

## تلفیق داده ها و اینترنت اشیا برای محیط های فراگیر هوشمند :

### یک مطالعه

#### چکیده

اینترنت اشیا (IoT) قرار است به یکی از پیشرفت های کلیدی فناوری زمان ما تبدیل شود؛ به شرطی که بتوانیم از پتانسیل کامل آن استفاده کنیم. انتظار می رود که تعداد اشیا متصل به IoT تا سال 2020 به 50 میلیارد برسد زیرا هجوم گسترده ای از اشیا متنوع به تدریج در حال ظهور است. از این رو، انتظار می رود که IoT یک تولید کننده اصلی داده (big data) باشد. به اشتراک گذاری و تعامل داده ها و سایر منابع می تواند کلیدی برای ایجاد محیط های پایدار فراگیر، مانند شهرهای هوشمند و جوامع باشد. تلفیق و تجزیه و تحلیل به موقع داده های بزرگ که از IoT و سایر منابع به دست می آید، برای ایجاد تصمیم گیری دقیق و با قابلیت اطمینان و مدیریت بسیار کارآمد محیط های فراگیر می تواند یک چالش بزرگ آینده باشد. هوش کامپیوتری در این چالش نقش کلیدی ایفا می کند. تعدادی از مطالعات بر روی تلفیق داده ها وجود دارد. با این حال، اینها عمدتاً بر روی حوزه های کاربردی خاص یا دسته بندی های خاص تمرکز می کنند. هدف از این مقاله بررسی ادبیات تلفیق داده ها برای IoT با یک توجه خاص به روش های ریاضی (از جمله روش های احتمالاتی، هوش مصنوعی و نظریه اعتقاد) و محیط های خاص IoT (محیط های توزیع شده، ناهمگن، غیر خطی و ردیابی شیء) می باشد. فرصت ها و چالش هایی برای هر یک از روش ها و محیط های ریاضی ارائه شده است. تحولات آینده، از جمله حوزه های نوظهور که ذاتاً از تلفیق داده ها و IoT، وسایل نقلیه مستقل (خود مختار)، یادگیری عمیق برای تلفیق داده ها و شهرهای هوشمند بهره می برند، مورد بحث قرار می گیرد.

**کلمات کلیدی:** اینترنت اشیا، داده های بزرگ، تلفیق داده ها، هوش محاسباتی و مصنوعی، محاسبات با کارایی بالا، شهر های هوشمند، جوامع هوشمند، محیط های فراگیر.

## 1. مقدمه

اینترنت اشیا (IoT) [1] قرار است به یکی از پیشرفت های مهم فناوری زمان ما تبدیل شود به شرطی که بتوانیم از پتانسیل کامل آن استفاده کنیم [2]. IoT "یک زیرساخت جهانی برای جامعه اطلاعاتی است و خدمات پیشرفته را بوسیله اتصالات (فیزیکی و مجازی) اشیا بر اساس تعامل فناوری اطلاعات و ارتباطات در حال توسعه موجود را به وجود آورد." [3] در گزارش سال 2008 [4] شورای ملی اطلاعات ایالات متحده (NIC) از میان شش کلید فن آوری های مدنی که می تواند به طور بالقوه بر قدرت ایالات متحده تأثیر بگذارد، به نام IoT نامگذاری شد. IoT یک ابزار محاسباتی فراگیر است که توسط Mark Weiser طراحی شده است. شکل 1 حوزه کاربردی IoT را نشان می دهد: خانه های هوشمند، سیستم های هشدار دهنده، خرید هوشمندانه، ابزارهای هوشمند، شهرهای هوشمند، جاده های هوشمند، مراقبت های بهداشتی، سیستم های آتش نشانی، سیستم های شناسایی تهدیدها، ردیابی و نظارت. انتظار می رود که تعداد اشیا متصل به IoT تا سال 2020 به 50 میلیارد برسد زیرا هجوم وسیعی از اشیای متنوع به تدریج در حال ظهور است [5]. هدف اصلی این افزایش تعداد و انواع اشیا IoT، تولید اطلاعات مفید درباره محیط اطراف است تا آنها را هوشمندانه تر کند. این امر با فراهم آوردن دسترسی محیط ها به اطلاعات مورد نیاز خود از طریق جمع آوری و تجزیه و تحلیل داده های گذشته، حال و آینده انجام می شود. داده ها اجازه می دهند تصمیم بهینه در مورد ما و محیط های ما در زمان واقعی انجام شود.

