن کردن تلویزیون» به عنوان ActionAggregate، حضور در پذیرایی یکی از شروط انجام آن (ActionCondition)، تغییر وضعیت تلویزیون به حالت روشن یکی از نتایج انجام آن (ActionResult)، افزایش یا کاهش صدای تلویزیون یکی از فعالیت‌های احتمالی بعدی (NextAction)، "روشن کردن تلیویزیون" نام آن، تاخیر برابر ۴ ثانیه، احتمال ۵۰ درصد بودن فعالیت نهایی و "روشن کردن تلویزیون به صورت دستی توسط افراد حاضر در خانه" به عنوان توضیحات بر اساس هستی‌شناسی این فعالیت داده شده است.



شکل ۴-۸: هستی‌شناسی کلی فعالیت‌های خانه هوشمند

در هستی‌شناسی شرایط انجام فعالیت که در شکل ۴-۹ قابل مشاهده است، نوع موجودیتی که شرط روی آن بررسی می‌شود(اجزای محیط، موقعیت مکانی شیء، حالت شیء، فراداده شیء، موقعیت مکانی فرد حاضر در خانه، احساس به شرایط آب و هوایی فرد حاضر در خانه و حالت بدن فرد حاضر در خانه)، نام آن موجودیت مانند تلویزیون، فرد حاضر در خانه و درب بالکن، نوع شرط مورد بررسی با توجه به موجودیت و شرط مورد بررسی مانند کمتر یا بیشتر، داخل یا خارج و برابر یا مخالف یک مقدار مشخص ذکر شده است. در هر شرط انجام فعالیت، آخرین مقادیر دریافتی از حسگرها

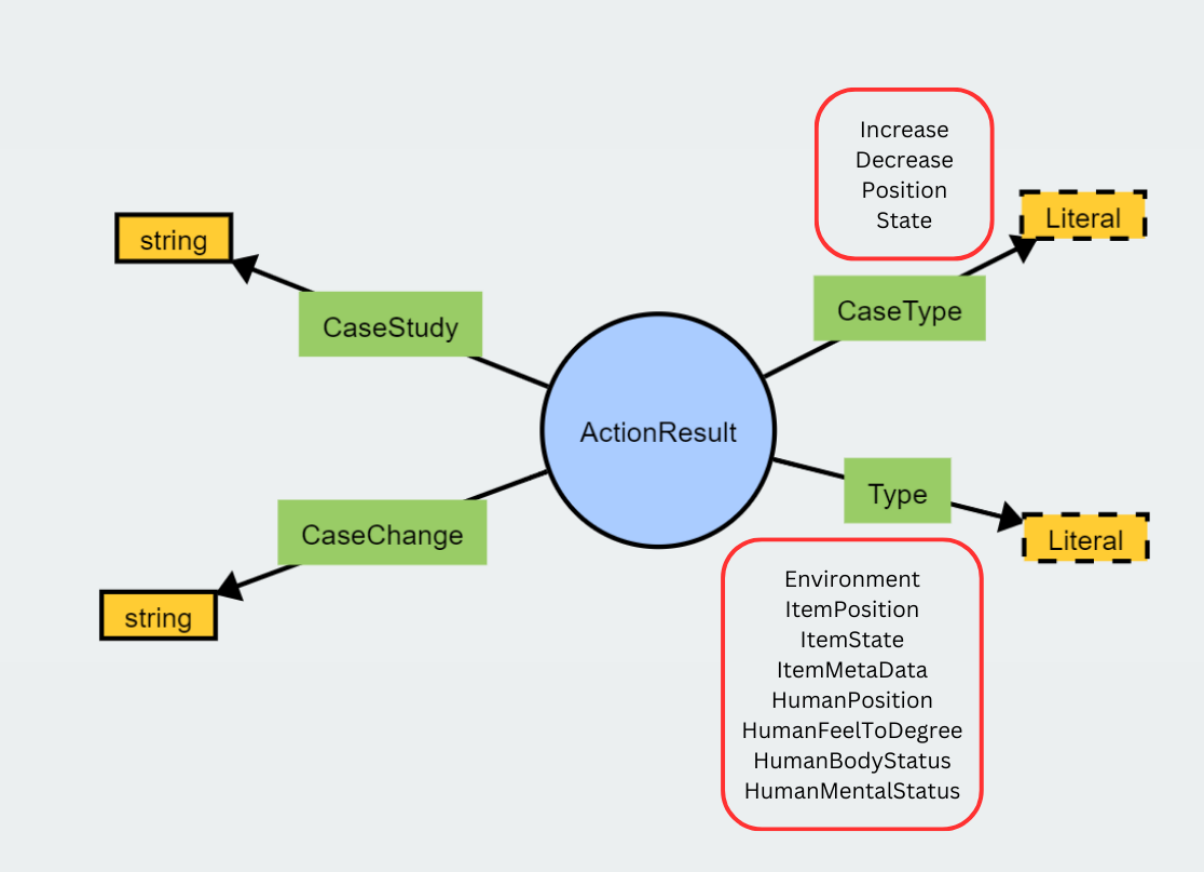
بررسی شده و بر اساس آن‌ها تصمیم گرفته می‌شود (به عنوان مثال اگر شرط انجام یک فعالیت این باشد که دمای خانه هوشمند کمتر از بیست درجه باشد، آخرین مقدار اندازه‌گیری شده توسط حسگر دما، مورد بررسی قرار می‌گیرد).



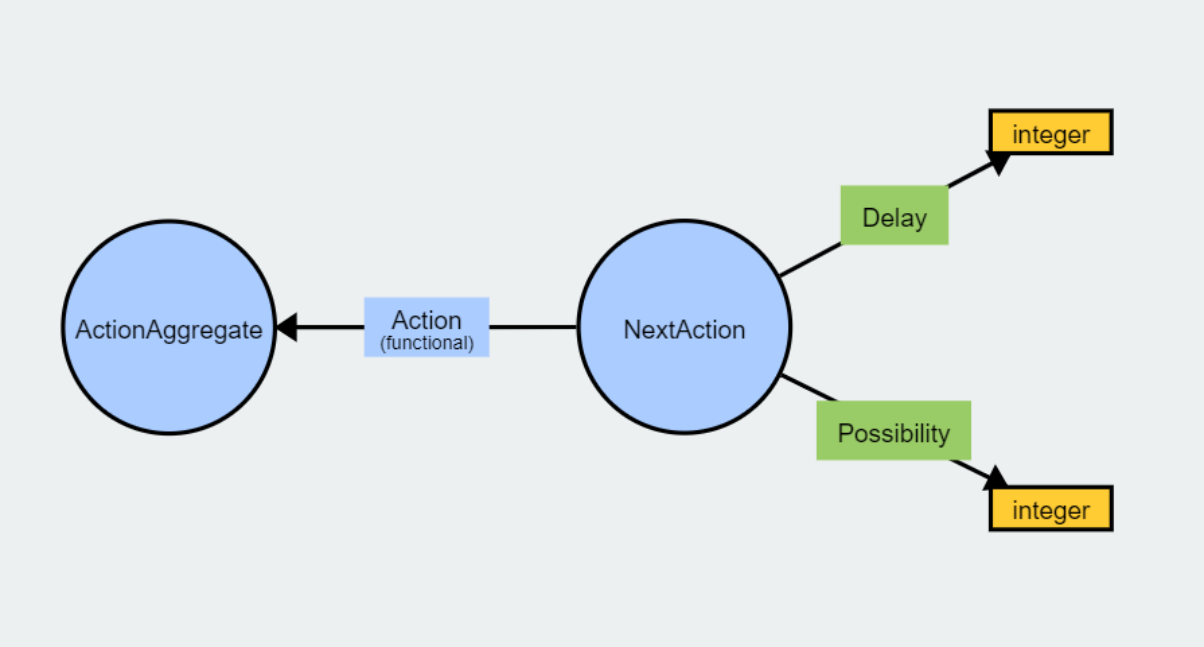
شکل ۴-۹: هستی‌شناسی شرایط فعالیت‌های خانه هوشمند

در هستی‌شناسی نتایج انجام فعالیت که در شکل ۴-۱۰ قابل مشاهده است، نوع موجودیت برای نتیجه اعمال شده، نام موجودیت، نوع نتیجه اعمال شده با توجه به موجودیت و نتیجه اعمال شده مانند افزایش یا کاهش، تغییر موقعیت مکانی و تغییر حالت به یک مقدار مشخص ذکر شده است. در هر نتیجه انجام فعالیت، آخرین مقادیر دریافتی از حسگرها تغییر می‌کند (به عنوان مثال اگر نتیجه انجام یک فعالیت این باشد که دمای خانه هوشمند پنج درجه افزایش پیدا کند، آخرین مقدار اندازه‌گیری شده توسط حسگر دما، تغییر خواهد کرد).

در هستی‌شناسی فعالیت‌های احتمالی بعدی که در شکل ۴-۱۱ قابل مشاهده است، تاخیر انجام فعالیت، احتمال انجام و خود فعالیت مشخص شده است.



شکل ۴-۱۰: هستی‌شناسی نتایج فعالیت‌های خانه هوشمند

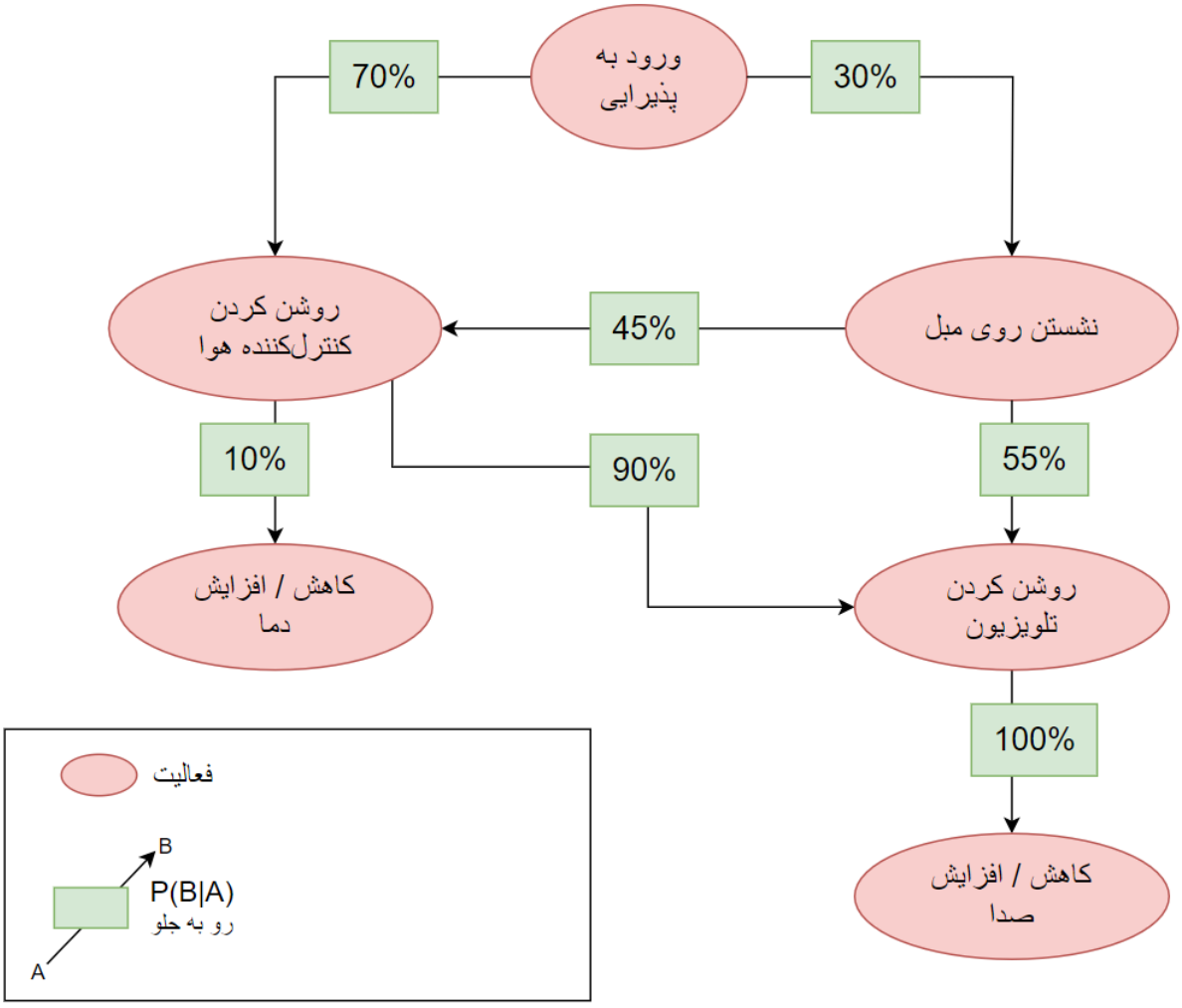


شکل ۴-۱۱: هستی‌شناسی فعالیت‌های احتمالی بعدی خانه هوشمند

۴-۴-۲ مدل‌سازی فعالیت کاربر با قوانین انجمنی

در این بخش به مدل‌سازی توالی فعالیت کاربر با استفاده از قوانین انجمنی خواهیم پرداخت و حرکت رو به جلو و رو به عقب سلسله فعالیت‌های کاربر را بررسی خواهیم کرد.