انتظار می رود IoT یک تولید کننده اصلی داده (big data) باشد. این داده ها توسط فروشندگان مختلف تولید می شود که داده ها را بعنوان سرویس ارائه می دهند. به اشتراک گذاری و تعامل داده ها و سایر منابع می تواند کلیدی برای ایجاد محیط های پایدار فراگیر مانند شهرهای هوشمند و جوامع باشد [6]. تلفیقی از انواع و فرمهای مختلف داده، به عنوان مثال، تلفیق داده ها، برای افزایش کیفیت داده ها و تصمیم گیری، اهمیت ویژه ای در محیط های

فراگیر دارد. تلفیق داده ها به عنوان «نظریه، تکنیک ها و ابزارهایی است که برای تلفیق داده های حسگر یا داده های مشتق شده از داده های حس شده به یک فرمت نمایشی رایج استفاده می شود» [7] تعریف می شود. تلفیق و تجزیه و تحلیل big data (حجم، سرعت، تنوع و صحت) که از IoT و سایر منابع به دست می آید، برای تصمیم گیری و مدیریت بسیار کارآمد، قابل اعتماد و دقیق محیط های فراگیر یک چالش بزرگ آینده می باشد. هوش محاسباتی در این چالش یک نقش کلیدی را ایفا می کند.

## A. مشارکت این کار

اصطلاح "اینترنت اشیاء" در ابتدا توسط کوین اشتون در سال 1999 معرفی شد [8]. با این حال، پژوهش و توسعه کلیدی کارها در IoT در سال 2010 آغاز شده است. یکی از اولین و مشهورترین مطالعات توسط Atzori و همکارانش بود. [1]، در سال 2010، جایی که آنها بلوک های بنیادی IoT و برنامه های کاربردی آن را توصیف کردند. این بررسی در [9] گسترش یافت. Al-Fuqaha و همکاران در سال 2015، یک مطالعه از فناوری ها، پروتکل ها و برنامه های کاربردی فعال سازی IoT ارائه شد [10]. در [11]، تصویر ماهواره ای روش های تلفیق چند حسگر بررسی شده است. این بررسی [12] ادبیات ادغام داده ها را به سه دسته تقسیم می کند: ارتباط داده ها، تلفیق تصمیم گیری و تخمین وضعیت. در همان سال 2013، یک مقاله دیگر [13] ادبیات تلفیق داده ها را براساس خواص داده های دریافت شده مانند نقص ها، همبستگی و ناسازگاری داده دسته بندی کرد. در [14]، پیشرفت و چالش های مختلف برای تلفیق داده ها در سیستم های حمل و نقل هوشمند مورد بحث قرار گرفته است. یک مطالعه اخیر [15] بررسی انتقادی تحولات تلفیق داده ها براساس سنسورهای تعبیه شده در دستگاه های تلفن همراه با تمرکز ویژه بر شناسایی فعالیت های انسانی ارائه کرده است. جدول 1 مطالعات مربوط به مقاله ما را بیشتر توضیح می دهد. خلاصه ای از تمرکز اصلی هر مقاله مربوطه مورد بررسی قرار گرفته است.