همانطور که در هستی‌شناسی شکل ۴-۱۱ مشاهده شد؛ هر فعالیت شامل لیستی از فعالیت‌های احتمالی بعدی است که احتمال انجام هر یک با توجه به سوابق کاربر در هستی‌شناسی فعالیت‌ها ثبت شده است. با توجه به قوانین انجمنی مطرح شده در بخش ۴-۲، احتمال انجام فعالیت Y پس از فعالیت X همان اطمینان است که بیان می‌کند در چند درصد مواقع پس از انجام فعالیت X، فعالیت Y انجام می‌شود. مثالی از نحوه مدل‌سازی احتمال توالی فعالیت‌ها با استفاده از قوانین انجمنی در شکل ۴-۱۲ قابل مشاهده است (احتمال توالی هر فعالیت پس از فعالیت دیگر به رنگ سبز نشان داده شده است).



شکل ۴-۱۲: مثالی از مدل‌سازی احتمال توالی فعالیت‌ها با استفاده از قوانین انجمنی

حال با توجه به مقدار اطمینان و با استفاده از تعریف پشتیبانی برای انجام فعالیت Y پس از فعالیت X، احتمال انجام فعالیت X قبل از فعالیت Y با فرض انجام Y را تعریف می‌کنیم. این تعریف برای حرکت

عقبگرد بین فعالیت‌ها می‌باشد تا بتوانیم انتخاب فعالیت قبل از یک فعالیت مشخص از بین فعالیت‌های ممکن را با دخیل کردن احتمال انجام هر یک انجام دهیم. این احتمال با استفاده از پشتیبانی انجام فعالیت Y پس از فعالیت X محاسبه می‌شود.

برای انجام محاسبات مربوط به احتمال انتخاب هر فعالیت در حرکت عقبگرد، در ابتدا برای هر فعالیت احتمال رخداد آن به صورت کلی را طبق فرمول زیر محاسبه می‌کنیم:

$$P(X) = \sum_{\text{Parents}} P(\text{Parent}) \cdot P(\text{Parent} \rightarrow X) \qquad (۱-۴)$$

در این فرمول، فعالیت‌هایی که به عنوان فعالیت بعدی در هستی‌شناسی هیچ فعالیت دیگری تعریف نشده‌اند، احتمال یک را دارند. برای دیگر فعالیت‌ها مانند فعالیت X، تمامی فعالیت‌هایی که پس از آن‌ها فعالیت X احتمال انجام دارد را پیدا می‌کنیم (لیست آن فعالیت‌ها را Parents می‌نامیم). سپس احتمال رخداد هر فعالیت Parent که طبق همین فرمول حساب شده را در احتمال رخداد فعالیت X به شرط رخداد فعالیت Parent ضرب کرده و در آخر مقدار به دست آمده به ازای هر فعالیت Parent را با هم جمع می‌کنیم. در ابتدای هر اجرای ابزار پیشنهادی این پژوهش، احتمال رخداد هر فعالیت به صورت کلی را محاسبه و ذخیره می‌کنیم. احتمال رخداد کلی هر فعالیت در مثال شکل ۴-۱۲، در شکل ۴-۱۳ قابل مشاهده است (احتمال کلی رخداد هر فعالیت به رنگ بنفش نشان داده شده است).

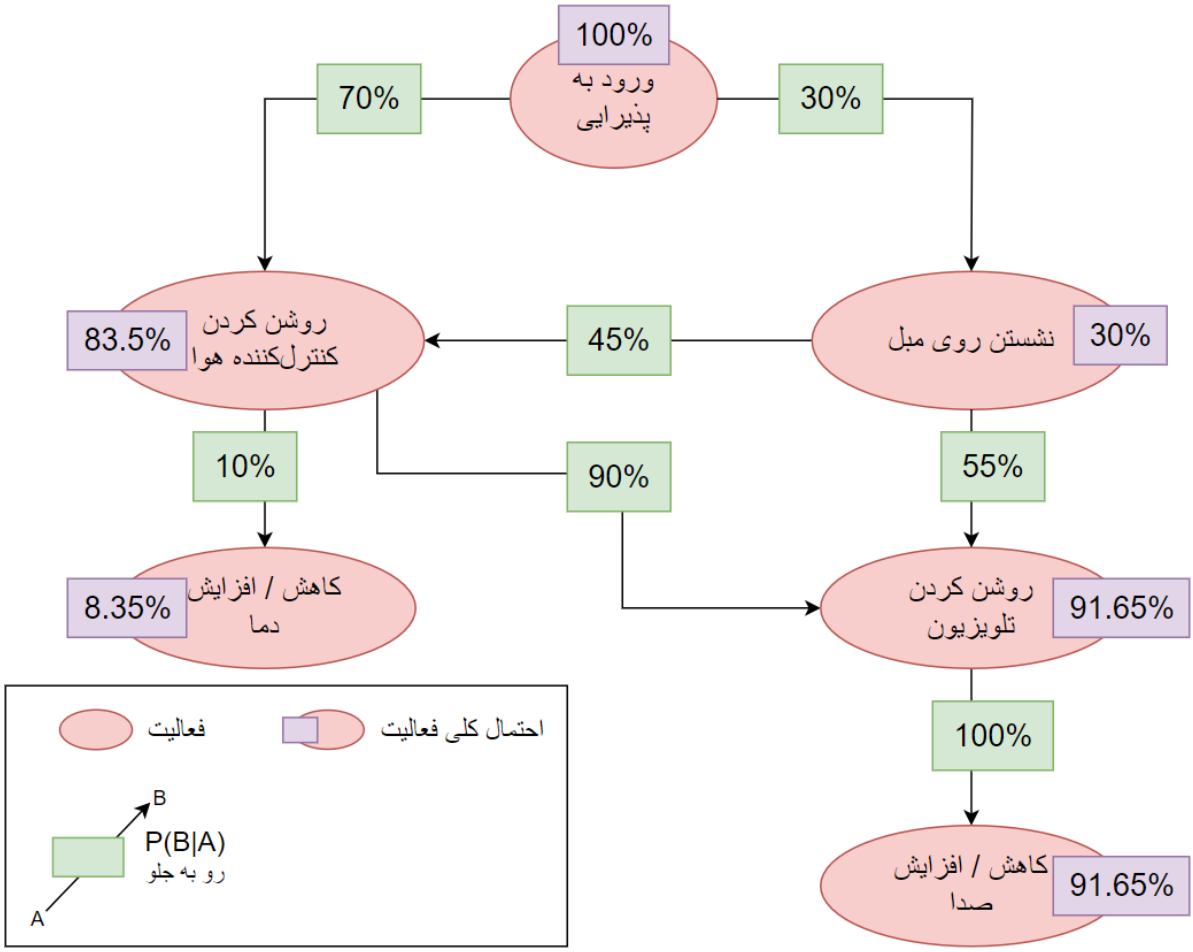
حال برای احتمال عقبگرد، از فرمول احتمالات شرطی استفاده می‌کنیم:

$$P(Y|X) = \frac{P(X|Y) \cdot P(Y)}{P(X)} \qquad (۲-۴)$$

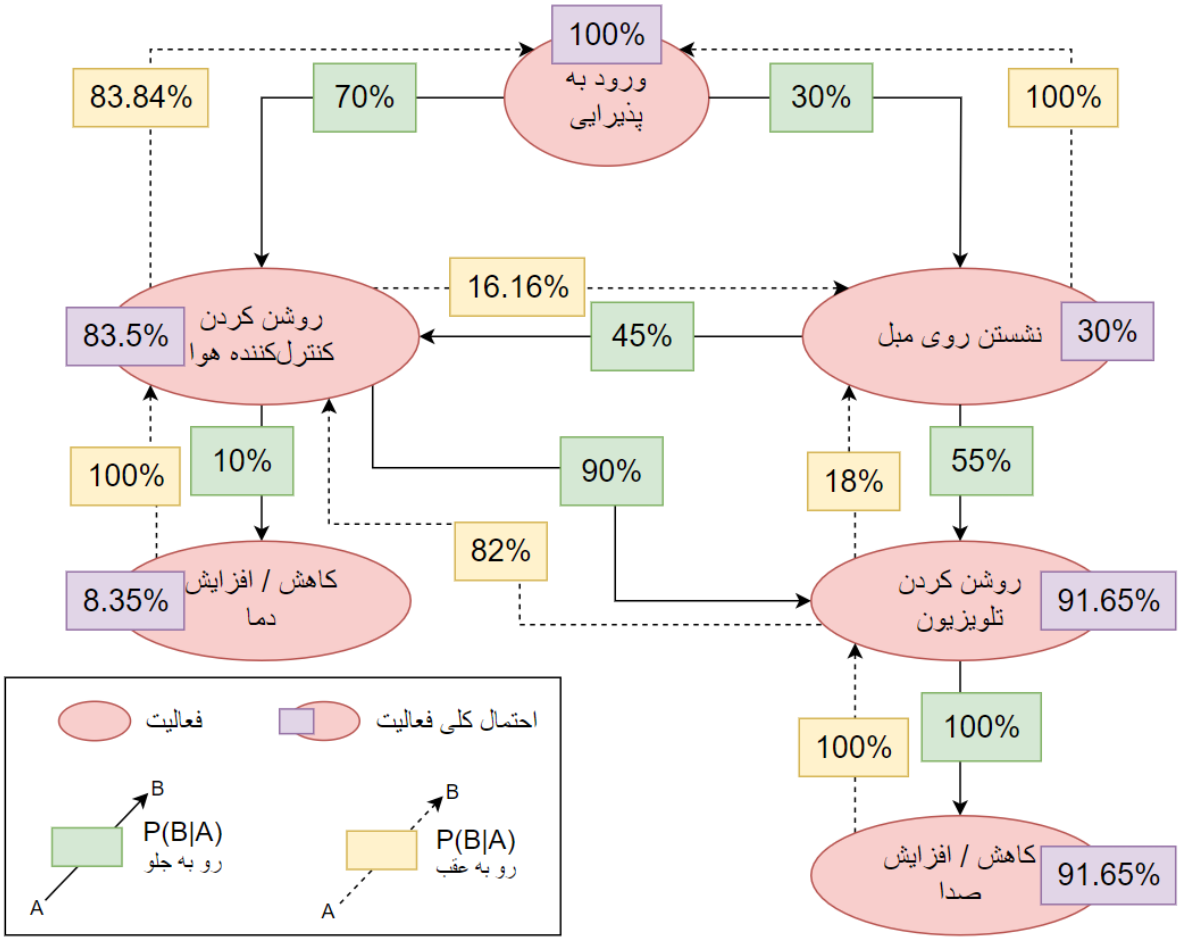
در این فرمول با استفاده از احتمال رخداد هر فعالیت مانند فعالیت X پس از هر فعالیت مانند Y و همچنین احتمال رخداد کلی فعالیت‌های X و Y که محاسبه شده است، احتمال رخداد فعالیت Y به صورت حرکت عقبگرد از فعالیت X قابل محاسبه است. احتمال انتخاب فعالیت‌ها در حرکت عقبگرد در مثال شکل ۴-۱۲، در شکل ۴-۱۴ قابل مشاهده است (احتمال انتخاب هر فعالیت برای حرکت عقبگرد از فعالیتی دیگر به رنگ زرد نشان داده شده است).

۴-۴-۳ الگوریتم تولید سلسله فعالیت جعلی

ابزار مبتنی بر راهکار پیشنهادی برای تولید سلسله فعالیت جعلی از الگوریتم نمایش داده شده در شکل ۴-۱۵ استفاده می‌کند. به طور کلی این فرایند به چند فاز مختلف تقسیم می‌شود که در ادامه به بررسی هر

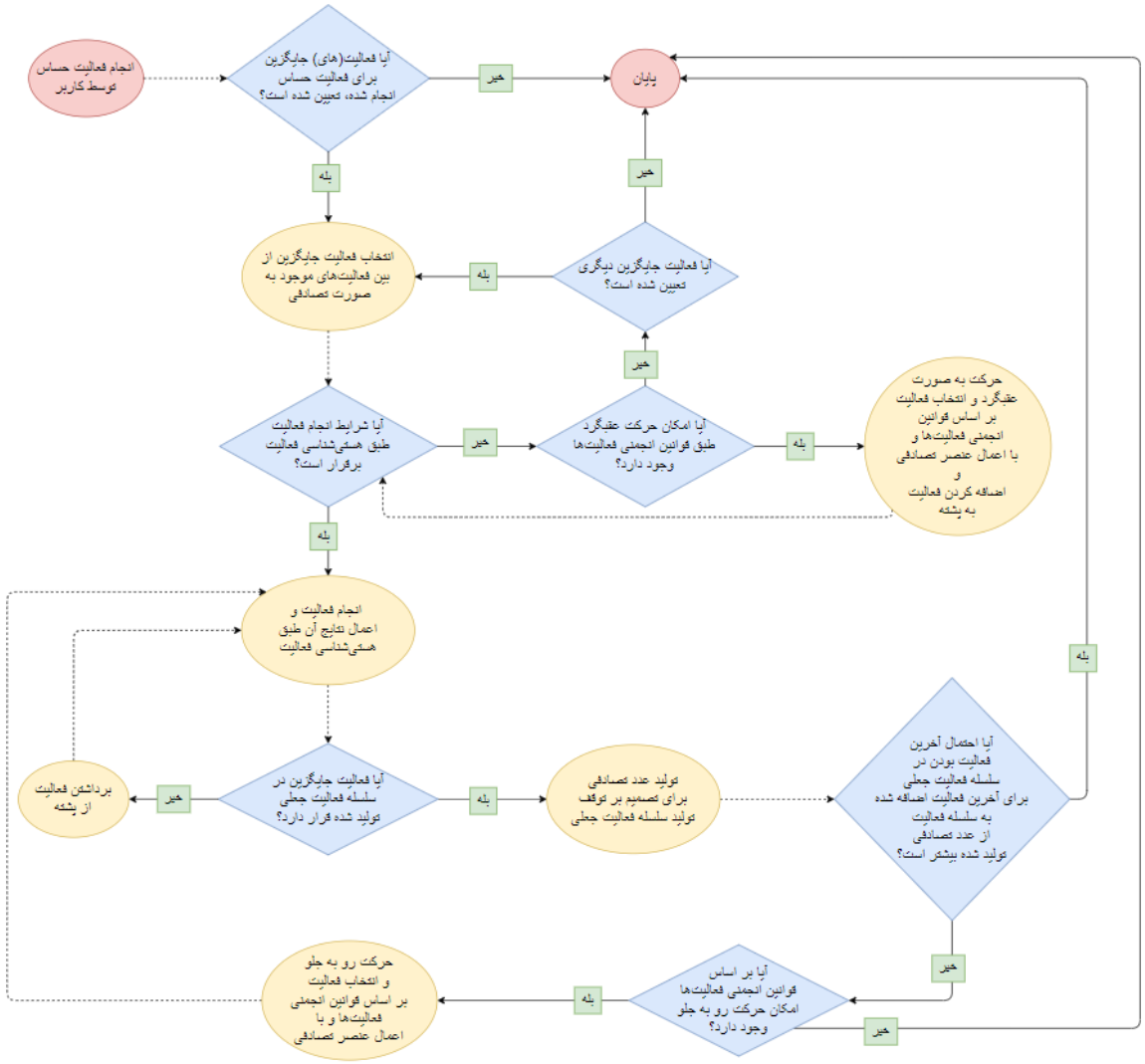


شکل ۴-۱۳: احتمال کلی فعالیت‌ها در مثال شکل ۴-۱۲



شکل ۴-۱۴: احتمال انتخاب فعالیت‌ها در حرکت عقبگرد در مثال شکل ۴-۱۲

یک خواهیم پرداخت.



شکل ۴-۱۵: الگوریتم تولید سلسله فعالیت جعلی

- انجام فعالیت حساس توسط کاربر: در اولین گام، فعالیتی حساس توسط کاربر انجام می‌شود و ابزار مبتنی بر راهکار پیشنهادی، از بین فعالیت‌های جایگزین برای این فعالیت حساس که توسط کاربر تعیین شده است، یکی را به صورت تصادفی انتخاب کرده و اقدام به تولید سلسله فعالیت جعلی می‌کند که این فعالیت جایگزین در آن حضور داشته باشد.

- حرکت عقبگرد: حال ابزار مبتنی بر راهکار پیشنهادی، با شروع از فعالیت جایگزین، حرکتی عقبگرد بر اساس قوانین انجمنی را شروع می‌کند تا جایی که به فعالیتی برسد که شرایط انجام آن طبق هستی‌شناسی فعالیت برقرار باشد. در این حرکت عقبگرد تمامی فعالیت‌ها در یک پشته ذخیره می‌شوند تا در صورت ارضا شدن شرایط یک فعالیت، آن‌ها را از پشته برداریم و به سلسله فعالیت جعلی اضافه کنیم. به طور مثال برای «روشن کردن تلویزیون»، ابتدا باید یک فرد در پذیرایی حضور

داشته باشد و برای حضور در پذیرایی باید از راهرو به آنجا وارد شد. به همین صورت حرکت عقبگرد انجام می‌دهیم تا به طور واقعی حضور فرد و ارضا شدن دیگر شرایط برای یک فعالیت را داشته باشیم؛ حال فعالیت‌های داخل پشته را برداشته و به سلسله فعالیت جعلی اضافه می‌کنیم. توجه شود که انتخاب فعالیت در حرکت عقبگرد بر اساس قوانین انجمنی و با اعمال عنصر تصادفی است که احتمال انتخاب هر فعالیت با استفاده از محاسبات مطرح شده در بخش ۴-۴-۲ انجام می‌شود.

- حرکت رو به جلو: حال که فعالیت جایگزین مورد نظر در سلسله فعالیت جعلی اضافه شده است؛ ابزار مبتنی بر راهکار پیشنهادی برای فریب سکو و یکسان نبودن سلسله فعالیت جعلی برای یک فعالیت جایگزین، حرکتی رو به جلو بر اساس قوانین انجمنی انجام می‌دهد. در ابتدا با استفاده از احتمال توقف آخرین فعالیت که در هستی‌شناسی آمده است تصمیم به ادامه تولید فعالیت جعلی یا توقف می‌گیرد؛ اگر تصمیم بر ادامه تولید فعالیت جعلی گرفته شد، با استفاده از قوانین انجمنی و با اعمال عنصر تصادفی یکی از فعالیت‌های ممکن برای کاربر با استفاده از محاسبات بخش ۴-۴-۲ را انتخاب کرده و آن را به سلسله فعالیت جعلی اضافه می‌کند.
- نقطه توقف نهایی: پس از حرکت رو به جلو جهت فریب سکو، حالت خانه هوشمند را به سمت حالت کنونی واقعی می‌بریم تا از حالات اشیاء و افراد حاضر در خانه پس از مدتی، سکو توانایی تشخیص رخداد سلسله فعالیت جعلی را نداشته باشد. برای این کار آخرین فعالیتی که تاثیر مهمی در حالات افراد حاضر در خانه و اشیاء گذاشته را از داده‌های حسگرها دریافت می‌کنیم، سپس اقدام به تولید مجدد سلسله فعالیت جعلی تا انجام آن فعالیت به عنوان فعالیت نهایی می‌کنیم. پس از تولید هر فعالیت جعلی، آخرین فعالیت با تاثیر مهم را بررسی مجدد می‌کنیم تا به صورت پویا و بدون تولید کامل سلسله فعالیت جعلی، حالت توقف را به حالت واقعی خانه هوشمند نزدیک کنیم. با استفاده از این روش همواره به حالت واقعی خانه هوشمند می‌رسیم که در بدترین حالت که تغییر سریع و مداوم حالت خانه هوشمند است، در هنگام پایان روز و زمان خوابیدن کاربران به حالت واقعی خانه هوشمند می‌رسیم.

در تولید سلسله فعالیت جعلی، جهت نزدیک بودن به واقعیت، مدت زمان میانگین انجام هر فعالیت را نیز در هستی‌شناسی لحاظ کرده‌ایم. علاوه بر آن، از آنجایی که تمامی موجودیت‌ها در خانه، هوشمند نیستند و ابزار مبتنی بر راهکار پیشنهادی از جزئیات کامل خبر ندارد؛ یک تاخیر بین فعالیت‌ها لحاظ می‌شود که زمان بین اتمام یک فعالیت و شروع فعالیت بعدی است. به طور مثال پس از «ورود به خانه»، «روشن کردن کولر» به طور میانگین پس از دو ثانیه انجام می‌شود که این زمان، حرکت کاربر از درب خانه به سمت کولر و روشن کردن آن است.

با استفاده از مدت زمان انجام هر فعالیت و فاصله زمانی آن تا فعالیت بعدی، هر فعالیتی که به سلسله فعالیت جعلی اضافه شود؛ یک زمان انجام فعالیت معینی دارد که آن زمان، زمان انجام فعالیت قبلی علاوه بر مدت زمان انجام آن و تاخیر بین فعالیت قبلی و فعالیت کنونی است. به طور مثال اگر «ورود به خانه» به مدت یک ثانیه طول می‌کشد و همچنین میانگین تاخیر بین باز کردن درب خانه و روشن کردن کولر هشت ثانیه باشد؛ زمان انجام فعالیت «روشن کردن کولر» ۱ + ۸ = ۹ ثانیه پس از «ورود به خانه» خواهد بود.

۴-۴-۴ عوامل تصادفی‌ساز

از آنجا که سکو قادر به ذخیره و مقایسه‌ی سلسله فعالیت‌ها و همچنین کشف شباهت بین آن‌هاست؛ در تولید سلسله فعالیت جعلی از عوامل تصادفی‌ساز استفاده می‌کنیم تا تنوع در تولید خروجی لحاظ شود. در این بخش به بررسی این عوامل خواهیم پرداخت.

- انتخاب فعالیت جایگزین: زمانی که فعالیتی حساس انجام می‌شود، فعالیت جایگزین از بین فعالیت‌های جایگزینی که ترجیح کاربر است به طور تصادفی انتخاب می‌شود.
- انتخاب فعالیت در حرکت عقبگرد: همانطور که اشاره شد، زمانی که به صورت عقبگرد حرکت می‌کنیم، احتمال انتخاب هر فعالیت به صورت تصادفی بوده اما احتمال محاسبه شده در قوانین انجمنی برای حرکت عقبگرد، به عنوان وزن در این انتخاب لحاظ می‌شود. برای این کار برای هر فعالیت سهمی برابر احتمال محاسبه شده در نظر می‌گیریم، سپس عددی تصادفی تولید کرده و با توجه به این که مقدار آن در سهم کدام فعالیت است، فعالیت مورد نیاز را انتخاب می‌کنیم. با این کار علاوه بر لحاظ کردن وزن به دست آمده با استفاده از قوانین انجمنی، از عنصر تصادفی نیز بهره بردیم.

  به عنوان مثال، سه فعالیت با احتمال محاسبه شده‌ی ۱۰، ۳۰ و ۶۰ درصد داریم، سهم فعالیت اول بازه ۰ تا ۱۰، سهم فعالیت دوم بازه ۱۰ تا ۴۰ و سهم فعالیت سوم بازه ۴۰ تا ۱۰۰ است. حال عدد تصادفی بین ۰ تا ۱۰۰ تولید می‌کنیم و فعالیت منتخب، فعالیتی خواهد بود که عدد تصادفی تولید شده متعلق به بازه‌ی آن است.
- انتخاب فعالیت در حرکت رو به جلو: با استفاده از قوانین انجمنی و هستی‌شناسی فعالیت‌ها، برای حرکت رو به جلو فعالیت بعدی در سلسله فعالیت جعلی را به صورت تصادفی انتخاب کرده اما مانند انتخاب فعالیت در حرکت عقبگرد، احتمال انجام فعالیت‌ها در تعیین فعالیت بعدی به صورت تصادفی لحاظ می‌شوند.

- تصمیم توقف: پس از آن که فعالیت جایگزین در سلسله فعالیت جعلی اضافه شد؛ امکان حرکت رو به جلو یا توقف داریم. این تصمیم با مقدار احتمال توقف که هر فعالیت در هستی‌شناسی خود آن را دارد انجام می‌پذیرد و به صورت وزن‌دار لحاظ می‌شود اما در نهایت این تصمیم به صورت تصادفی بوده و در هر بار اجرا امکان توقف یا حرکت رو به جلو داریم.
- مدت زمان انجام فعالیت و تاخیر بین فعالیت‌ها: همانطور که اشاره شد، زمان انجام هر فعالیت در سلسله فعالیت جعلی با استفاده از زمان انجام فعالیت قبلی، مدت زمان انجام آن و تاخیر بین این دو فعالیت محاسبه خواهد شد. حال برای آن که فاصله‌ی انجام بین دو فعالیت مشخص همیشه یکسان نباشد، مدت زمان انجام فعالیت‌ها و تاخیر بین انجام فعالیت‌های مختلف در سلسله فعالیت جعلی، با مقداری تصادفی بین حداقل زمان مورد نیاز تا پنج برابر حداقل زمان مورد نیاز جمع زده می‌شوند تا هر بار فاصله‌ی بین انجام فعالیت‌ها نیز متفاوت باشد.

۵-۴ جمع‌بندی

در این فصل به طور کامل راهکار پیشنهادی را بررسی کرده و توصیف اجمالی راه حل را ارائه کردیم. برای طراحی هر چه بهتر این نرم‌افزار مدل تهدید را بررسی کرده و بر اساس آن معماری کلان و مدل‌سازی را با جزئیات توصیف کردیم.

فصل ۵

پیاده‌سازی و ارزیابی

برای تولید سلسله فعالیت جعلی به نحوی که از دید سکو قابل تمایز با سلسله فعالیت واقعی نباشد؛ نیاز به پیاده‌سازی دقیق مدل مطرح شده در بخش ۴-۴ و ارزیابی کامل بر اساس مدل تهدید گفته شده در بخش ۴-۲ است. در این فصل به توصیف کلی پیاده‌سازی نرم‌افزار و ارزیابی کامل آن با نمونه داده‌های متنوع بر اساس مدل تهدید سکوی صادق ولی کنجکاو خواهیم پرداخت.

۵-۱ پیاده‌سازی

در این بخش به بررسی پیاده‌سازی راهکار پیشنهادی این پژوهش طبق الگوریتم ارائه شده در بخش ۴-۴-۳ می‌پردازیم.

۵-۱-۱ معماری و ساختار پیاده‌سازی

راهکار پیشنهادی این پژوهش به صورت یک پروژه Web API است که به زبان سی‌شارپ[۱] توسعه یافته است که با استفاده از معماری سه لایه[۲] طراحی شده است که شامل لایه‌های API، دامنه[۳] و زیرساخت[۴] می‌باشد. لایه API به عنوان رابط کاربری عمل می‌کند و درخواست‌ها را از کاربران دریافت کرده و به لایه دامنه ارسال می‌کند. لایه دامنه مسئولیت مدیریت منطق کسب‌وکار و پردازش داده‌ها را بر عهده دارد. لایه زیرساخت نیز ارتباط با پایگاه داده و منابع خارجی را فراهم می‌سازد.

---
[۱] C#
[۲] Three Layered Architecture
[۳] Domain
[۴] Infrastructure

برای مدیریت داده‌ها در این راهکار پیشنهادی این پژوهش از SQL Server استفاده شده است. این پایگاه داده به عنوان سیستم مدیریت پایگاه داده[۵] عمل می‌کند و داده‌های پروژه را به صورت ساختاریافته ذخیره و مدیریت می‌نماید، که این امر باعث افزایش کارایی و سرعت دسترسی به داده‌ها می‌شود.

۵-۱-۲ کد برنامه

عملکرد اصلی راهکار پیشنهادی این پژوهش در کلاس CoreService انجام شده است که بخش‌های اصلی آن را به صورت دقیق تعریف می‌کنیم.

عملکرد اصلی راهکار پیشنهادی این پژوهش در کلاس CoreService انجام شده است که بخش‌های اصلی آن را به صورت دقیق تعریف می‌کنیم. کد مربوط به این بخش‌ها در آدرس GitHub[۶] قرار داده شده است.

- **GenerateFakeActivities**: در این تابع با توجه به شرایط کنونی خانه هوشمند، به صورت عقبگرد حرکت کرده تا به نقطه شروع تولید سلسله فعالیت جعلی برسیم (ممکن است شرایط کنونی ورود به خانه باشد و نیازی به حرکت عقبگرد نداشته باشیم).
- **CheckConditions**: در این تابع شرایط انجام فعالیتی که می‌خواهیم به سلسله فعالیت جعلی اضافه شود را بررسی می‌کنیم تا با توجه به حالت خانه هوشمند، فعالیت قابل انجام باشد.
- **GenerateRecursively**: در این تابع تولید سلسله فعالیت جعلی انجام می‌شود و در هر مرحله با توجه به فعالیت‌های بعدی ممکن، یک مرحله جلو می‌رویم تا به نقطه توقف برسیم.
- **SetResults**: پس از تولید هر فعالیت جعلی، تغییر حالت خانه هوشمند که نتیجه آن فعالیت است را اعمال می‌کنیم تا حالت خانه هوشمند از دید سکو، حالتی منطقی و درست باشد.
- **FindNextActivity**: در این تابع از بین فعالیت ممکن قابل انجام پس از فعالیت جعلی تولید شده، یکی را به صورت تصادفی (با احتمال انجام آن بر اساس قوانین انجمنی) انتخاب می‌کنیم.
- **GenerateFinalResult**: در این تابع که پس از اتمام تابع GenerateRecursively اجرا می‌شود، هر فعالیت به داده‌های ارسالی حسگرها نگاشت می‌شود تا به سکو ارسال شوند.

---

[۵] Database Management System (DBMS)

[۶] https://github.com/BehzadDara/FakeEventGenerator

۵-۲ ارزیابی

برای ارزیابی این پژوهش فرض شده است که سکو با استفاده از یک دسته‌بند به شناسایی فعالیت‌های کاربر با استفاده از اطلاعات حسگرهای خانه هوشمند می‌پردازد. در این ارزیابی میزان گمراه‌سازی سکو با تولید رخدادهای جعلی بر اساس مجموعه داده‌ای از فعالیت‌های کاربر توسط ابزار پیاده‌سازی شده اندازه‌گیری شده است. بدین منظور ابتدا هستی‌شناسی متناسب با مجموعه داده تعریف شده و پس از تولید سلسله فعالیت جعلی، ترکیب آن با سلسله فعالیت واقعی به دسته‌بند داده شده و میزان کاهش دقت آن به عنوان معیار ارزیابی در نظر گرفته شده است. این معیار نشان‌دهنده قدرت راهکار پیشنهادی در گمراه‌سازی سکو است.

۵-۲-۱ مجموعه دادگان

مجموعه داده مورد استفاده در این پژوهش مجموعه داده Orange4Home [۱۲۸] است که شامل تقریباً ۱۸۰ ساعت فعالیت‌های روزمره یک فرد است که به مدت ۴ هفته متوالی در روزهای کاری انجام شده است. این مجموعه داده شامل داده‌های ۲۳۶ حسگر ناهمگن است که به‌طور یکپارچه در سراسر یک آپارتمان پخش شده‌اند و ۲۰ دسته فعالیت که توسط فرد به‌طور دقیق در محل برچسب‌گذاری شده‌اند، و در مجموع ۴۹۳ نمونه از فعالیت‌ها را شامل می‌شود. این ویژگی‌ها، Orange4Home را به یک مجموعه داده مناسب برای ارزیابی الگوریتم‌های شناسایی فعالیت، ارزیابی الگوریتم‌های پیش‌بینی فعالیت و سایر مسائل پژوهشی مرتبط با الگوریتم‌ها و خانه‌های هوشمند تبدیل می‌کند.

خانه هوشمند استفاده شده در مجموعه داده Orange4Home شامل بخش‌های ورودی، آشپزخانه، سرویس بهداشتی، پذیرایی و راه پله در طبقه همکف و راه پله، راهرو، دفتر کار، حمام و اتاق خواب در طبقه اول است که این نقشه در اشکال ۵-۱ و ۵-۲ قابل مشاهده است.

شکل ۵-۱: نقشه خانه مجموعه داده **Orange4Home**، طبقه همکف [۱۲۸]

شکل ۵-۲: نقشه خانه مجموعه داده **Orange4Home**، طبقه اول [۱۲۸]

در هر بخش از این خانه هوشمند، حسگرهایی با انواع دودویی[۷]، عدد صحیح[۸]، عددی[۹] و دسته‌ای[۱۰] وجود دارد که تعداد این حسگرها به تفکیک بخش‌های مختلف خانه هوشمند، در شکل ۵-۱ قابل مشاهده است. تعدادی از حسگرها در مجموعه داده Orange4Home سراسری و برای کل خانه است که این مقادیر رابطه مستقیم با فعالیت فعلی کاربر و داده حسگرهای مربوط به فعالیت‌ها ندارند، در نتیجه در مجموعه داده و تولید سلسله فعالیت جعلی از این داده‌ها صرف نظر شده است و به جای ۲۳۶ حسگر، در این پژوهش از ۱۹۶ حسگر (با حذف حسگرهای سراسری) استفاده شده است.

در این مجموعه داده فعالیت‌های برچسب‌گذاری شده شامل داده‌های حسگرها می‌باشند و تمامی

[۷] Binary
[۸] Integer
[۹] Real number
[۱۰] Categorical

جدول ۵-۱: حسگرهای هر بخش خانه هوشمند

محل	دودویی	عدد صحیح	عددی	دسته‌ای	مجموع
ورودی	۳	۱	۲	۳	۹
آشپزخانه	۱۳	۲۱	۱۸	۰	۵۲
پذیرایی	۱۶	۶	۸	۷	۳۷
سرویس بهداشتی	۳	۱	۱	۰	۵
راه پله	۳	۰	۰	۰	۳
راهرو	۹	۰	۱	۰	۱۰
حمام	۹	۶	۸	۳	۲۶
دفتر کار	۹	۳	۳	۵	۲۰
اتاق خواب	۱۷	۴	۶	۷	۳۴
سراسری	۱	۱۳	۲۰	۶	۴۰
مجموع	۸۳	۵۵	۶۷	۳۱	۲۳۶

اطلاعات ارسالی از حسگرها بین شروع و پایان برچسب یک فعالیت است. هر فعالیت برچسب‌گذاری شده متعلق به یک بخش خانه است که در ادامه لیستی از این فعالیت‌ها ذکر شده است:

- ورودی: ورود، خروج
- آشپزخانه: آماده‌سازی، آشپزی، شستن ظروف
- پذیرایی: خوردن، تلویزیون دیدن، کار با رایانه
- سرویس بهداشتی: استفاده از سرویس بهداشتی
- راه پله: بالا رفتن، پایین آمدن
- حمام: استفاده از روشویی، استفاده از سرویس بهداشتی، حمام کردن
- دفتر کار: کار با رایانه، تلویزیون دیدن
- اتاق خواب: لباس عوض کردن، کتاب خواندن، خوابیدن

- تمامی بخش‌ها: تمیز کاری

همانطور که گفته شد، داده‌های حسگرها بین شروع و پایان یک فعالیت برچسب‌گذاری شده ارسال می‌شوند. در مجموعه داده Orange4Home، یک فایل csv قرار دارد که داده حسگرها با را در خود دارد. این داده‌ها شامل زمان دقیق ارسال داده توسط حسگر، نام حسگر و مقدار ارسال شده توسط آن حسگر است که نمونه داده موجود در مجموعه داده در شکل ۵-۳ قابل مشاهده است. در این شکل شروع فعالیت ورود به خانه است که در همان دقیقه ورود و باز کردن درب، داده‌های مربوط به صدای محیط، روشن شدن چراغ و دیگر عوامل محیطی به صورت خودکار به سکو ارسال می‌گردد.

Time	ItemName	Value
1/30/2017 7:58	label	START:Entrance|Entering
1/30/2017 7:58	livingroom_couch_plug_consumption	0
1/30/2017 7:58	office_desk_plug_consumption	0
1/30/2017 7:58	office_tv_plug_consumption	0
1/30/2017 7:58	livingroom_tv_plug_consumption	1
1/30/2017 7:58	livingroom_table_plug_consumption	0
1/30/2017 7:58	livingroom_table_noise	0.278974
1/30/2017 7:58	livingroom_table_noise	0.640542
1/30/2017 7:58	livingroom_table_noise	2.10256
1/30/2017 7:58	livingroom_table_noise	0.429033
1/30/2017 7:58	livingroom_table_noise	0.182365
1/30/2017 7:58	livingroom_couch_plug_consumption	0
1/30/2017 7:58	office_desk_plug_consumption	0
1/30/2017 7:58	office_tv_plug_consumption	0
1/30/2017 7:58	livingroom_tv_plug_consumption	1
1/30/2017 7:58	livingroom_table_plug_consumption	0
1/30/2017 7:58	livingroom_couch_noise	0.354432
1/30/2017 7:58	bedroom_CO2	490.88
1/30/2017 7:58	entrance_door	OPEN
1/30/2017 7:58	livingroom_couch_noise	0.165306
1/30/2017 7:58	livingroom_couch_plug_consumption	0
1/30/2017 7:58	office_desk_plug_consumption	0
1/30/2017 7:58	office_tv_plug_consumption	0
1/30/2017 7:58	livingroom_tv_plug_consumption	1
1/30/2017 7:58	livingroom_table_plug_consumption	0
1/30/2017 7:58	entrance_switch_left	ON
1/30/2017 7:58	entrance_light1	100
1/30/2017 7:58	entrance_switch_left	OFF
1/30/2017 7:58	staircase_light	100
1/30/2017 7:58	kitchen_luminosity	20
1/30/2017 7:58	entrance_door	CLOSED

شکل ۵-۳: نمونه داده ارسالی از حسگرها در مجموعه داده **Orange4Home**

۵-۲-۲ هستی‌شناسی

برای استخراج هستی‌شناسی از مجموعه داده Orange4Home، ابتدا مجاورت بخش‌های مختلف خانه

بر اساس نقشه تعریف شده است. سپس نسخه‌ای فیلتر شده از مجموعه داده که فقط دارای فعالیت‌های برچسب‌گذاری شده است را تهیه کردیم که بخشی از آن در شکل ۴-۵ قابل مشاهده است. از این مجموعه داده برای محاسبه احتمال رخداد یک فعالیت پس از دیگری استفاده شده است تا بتوانیم سلسله فعالیت جعلی غیر قابل تمایز با استفاده از قوانین انجمنی تولید کنیم (جزئیات استفاده از احتمالات در قوانین انجمنی، در بخش ۴-۴-۲ آورده شده است).

Time	ItemName	Value
1/30/2017 7:58	label	START:Entrance|Entering
1/30/2017 8:01	label	STOP:Entrance|Entering
1/30/2017 8:01	label	START:Staircase|Going_up
1/30/2017 8:02	label	STOP:Staircase|Going_up
1/30/2017 8:02	label	START:Bathroom|Showering
1/30/2017 8:18	label	STOP:Bathroom|Showering
1/30/2017 8:18	label	START:Bathroom|Using_the_sink
1/30/2017 8:22	label	STOP:Bathroom|Using_the_sink
1/30/2017 8:22	label	START:Staircase|Going_down
1/30/2017 8:23	label	STOP:Staircase|Going_down
1/30/2017 8:23	label	START:Living_room|Watching_TV
1/30/2017 8:45	label	STOP:Living_room|Watching_TV
1/30/2017 8:45	label	START:Toilet|Using_the_toilet
1/30/2017 8:48	label	STOP:Toilet|Using_the_toilet
1/30/2017 8:48	label	START:Staircase|Going_up
1/30/2017 8:48	label	STOP:Staircase|Going_up
1/30/2017 8:48	label	START:Office|Computing
1/30/2017 11:45	label	STOP:Office|Computing
1/30/2017 11:45	label	START:Staircase|Going_down
1/30/2017 11:46	label	STOP:Staircase|Going_down
1/30/2017 11:46	label	START:Kitchen|Preparing
1/30/2017 11:48	label	STOP:Kitchen|Preparing
1/30/2017 11:48	label	START:Kitchen|Cooking
1/30/2017 12:02	label	STOP:Kitchen|Cooking
1/30/2017 12:02	label	START:Living_room|Eating
1/30/2017 12:16	label	STOP:Living_room|Eating
1/30/2017 12:16	label	START:Kitchen|Washing_the_dishes
1/30/2017 12:25	label	STOP:Kitchen|Washing_the_dishes
1/30/2017 12:25	label	START:Living_room|Cleaning
1/30/2017 12:26	label	STOP:Living_room|Cleaning
1/30/2017 12:26	label	START:Living_room|Computing

شکل ۵-۴: بخشی از مجموعه داده فیلتر شده **Orange4Home** که فقط فعالیت‌ها در آن هستند.

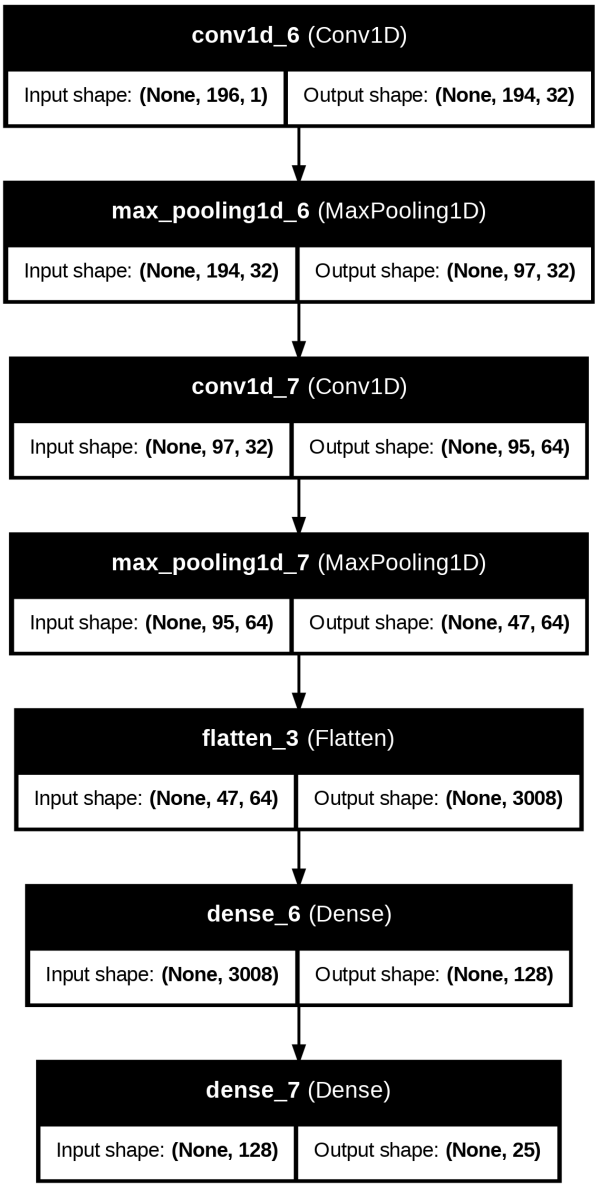
سپس، داده حسگرها مربوط به هر فعالیت برچسب‌گذاری شده را استخراج می‌کنیم و پس از تولید سلسله فعالیت جعلی بر اساس هستی‌شناسی و قوانین انجمنی فعالیت‌ها، به جای هر فعالیت یکی از سلسله داده حسگرهای موجود در مجموعه داده برای آن فعالیت را به صورت تصادفی انتخاب و به سکو ارسال می‌کنیم. سلسله داده‌های ارسالی حسگرها به ازای هر فعالیت در هستی‌شناسی آن فعالیت و ارتباط با حسگرها قرار

داده شده است.

۵-۲-۳ دسته‌بند

همان‌طور که بیان شد، برای ارزیابی کیفیت سلسله فعالیت جعلی تولید شده توسط ابزار پیشنهادی، میزان افت دقت تشخیص فعالیت توسط یک دسته‌بند اندازه‌گیری می‌شود. برای این منظور از دسته‌بند استفاده شده در کار پژوهشی اقوامی و همکاران [۱۲۶] به عنوان معیار ارزیابی استفاده شده است که با دقت ۹۸ درصد فعالیت متناظر با رخدادهای مجموعه‌داده Orange4Home را تشخیص می‌دهد. این دسته‌بند یک شبکه عصبی پیچشی است که معماری آن در شکل ۵-۵ آورده شده است. ورودی این دسته‌بند یک بردار شامل ۱۹۶ مقدار است که هر کدام نشانگر یک حسگر در خانه است و مقدار آن بیانگر آخرین رخداد گزارش شده از آن حسگر است. این مقادیر با پیش‌پردازش مجموعه داده محاسبه شده است و برای پر کردن مقادیر ناموجود در هر لحظه، از مقدار قبل از آن و یا در صورت عدم وجود، از مقدار بعد از آن استفاده شده است. بدین ترتیب دسته‌بند با داشتن یک وضعیت از تمام حسگرها، فعالیت کاربر در آن زمان را تشخیص می‌دهد.

روش Stratified که برای آموزش این دسته‌بند مورد استفاده قرار می‌گیرد، به ویژه در مواجهه با داده‌های غیرمتوازن بسیار مناسب است. این روش تضمین می‌کند که در فرآیند آموزش، تمامی کلاس‌ها حضور داشته باشند، که این امر بهبود دقت و قدرت تشخیص مدل را به همراه دارد [۱۲۹]. برای ارزیابی راهکار پیشنهادی، آموزش با کل مجموعه داده Orange4Home انجام شده است.

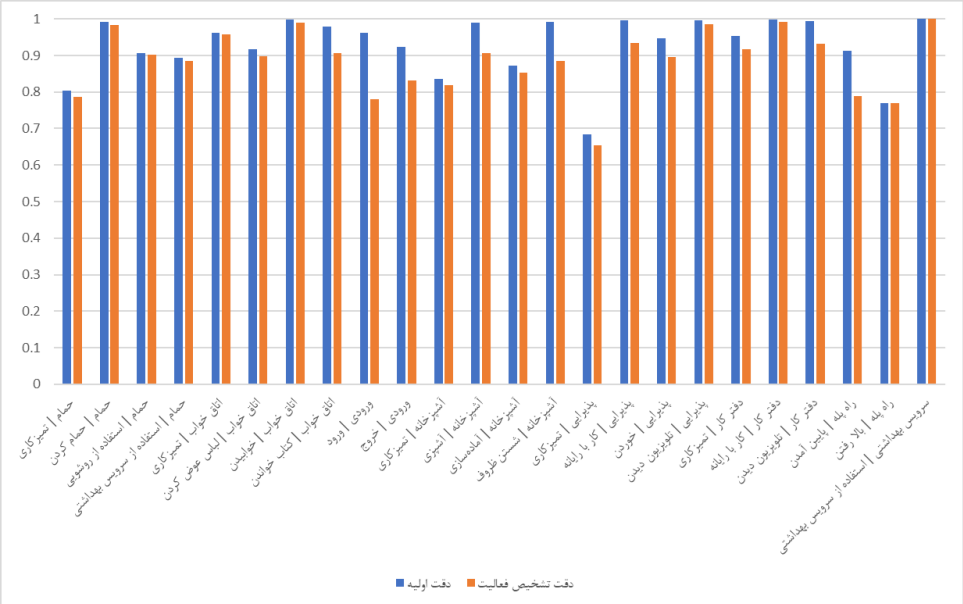


شکل ۵-۵: معماری دسته‌بند استفاده شده

۵-۲-۴ نتایج

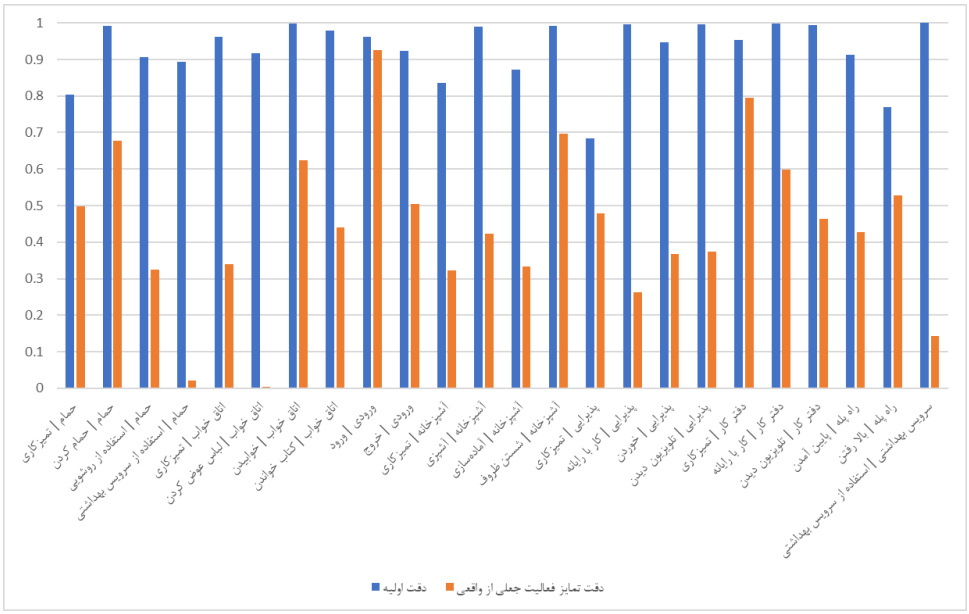
دسته‌بند مورد ارزیابی، فعالیت‌های مجموعه داده Orange4Home را با دقت ۹۸ درصد تشخیص می‌دهد. برای ارزیابی راهکار ارائه شده این پژوهش دو عامل دقت تشخیص فعالیت جعلی به عنوان یک فعالیت کامل به صورت مستقل و دقت تشخیص فعالیت واقعی در زمان دریافت فعالیت واقعی و جعلی به صورت در هم تنیده شده را بررسی می‌کنیم.

در صورتی که دسته‌بند سلسله فعالیت جعلی را به صورت مجزا دریافت کند، بالا بودن دقت تشخیص آن به عنوان یک فعالیت مستقل و کامل به معنی نزدیکی سلسله فعالیت جعلی به واقعیت است که دقت دسته‌بند از ۹۸ درصد به ۸۸ درصد رسیده که به معنی عملکرد قابل قبول از راهکار ارائه شده در پژوهش است. برای ارزیابی در این بخش به دسته‌بند که با مجموعه داده Orange4Home آموزش دیده شده است، چندین بار سلسله فعالیت جعلی تولید شده برای یک یا دو روز کامل را دادیم و دقت تشخیص آن را اندازه‌گیری کردیم. دقت تشخیص فعالیت به تفکیک نوع فعالیت در شکل ۵-۶ قابل مشاهده است.



شکل ۵-۶: دقت تشخیص هر نوع فعالیت جعلی به عنوان یک فعالیت مستقل در مقایسه با دقت اولیه

در صورتی که دسته‌بند سلسله فعالیت جعلی و واقعی را همزمان و به صورت در هم تنیده شده دریافت کند، پایین بودن دقت تشخیص آن به عنوان یک فعالیت به معنی عدم تمایز سلسله فعالیت جعلی از واقعی است که دقت دسته‌بند در این حالت از ۹۸ درصد به ۴۳ کاهش پیدا کرده که اثباتی بر عدم توانایی دسته‌بند در تمییز فعالیت جعلی و واقعی از یکدیگر است. برای ارزیابی در این بخش به دسته‌بند که با مجموعه داده Orange4Home آموزش دیده شده است، چندین بار به سلسله فعالیت واقعی، سلسله فعالیت جعلی تزریق کرده و آن سلسله فعالیت جعلی را با برچسب فعالیت‌های واقعی برای یک یا دو روز کامل را دادیم و دقت تشخیص آن را اندازه‌گیری کردیم. دسته‌بند در این حالت برچسب فعالیت واقعی را به همراه داده واقعی آن فعالیت و داده جعلی فعالیتی دیگر دریافت می‌کند. دقت به دست آمده از دسته‌بند به تفکیک نوع فعالیت در شکل ۵-۷ قابل مشاهده است.



شکل ۵-۷: دقت تشخیص هر نوع فعالیت در ترکیب فعالیت جعلی و واقعی و تمایز آن در مقایسه با دقت اولیه

فصل ۶

نتیجه‌گیری

در این پژوهش راهکاری مبتنی بر تولید رویداد جعلی برای حفظ حریم خصوصی کاربر در برابر سکوی نامعتمد در خانه هوشمند ارائه شد. این راهکار برای فریب سکو و پنهان‌سازی فعالیت‌های حساس کاربر، از تولید سلسله فعالیت جعلی بنا به درخواست کاربر بهره می‌برد تا سکوی اینترنت اشیاء متوجه تفاوت بین فعالیت‌های جعلی و واقعی نشود. در ادامه ضمن جمع‌بندی، به طرح پیشنهادهایی برای پژوهش‌های آینده و تکمیل راهکار خواهیم پرداخت.

۶-۱ جمع‌بندی

هدف این پژوهش تولید سلسله فعالیت جعلی برای حفظ حریم خصوصی کاربر بوده که برای این هدف از روش تولید رویداد جعلی بهره برداری شده است. تولید رویداد جعلی بر اساس هستی‌شناسی خانه هوشمند بوده که دانش اولیه‌ی آن توسط فرد خبره با روش دانش‌محور فراهم شده است. در این پژوهش از هستی‌شناسی برای مدل‌سازی معنایی رویدادهای مربوط به دستگاه‌های اینترنت اشیاء و سناریوهای رفتارهای کاربران در محیط خانه هوشمند استفاده شد تا بر مبنای آن امکان تولید رویدادها و رفتارهای جعلی به صورت غیرقابل تمایز با رویدادها و رفتارهای واقعی در محیط خانه هوشمند فراهم گردد.

برای عدم تشخیص سکوی نامعتمد، از عوامل تصادفی‌ساز در انجام و زمان‌بندی هر رویداد جعلی استفاده شده و هر رویداد توسط راهکار پیشنهادی پژوهش، برچسب جعلی می‌خورد و به سکو ارسال می‌شود. کنش مربوط به رویدادها توسط راهکار پیشنهادی پژوهش نظارت شده و کنش‌های مربوط به رویدادهای جعلی کنار گذاشته می‌شود تا بهره‌وری خانه هوشمند پایین نیاید.

رویکرد استفاده شده در این پژوهش که بر مبنای هستی‌شناسی بوده، تا کنون کمتر مورد توجه قرار گرفته است و امید است تا حریم خصوصی کاربران در خانه هوشمند با استفاده از نتایج این پژوهش، حفظ شود. بر اساس نتایج به دست آمده، دسته‌بند مورد ارزیابی که دقت تشخیص فعالیتش برای مجموعه داده Orange4Home برابر با ۹۸ درصد بود، در صورت دریافت سلسله فعالیت جعلی که با استفاده از راهکار پیشنهادی این پژوهش تولید شده است دقتش به طور میانگین برابر با ۸۸ درصد است که عدم کاهش دقت در تشخیص فعالیت به معنی نزدیک به واقعیت بودن سلسله فعالیت جعلی است و در صورت دریافت سلسله فعالیت جعلی و واقعی به همراه یکدیگر، دقتش برابر با ۴۳ درصد است که کاهش دقت در تشخیص فعالیت به معنی تمایزناپذیری فعالیت‌های جعلی از فعالیت‌های واقعی می‌باشد.

۶-۲ پیشنهادهایی برای پژوهش‌های آینده

جهت تکمیل راهکار ارائه شده در این پژوهش جهت حفظ حریم خصوصی کاربر در خانه هوشمند، پیشنهادهای زیر ارائه می‌گردد:

- ذخیره‌ی وضعیت موجودیت‌ها از دیدگاه سکو: اگر سکوی اینترنت اشیاء توانایی ذخیره و استنتاج وضعیت فعلی موجودیت‌های خانه هوشمند را داشته باشد، پس از مدتی به دلیل متغایر بودن وضعیت فعلی یک موجودیت و وضعیت مورد انتظار، توانایی تشخیص جعلی بودن آخرین فعالیتی که اقدام به تغییر وضعیت آن موجودیت کرده را دارد. برای حل این مشکل، نیاز است تا راهکارهای تولیدکننده سلسله فعالیت جعلی علاوه بر استفاده از هستی‌شناسی، وضعیت مورد انتظار هر موجودیت از دید سکو را ذخیره کرده و سلسله فعالیت جعلی را بر اساس آن تولید کند.
- ترکیب روش دانش‌محور ارائه شده با روش‌های داده‌محور: رفتار کلی افراد در خانه هوشمند با گذر زمان تغییر می‌کند و سکو همواره در حال به‌روزرسانی اطلاعات خود از خانه هوشمند است. هستی‌شناسی فعالیت افراد در خانه هوشمند و احتمال توالی آن‌ها که در قوانین انجمنی و تولید سلسله فعالیت جعلی مورد استفاده قرار می‌گیرد می‌تواند با استفاده از روش‌های داده‌محور به‌روزرسانی شده تا سلسله فعالیت جعلی مبتنی بر آخرین فعالیت‌های افراد باشند. برای نزدیک بودن هر چه بیشتر سلسله فعالیت جعلی تولید شده می‌توان به فعالیت‌های جدیدتر وزن بیشتری اختصاص داد تا مبنای اصلی آخرین فعالیت‌ها باشند. به‌روزرسانی هستی‌شناسی به صورت روزانه و حتی هفتگی توسط فرد خبره امری دشوار است که می‌توان با استفاده از روش‌های دانش‌محور این کار را به صورت خودکار انجام داد.

مراجع

[1] S. R. Department, "Internet of things - number of connected devices worldwide 2015-2025," 2019. Statista. https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide (Accessed: 2023-11-12).

[2] H. Park *et al.*, "Energy-efficient privacy protection for smart home environments using behavioral semantics," *Sensors*, vol.14, no.9, pp.16235–16257, 2014.

[3] K. L. Courtney, "Privacy and senior willingness to adopt smart home information technology in residential care facilities," *Methods of Information in Medicine*, vol.47, no.01, pp.76–81, 2008.

[4] C. Goyal, "Intuition behind correlation – definition and its types," 2021. Statista. https://www.analyticsvidhya.com/blog/2021/04/intuition-behind-correlation-definition-and-its-types/(Accessed: 2023-11-12).

[5] J. y. Hong, E. h. Suh, and S.-J. Kim, "Context-aware systems: A literature review and classification," *Expert Systems with Applications*, vol.36, no.4, p.8509–8522, 2009.

[6] N. D. Rodriguez, M. P. Cuellar, J. Lilius, and M. D. Calvo-Flores, "A survey on ontologies for human behavior recognition," *ACM Computing Surveys (CSUR)*, vol.46, no.4, p.1–33, 2014.

[7] R. Yasaei, F. Hernandez, and M. A. A. Faruque, "Iot-cad: Context-aware adaptive anomaly detection in iot systems through sensor association," in *Proceedings of the 39th International Conference on Computer-Aided Design*, p.1–9, 2020.

[8] N. D. Rodriguez *et al.*, "A fuzzy ontology for semantic modelling and recognition of human behaviour," *Knowledge-Based Systems*, vol.66, p.46–60, 2014.

[9] P. Lago, C. Jiménez-Guarín, and C. Roncancio, "Contextualized behavior patterns for ambient assisted living," in *Proceedings of the Conference*, pp.132–145, 2015.

[10] P. Rashidi and D. J. Cook, "Keeping the resident in the loop: Adapting the smart home to the user," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol.39, no.5, p.949–959, 2009.

[11] P. Remagnino, H. Hagras, N. Monekosso, and S. Velastin, "Ambient intelligence," in *Ambient Intelligence: A Novel Paradigm*, p.1, 2006.

[12] D. J. Cook, J. C. Augusto, and V. R. Jakkula, "Ambient intelligence: Technologies, applications, and opportunities," *Pervasive and Mobile Computing*, vol.5, no.4, pp.277–298, 2009.

[13] R. Mojarad, F. Attal, A. Chibani, and Y. Amirat, "A context-aware hybrid framework for human behavior analysis," in *Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI*, vol.2020-Novem, pp.460–465, 2020.

[14] N. Díaz Rodríguez, M. P. Cuéllar, J. Lilius, and M. Delgado Calvo-Flores, "A fuzzy ontology for semantic modelling and recognition of human behaviour," *Knowledge-Based Systems*, vol.66, pp.46–60, Aug 2014.

[15] J. Mattioli, G. Pedroza, S. Khalfaoui, and B. Leroy, "Combining data-driven and knowledge-based ai paradigms for engineering ai-based safety-critical systems," in *Workshop on Artificial Intelligence Safety (SafeAI)*, 2022.

[16] W. Wang and Y. Yang, "Towards data- and knowledge-driven artificial intelligence: A survey on neuro-symbolic computing," *arXiv preprint arXiv:2210.15889*, 2022.

[17] N. F. Noy and D. L. McGuinness, "Ontology development 101: A guide to creating your first ontology," 2001. Accessed: 2023-11-12.

[18] E. R. Harold. *Effective XML: 50 Specific Ways to Improve Your XML*. Addison-Wesley Professional, 2004.

[19] World Wide Web Consortium *et al.*, "Rdf 1.1 concepts and abstract syntax," 2014.

[20] A. Gomez-Perez and O. Corcho, "Ontology languages for the semantic web," *IEEE Intelligent Systems*, vol.17, no.1, p.54–60, 2002.

[21] D. L. McGuinness, F. V. Harmelen, *et al.*, "Owl web ontology language overview," Tech. Rep. 10, W3C Recommendation, 2004.

[22] B. Motik, P. F. Patel-Schneider, B. Parsia, *et al.*, "Owl 2 web ontology language: Structural specification and functional-style syntax," Tech. Rep. 65, W3C Recommendation, 2009.

[23] I. Horrocks, P. F. Patel-Schneider, H. Boley, *et al.*, "Swrl: A semantic web rule language combining owl and ruleml," Tech. Rep. 79, W3C Member Submission, 2004.

[24] TechTarget, "Association rules in data mining," Statista. https://www.techtarget.com/searchbusinessanalytics/definition/association-rules-in-data-mining (Accessed: 2023-03-20).

[25] W. Zhou, Y. Jia, Y. Yao, L. Zhu, L. Guan, Y. Mao, P. Liu, and Y. Zhang, "Discovering and understanding the security hazards in the interactions between iot devices, mobile apps, and clouds on smart home platforms," in *28th USENIX Security Symposium (USENIX Security 19)*, pp.1133–1150, 2019.

[26] Z. Wang, D. Liu, Y. Sun, X. Pang, P. Sun, F. Lin, J. C. Lui, and K. Ren, "A survey on iot-enabled home automation systems: Attacks and defenses," *IEEE Communications Surveys & Tutorials*, 2022.

[27] "IFTTT - connect your apps," https://ifttt.com/ (Accessed: 2023-10-01).

[28] "SmartThings," https://www.smartthings.com/ (Accessed: 2023-10-01).

[29] "openHAB," https://www.openhab.org/ (Accessed: 2023-10-01).

[30] "Easier automation, bigger impact," https://zapier.com/ (Accessed: 2023-10-01).

[31] "Home app," https://www.apple.com/home-app/ (Accessed: 2023-10-01).

[32] "Power automate," https://powerautomate.microsoft.com/en-us/ (Accessed: 2023-10-01).

[33] I. Zavalyshyn, A. Legay, A. Rath, and E. Riviere, "SoK: privacy-enhancing smart home hubs," *Proceedings on Privacy Enhancing Technologies*, vol.4, pp.24–43, 2022.

[34] G. Bajaj *et al.*, "A study of existing ontologies in the iot-domain," *arXiv preprint arXiv:1707.00112*, 2017.

[35] M. Compton, P. Barnaghi, L. Bermudez, R. Garcia-Castro, O. Corcho, S. Cox, J. Graybeal, M. Hauswirth, C. Henson, A. Herzog, *et al.*, "The ssn ontology of the w3c semantic sensor network incubator group," *Web Semantics: Science, Services and Agents on the World Wide Web*, 2012.

[36] L. Xue, Y. Liu, P. Zeng, H. Yu, and Z. Shi, "An ontology based scheme for sensor description in context awareness system," in *Information and Automation, 2015 IEEE International Conference on*, pp.817–820, IEEE, 2015.

[37] A. Gyrard, S. K. Datta, C. Bonnet, and K. Boudaoud, "Standardizing generic cross-domain applications in internet of things," in *Globecom Workshops (GC Wkshps), 2014*, pp.589–594, IEEE, 2014.

[38] D. J. Russomanno, C. Kothari, and O. Thomas, "Sensor ontologies: from shallow to deep models," in *Proceedings of the Thirty-Seventh Southeastern Symposium on System Theory, 2005. SSST'05.*, pp.107–112, IEEE, 2005.

[39] I. Niles and A. Pease, "Towards a standard upper ontology," in *Proceedings of the international conference on Formal Ontology in Information SystemsVolume 2001*, pp.2–9, ACM, 2001.

[40] P. Hirmer, M. Wieland, U. Breitenbücher, and B. Mitschang, "Dynamic ontology-based sensor binding," in *Advances in Databases and Information Systems: 20th East European Conference, ADBIS 2016, Lecture Notes in Computer Science, Vol. 9809* (J. Pokorný, M. Ivanović, B. Thalheim, and P. Saloun, eds.), (Cham), pp.323–337, Springer International Publishing, 2016.

[41] Y. Shi, G. Li, X. Zhou, and X. Zhang, "Sensor ontology building in semantic sensor web," in *Internet of Things*, pp.277–284, Springer, 2012.

[42] L. Daniele, F. den Hartog, and J. Roes, "Study on semantic assets for smart appliances interoperability," 2015.

[43] L. Daniele, M. Solanki, F. den Hartog, and J. Roes, "Interoperability for smart appliances in the iot world," in *International Semantic Web Conference*, pp.21–29, Springer, 2016.

[44] S. Dey and R. Dasgupta, "Sensor knowledge representation with spatiotemporal annotation: An energy sensor ontology use case," in *Pervasive Computing and Communications Workshops (PERCOM Workshops)*, pp.455–459, IEEE, 2014.

[45] B. Balaji, A. Bhattacharya, G. Fierro, J. Gao, J. Gluck, D. Hong, A. Johansen, J. Koh, J. Ploennigs, and e. a. Y. Agarwal, "Brick: Towards a unified metadata schema for buildings," in *Proceedings of the ACM International Conference on Embedded Systems for Energy-Efficient Built Environments (BuildSys)*, ACM, 2016.

[46] S. Hachem, T. Teixeira, and V. Issarny, "Ontologies for the internet of things," in *Proceedings of the 8th Middleware Doctoral Symposium*, 2011.

[47] M. Baldauf, S. Dustdar, and F. Rosenberg, "A survey on context-aware systems," *International Journal of Ad Hoc and Ubiquitous Computing*, vol.2, no.4, pp.263–277, 2007.

[48] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, "Context aware computing for the internet of things: A survey," *IEEE Communications Surveys & Tutorials*, vol.16, no.1, pp.414–454, 2014.

[49] H. Chen, T. Finin, and A. Joshi, "An ontology for context-aware pervasive computing environments," *The Knowledge Engineering Review*, vol.18, no.03, pp.197–207, 2003.

[50] J. R. Hobbs and F. Pan, "An ontology of time for the semantic web," *ACM Transactions on Asian Language Information Processing (TALIP)*, vol.3, no.1, pp.66–85, 2004.

[51] G. Okeyo, L. Chen, H. Wang, and R. Sterritt, "Dynamic sensor data segmentation for real-time knowledge-driven activity recognition," *Pervasive and Mobile Computing*, vol.10, no.Part B, pp.155–172, 2014.

[52] I.-H. Bae, "An ontology-based approach to ADL recognition in smart homes," *Future Generation Computer Systems*, vol.33, pp.32–41, 2014.

[53] K. Lee, J. Lee, and M.-P. Kwan, "Location-based service using ontology-based semantic queries: A study with a focus on indoor activities in a university context," *Computers, Environment and Urban Systems*, vol.62, pp.41–52, 2017.

[54] L. Chen, C. Nugent, and A. Al-Bashrawi, "Semantic data management for situation-aware assistance in ambient assisted living," in *Proceedings of the 11th International Conference on Information Integration and Web-based Applications & Services*, 2009.

[55] L. Chen, C. D. Nugent, and H. Wang, "A knowledge-driven approach to activity recognition in smart homes," *IEEE Transactions on Knowledge and Data Engineering*, vol.24, no.6, pp.961–974, 2011.

[56] D. Brickley, "Basic geo (wgs84 lat/long) vocabulary," Documento informal escrito en colaboración, 2003.

[57] T. Flury, G. Privat, and F. Ramparany, "Owl-based location ontology for context-aware services," in *Proceedings of the Artificial Intelligence in Mobile Systems (AIMS 2004)*, pp.52–57, 2004.

[58] S. I. Kim and H. S. Kim, "Ontology based location reasoning method using smart phone data," in *Information Networking (ICOIN), 2015 International Conference on*, pp.509–514, IEEE, 2015.

[59] B. Szász, R. Fleiner, and A. Micsik, "iloc–building indoor navigation services using linked data," *Add journal name here*, p.Add pages here, Add year here.

[60] R. Fikes and Q. Zhou, "A reusable time ontology," Tech. Rep., Tech. rep., 2000.

[61] J. Pustejovsky, J. Castaño, R. Ingria, R. Saurí, R. Gaizauskas, A. Setzer, and G. Katz, "Timeml: Robust specification of event and temporal expressions in text," in *Fifth International Workshop on Computational Semantics*, pp.28–34, 2003.

[62] S. Dey and R. Dasgupta, "Sensor knowledge representation with spatio-temporal annotation: An energy sensor ontology use case," in *IEEE International Conference on Pervasive Computing and Communication Workshops (PERCOM WORKSHOPS)*, (Budapest), pp.455–459, IEEE, 2014.

[63] C. Zhang, C. Cao, Y. Sui, and X. Wu, "A chinese time ontology for the semantic web," *Knowledge-Based Systems*, vol.24, no.7, pp.1057–1074, 2011.

[64] Q. Ni, I. Pau de la Cruz, and A. B. García Hernando, "A foundational ontology-based model for human activity representation in smart homes," *Journal of Ambient Intelligence and Smart Environments*, 2016.

[65] E. M. Tapia, S. S. Intille, and K. Larson, "Activity recognition in the home using simple and ubiquitous sensors," in *International Conference on Pervasive Computing*, p.158–175, Springer, 2004.

[66] T. van Kasteren and B. Krose, "Bayesian activity recognition in residence for elders," in *2007 3rd IET International Conference on Intelligent Environments*, p.209–212, IET, 2007.

[67] J. Boger *et al.*, "A planning system based on markov decision processes to guide people with dementia through activities of daily living," *IEEE Transactions on Information Technology in Biomedicine*, vol.10, no.2, p.323–333, 2006.

[68] X. Meng *et al.*, "Human driving behavior recognition based on hidden markov models," in *2006 IEEE International Conference on Robotics and Biomimetics*, p.274–279, IEEE, 2006.

[69] D. Zhang *et al.*, "Modeling individual and group actions in meetings with layered hmms," *IEEE Transactions on Multimedia*, vol.8, no.3, p.509–520, 2006.

[70] T. L. van Kasteren *et al.*, "Hierarchical activity recognition using automatically clustered actions," in *Ambient Intelligence: Second International Joint Conference on AmI 2011*, p.82–91, Springer, 2011.

[71] J. Hoey *et al.*, "Rapid specification and automated generation of prompting systems to assist people with dementia," *Pervasive and Mobile Computing*, vol.7, no.3, p.299–318, 2011.

[72] M. Brand *et al.*, "Coupled hidden markov models for complex action recognition," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, p.994–999, IEEE, 1997.

[73] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *International Conference on Pervasive Computing*, p.1–17, Springer, 2004.

[74] O. Brdiczka *et al.*, "Learning situation models in a smart home," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol.39, no.1, p.56–63, 2008.

[75] U. Maurer *et al.*, "Location and activity recognition using ewatch: A wearable sensor platform," *Ambient Intelligence in Everyday Life: Foreword by Emile Aarts*, p.86–102, 2006.

[76] L. Liao *et al.*, "Hierarchical conditional random fields for gps-based activity recognition," in *Robotics Research: Results of the 12th International Symposium ISRR*, p.487–506, Springer, 2007.

[77] J. Lester *et al.*, "A hybrid discriminative/generative approach for modeling human activities," in *Proceedings of the 19th International Joint Conference on Artificial Intelligence*, p.766–772, 2005.

[78] D. H. Wilson, D. Wyatt, and M. Philipose, "Using context history for data collection in the home," Tech. Rep. 577, Cognitive Science Research Paper-University Of SUSSEX CSRP, 2005.

[79] U. Maurer *et al.*, "Activity recognition and monitoring using multiple sensors on different body positions," in *International Workshop on Wearable and Implantable Body Sensor Networks (BSN'06)*, IEEE, 2006.

[80] M. Delgado, M. Ros, and M. A. Vila, "Correct behavior identification system in a tagged world," *Expert Systems with Applications*, vol.36, no.6, p.9899–9906, 2009.

[81] M. Ermes *et al.*, "Detection of daily activities and sports with wearable sensors in controlled and uncontrolled conditions," *IEEE Transactions on Information Technology in Biomedicine*, vol.12, no.1, p.20–26, 2008.

[82] M. Perkowitz, M. Philipose, K. Fishkin, and D. J. Patterson, "Mining models of human activities from the web," in *Proceedings of the 13th International Conference on World Wide Web*, p.573–582, 2004.

[83] E. M. Tapia, T. Choudhury, and M. Philipose, "Building reliable activity models using hierarchical shrinkage and mined ontology," in *International Conference on Pervasive Computing*, p.17–32, Springer, 2006.

[84] P. Palmes, H. K. Pung, T. Gu, W. Xue, and S. Chen, "Object relevance weight pattern mining for activity recognition and segmentation," *Pervasive and Mobile Computing*, vol.6, no.1, p.43–57, 2010.

[85] H. K. Ngankam, H. Pigot, and S. Giroux, "Ontodomus: A semantic model for ambient assisted living system based on smart homes," *Electronics*, vol.11, no.7, p.1143, 2022.

[86] I.-H. Bae, "An ontology-based approach to adl recognition in smart homes," *Future Generation Computer Systems*, vol.33, p.32–41, 2014.

[87] K. Wongpatikaseree *et al.*, "Activity recognition using context-aware infrastructure ontology in smart home domain," in *2012 Seventh International Conference on Knowledge, Information and Creativity Support Systems*, p.50–57, IEEE, 2012.

[88] L. Chen, C. D. Nugent, and H. Wang, "A knowledge-driven approach to activity recognition in smart homes," *IEEE Transactions on Knowledge and Data Engineering*, vol.24, no.6, p.961–974, 2011.

[89] D. Riboni and C. Bettini, "Owl 2 modeling and reasoning with complex human activities," *Pervasive and Mobile Computing*, vol.7, no.3, p.379–395, 2011.

[90] L. Chen and C. Nugent, "Ontology-based activity recognition in intelligent pervasive environments," *International Journal of Web Information Systems*, vol.5, no.4, p.410–430, 2009.

[91] D. Riboni *et al.*, "Is ontology-based activity recognition really effective?," in *2011 IEEE International Conference on Pervasive Computing and Communications Workshops*, p.427–431, IEEE, 2011.

[92] S. Zhang, P. McCullagh, C. Nugent, H. Zheng, and N. Black, "An ontological framework for activity monitoring and reminder reasoning in an assisted environment," *Journal of Ambient Intelligence and Humanized Computing*, vol.4, p.157–168, 2013.

[93] L. Chen, C. Nugent, and A. Al-Bashrawi, "Semantic data management for situation-aware assistance in ambient assisted living," in *Proceedings of the 11th International Conference on Information Integration and Web-Based Applications & Services*, p.298–305, 2009.

[94] G. Okeyo, L. Chen, and H. Wang, "Combining ontological and temporal formalisms for composite activity modelling and recognition in smart homes," *Future Generation Computer Systems*, vol.39, p.29–43, 2014.

[95] G. Meditskos *et al.*, "Ontology patterns for complex activity modelling," in *Theory, Practice, and Applications of Rules on the Web: 7th International Symposium, RuleML 2013*, p.144–157, Springer, 2013.

[96] D. Riboni and C. Bettini, "Cosar: Hybrid reasoning for context-aware activity recognition," *Personal and Ubiquitous Computing*, vol.15, p.271–289, 2011.

[97] L. Chen, G. Okeyo, H. Wang, R. Sterritt, and C. Nugent, "A systematic approach to adaptive activity modeling and discovery in smart homes," in *2011 4th International Conference on Biomedical Engineering and Informatics (BMEI)*, vol.4, p.2192–2196, IEEE, 2011.

[98] L. Chen, C. Nugent, and G. Okeyo, "An ontology-based hybrid approach to activity modeling for smart homes," *IEEE Transactions on Human-Machine Systems*, vol.44, no.1, p.92–105, 2013.

[99] A. S. A. Sukor, A. Zakaria, N. A. Rahim, L. M. Kamarudin, R. Setchi, and H. Nishizaki, "A hybrid approach of knowledge-driven and data-driven reasoning for activity recognition in smart homes," *Journal of Intelligent & Fuzzy Systems*, vol.36, no.5, p.4177–4188, 2019.

[100] K. Gayathri, K. Easwarakumar, and S. Elias, "Probabilistic ontology based activity recognition in smart homes using markov logic network," *Knowledge-Based Systems*, vol.121, p.173–184, 2017.

[101] D. Riboni, T. Sztyler, G. Civitarese, and H. Stuckenschmidt, "Unsupervised recognition of interleaved activities of daily living through ontological and probabilistic reasoning," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, p.1–12, 2016.

[102] C. Bettini, G. Civitarese, and R. Presotto, "Caviar: Context-driven active and incremental activity recognition," *Knowledge-Based Systems*, vol.196, p.105816, 2020.

[103] G. Zhang, X. Liu, and Y. Yang, "Time-series pattern based effective noise generation for privacy protection on cloud," *IEEE Transactions on Computers*, vol.64, no.5, pp.1456–1469, 2014.

[104] "The 15 biggest data breaches of the 21st century | cso online," https://www.csoonline.com/article/2130877/the-biggest-data-breaches-of-the-21st-century.html (Accessed: 2023-04-15).

[105] "World's biggest data breaches & hacks — information is beautiful," https://www.informationisbeautiful.net/visualizations/worlds-biggest-data-breaches-hacks/ (Accessed: 2023-04-15).

[106] "Privacy policy - smartthings," https://www.smartthings.com/privacy(Accessed: 2023-04-15).

[107] "Privacy policy - ifttt," https://ifttt.com/terms(Accessed: 2023-04-15).

[108] S. Schoettler, A. Thompson, R. Gopalakrishna, and T. Gupta, "Walnut: A low-trust trigger-action platform," 2020.

[109] I. Zavalyshyn, N. Santos, R. Sadre, and A. Legay, "My house, my rules: A private-by-design smart home platform," in *EAI MobiQuitous*, 2020.

[110] Y.-H. Chiang, H.-C. Hsiao, C.-M. Yu, and T. H.-J. Kim, "On the privacy risks of compromised trigger-action platforms," in *European Symposium on Research in Computer Security*, pp.251–271, Springer, 2020.

[111] R. Xu, Q. Zeng, L. Zhu, H. Chi, X. Du, and M. Guizani, "Privacy leakage in smart homes and its mitigation: Ifttt as a case study," *IEEE Access*, vol.7, pp.63457–63471, 2019.

[112] Y. Chen, A. R. Chowdhury, R. Wang, A. Sabelfeld, R. Chatterjee, and E. Fernandes, "Data privacy in trigger-action systems," in *2021 IEEE Symposium on Security and Privacy (SP)*, pp.501–518, IEEE, 2021.

[113] M. Götz, S. Nath, and J. Gehrke, "Maskit: Privately releasing user context streams for personalized mobile applications," in *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*, 2012.

[114] H. Chi, Q. Zeng, X. Du, and L. Luo, "Pfirewall: Semantics-aware customizable data flow control for smart home privacy protection," 2021.

[115] Y. Chen, M. Alhanahnah, A. Sabelfeld, R. Chatterjee, and E. Fernandes, "Practical data access minimization in trigger-action platforms," Online, 2022.

[116] S. A. Osia *et al.*, "A hybrid deep learning architecture for privacy-preserving mobile analytics," *IEEE Internet of Things Journal*, vol.7, no.5, pp.4505–4518, 2020.

[117] M. A. Erdogdu, N. Fawaz, and A. Montanari, "Privacy-utility trade-off for time-series with application to smart-meter data," in *AAAI Workshop: Computational Sustainability*, 2015.

[118] Y.-S. Moon *et al.*, "Publishing time-series data under preservation of privacy and distance orders," in *Database and Expert Systems Applications: 21th International Conference, DEXA 2010, Bilbao, Spain, August 30-September 3, 2010, Proceedings, Part II*, vol.21, Springer Berlin Heidelberg, 2010.

[119] V. Rastogi and S. Nath, "Differentially private aggregation of distributed time-series with transformation and encryption," in *Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data*, 2010.

[120] C. Yin *et al.*, "Improved collaborative filtering recommendation algorithm based on differential privacy protection," *The Journal of Supercomputing*, vol.76, pp.5161–5174, 2020.

[121] R. Masood *et al.*, "Privacy preserving release of mobile sensor data," *arXiv preprint arXiv:2205.06641*, 2022.

[122] H. Kargupta *et al.*, "Random-data perturbation techniques and privacy-preserving data mining," *Knowledge and Information Systems*, vol.7, pp.387–414, 2005.

[123] D. Zheng *et al.*, "An enhanced differential private protection method based on adaptive iterative wiener filtering in discrete time series," *International Journal of Network Security*, vol.23, no.2, pp.351–358, 2021.

[124] N. Saleheen *et al.*, "msieve: Differential behavioral privacy in time series of mobile sensor data," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2016.

[125] M. Malekzadeh, R. G. Clegg, and H. Haddadi, "Replacement autoencoder: A privacy-preserving algorithm for sensory data analysis," *arXiv preprint arXiv:1710.06564*, 2017.

[126] M. Aghvamipanah and M. Amini, "Activity recognition protection for iot trigger-action platforms," in *Proceedings of IEEE European Symposium on Security and Privacy*, IEEE, 2024.

[127] X. Yuan, P. He, Q. Zhu, and X. Li, "Adversarial examples: Attacks and defenses for deep learning," *IEEE Transactions on Neural Networks and Learning Systems*, vol.30, no.9, pp.2805–2824, 2019.

[128] Amiqual4Home, “Orange4home,” 2024. Accessed: 2024-08-12.

[129] S. learn Developers, “sklearn model selection stratifiedshufflesplit,” 2024. Accessed: 29-Aug-2024.

واژه‌نامه فارسی به انگلیسی

آ

آشفته‌سازی ..... Randomization

ا

ابر ..... Cloud

استدلال منطقی ..... Logical reasoning

اطلاعات حساس ..... Sensitive information

اطمینان ..... Confidence

امنیت ..... Safety

انتها به انتها ..... End-to-End

اینترنت اشیاء ..... Internet of things

ب

بازیابی اطلاعات ..... Data recovery

برخط ..... Online

بیز ساده ..... Naive bayes

پ

پایگاه دانش ..... Knowledge base

پردازش زبان طبیعی Natural Language Processing

پشتیبانی ..... Support

ت

تا حدودی مشاهده‌پذیر ..... Partially observable

تجارت الکترونیک ..... Electronic commerce

تجهیزات پوشیدنی ..... Bearable device

ج

جابجایی ..... Replacement

ح

حریم خصوصی ..... Privacy

حسگر ..... Sensor

حمله کانال جانبی ..... Side channel attack

خ

خانه هوشمند ..... Smart home

خودرمزگذار ..... AutoEncoder

خودکارسازی ..... Automation

خوشه‌بندی ..... Clustering

د

داده کاوی ..... Data mining

داده محور ..... Data-driven

دانش محور ..... Knowledge-driven

درخت تصمیم ..... Decision tree

درشت‌دانه ..... Coarse-grained

دسته‌بند ..... Classifier

ر

رفتار جعلی ..... Fake behavior

رفتار کاربر ..... User behavior

روابط ..... Relations

رهانا ..... Trigger

ریزدانه ..... Fine-grained

ز

زمینه ..... Context

س

سکوی اینترنت اشیاء . Internet of things platform

ش

شبکه بیزین ..... Bayesian network

شبکه عصبی ..... Neural network

شبکه عصبی پرسپترون چندلایه ..... Multi-layer perceptron neural network

شبکه منطق مارکوف ..... Markov login network

شروع سرد ..... Cold-start

شناسایی ..... Recognition

ص

صوری ..... Formal

ط

طول جغرافیایی ..... Longitude

ع

عرض جغرافیایی ..... Latitude

عملگر ..... Actuator

عوامل تصادفی‌ساز ..... Randomizing factors

ف

فراداده ..... Metadata

فراسطح ..... Meta-level

فرد خبره ..... Expert

فیلترینگ ..... Filtering

ق

قواعد ..... Rules

قوانین انجمنی ..... Associative rules

ک

کلاس‌بندی ..... Classification

کنش ..... Action

م

ماشین بردار پشتیبان .... Support vector machine

مدل پنهان مارکوف ..... Hidden markov model

مدل پنهان مارکوف جفت‌شده .... Coupled hidden markov model

مراقبت‌های بهداشتی ..... Health care

مشمول‌سازی ..... Subsumption reasoning

معناشناسی ..... Semantics

مفاهیم ..... Concepts

منطق توصیفی ..... Description logic

منطق دودویی ..... Binary logic

منطق فازی ..... Fuzzy logic

منطق مرتبه اول ..... First order logic

موتور جستجو ..... Search engine

موجودیت ..... Entity

موجودیت‌های زمانی ..... TemporalThings

موجودیت‌های فضایی ..... SpatialThings

مهاجم ..... Attacker

مهندسی دانش ..... Knowledge engineering

ن

ناهمگونی ..... Heterogeneity

نزدیکترین همسایه ..... Nearest neighbor

نگاشت ..... Mapping

نمونه‌ها ..... Instances

نمونه خصمانه ..... Adversarial example

و

وب معنایی ..... Semantic web

ه

هاب ..... Hub

هستی‌شناسی ..... Ontology

همبستگی ..... Correlation

ی

یادگیری استقرایی ..... Inductive learning

یادگیری ماشین ..... Machine learning

یادگیری نیمه‌نظارتی ..... Semi-supervised learning

واژه‌نامه انگلیسی به فارسی

A

Action ..... کنش

Actuator ..... عملگر

Adversarial example ..... نمونه خصمانه

Associative rules ..... قوانین انجمنی

Attacker ..... مهاجم

AutoEncoder ..... خودرمزگذار

Automation ..... خودکارسازی

B

Bayesian network ..... شبکه بیزین

Bearable device ..... تجهیزات پوشیدنی

Binary logic ..... منطق دودویی

C

Classification ..... کلاس‌بندی

Classifier ..... دسته‌بند

Cloud ..... ابر

Clustering ..... خوشه‌بندی

Coarse-grained ..... درشت‌دانه

Cold-start ..... شروع سرد

Concepts ..... مفاهیم

Confidence ..... اطمینان

Context ..... زمینه

Correlation ..... همبستگی

Coupled hidden markov model ..... مدل پنهان مارکوف جفت‌شده

D

Data mining ..... داده کاوی

Data recovery ..... بازیابی اطلاعات

Data-driven ..... داده محور

Description logic ..... منطق توصیفی

Decision tree ..... درخت تصمیم

E

Electronic commerce ..... تجارت الکترونیک

End-to-End ..... انتها به انتها

Entity ..... موجودیت

Expert ..... فرد خبره

F

Fake behavior ..... رفتار جعلی

Filtering ..... فیلترینگ

Fine-grained ..... ریزدانه

First order logic ..... منطق مرتبه اول

Formal ..... صوری

Fuzzy logic ..... منطق فازی

H

Health care ..... مراقبت‌های بهداشتی

Heterogeneity ..... ناهمگونی

Hidden markov model..... مدل پنهان مارکوف

Hub..... هاب

I

Inductive learning ..... یادگیری استقرایی

Instances..... نمونه‌ها

Internet of things ..... اینترنت اشیاء

Internet of things platform . سکوی اینترنت اشیاء

K

Knowledge base..... پایگاه دانش

Knowledge engineering ..... مهندسی دانش

Knowledge-driven..... دانش محور

L

Latitude ..... عرض جغرافیایی

Logical reasoning..... استدلال منطقی

Longitude ..... طول جغرافیایی

M

Machine learning ..... یادگیری ماشین

Mapping ..... نگاشت

Markov login network ..... شبکه منطق مارکوف

Metadata ..... فراداده

Meta-level ..... فراسطح

Multi-layer..... شبکه عصبی پرسپترون چندلایه

perceptron neural network

N

Naive bayes ..... بیز ساده

Natural Language Processing پردازش زبان طبیعی

Nearest neighbor ..... نزدیکترین همسایه

Neural network..... شبکه عصبی

O

Online..... برخط

Ontology ..... هستی‌شناسی

P

Privacy ..... حریم خصوصی

R

Randomization..... آشفته‌سازی

Randomizing factors ..... عوامل تصادفی‌ساز

Recognition..... شناسایی

Relations ..... روابط

Replacement..... جابجایی

Rules..... قواعد

S

Safety ..... امنیت

Search engine..... موتور جستجو

Semantic web ..... وب معنایی

Semantics..... معناشناسی

Semi-supervised learning.... یادگیری نیمه‌نظارتی

Sensitive information..... اطلاعات حساس

Sensor ..... حسگر

Side channel attack..... حمله کانال جانبی

T

U

Abstract

...

Keywords: Internet of Things, Privacy, Smart Home, Ontology, IoT Platform

Sharif University of Technology

Department of Computer Engineering

M.Sc. Thesis

Ontology based Fake User Behavior Generation for Privacy Preservation in Smart Home

By:

Behzad Dara

Supervisor:

Dr. Morteza Amini

December 2023