واضح است که هرچند تعدادی از مطالعه ها بر روی تلفیق داده ها وجود دارد، این ها عمدتاً بر روی برنامه های خاص (مانند ITS، سنسورهای تعبیه شده در دستگاه های تلفن همراه)، حوزه یا دسته بندی هایی (به عنوان مثال، خواص داده ها، میان افزار) تمرکز می کنند. هدف این مقاله این است که:

✓ بررسی ادبیات مربوط به تلفیق داده ها برای IoT با تمرکز خاص بر روی

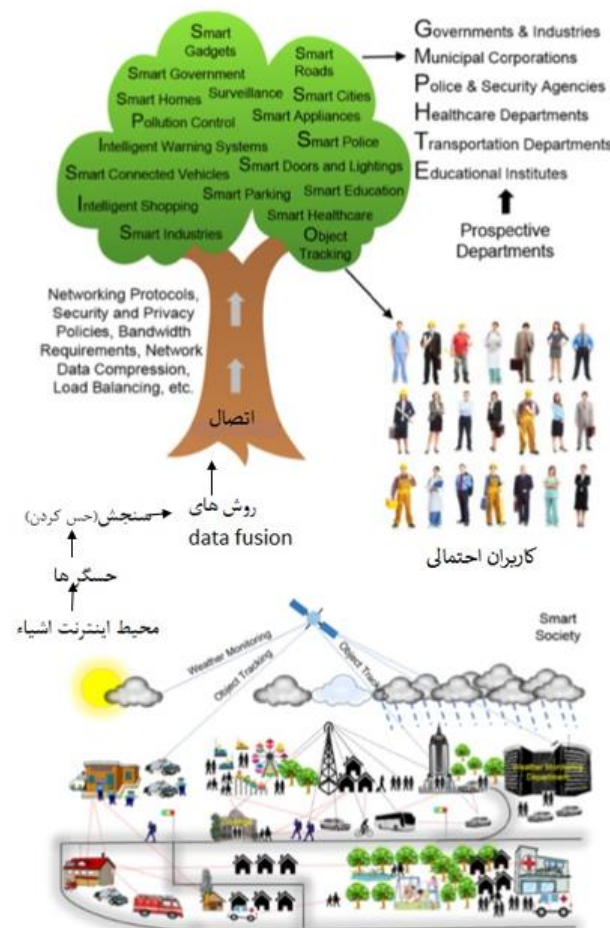
روش های ریاضی (از جمله روش های احتمالاتی، هوش مصنوعی، و نظریه اعتقاد)، و

محیط های ویژه IoT (محیط های توزیع شده، ناهمگن، غیر خطی و ردیابی شیء).

✓ برای کشف فرصت ها و چالش ها برای هر یک از روش های ریاضی بازبینی شده و محیط های IoT

✓ برای بررسی حوزه در حال ظهور که به طور ذاتی از تلفیق داده ها و IoT (از جمله شهرهای هوشمند و وسایل نقلیه

مستقل) بهره مند می شوند



شکل 1: اطلاعات گرافیکی برای نشان دادن چشم انداز IoT با توجه به Data Fusion.

## B. ساختار مقاله

مقاله به شش بخش تقسیم شده است. بخش دوم در مورد فرصت ها و چالش های تلفیق داده ها و همچنین مشخصه های IoT بحث می کند. بخش سوم ادبیات تلفیقی داده ها را براساس روش های ریاضی مورد استفاده برای تلفیق داده ها بررسی می کند. اینها شامل روش های احتمالی، هوش مصنوعی و نظریه اعتقاد است. بخش چهارم ادبیات را بر اساس محیط های خاص IoT، از جمله توزیعی، ناهمگن و غیر خطی و ردیابی شیء مورد بحث قرار می دهد. هر روش ریاضی و محیط تلفیق داده ها برای IoT که در بخش های سوم و چهارم مورد بحث قرار گرفته است، با فرصت ها و چالش های مربوطه تعریف شده است. بخش پنجم درمورد حوزه نو ظهور که به طور ذاتی از تلفیق داده ها و IoT سود می برد بحث می کند؛ این شامل وسایل نقلیه مستقل، یادگیری عمیق برای تلفیق داده ها و شهرهای هوشمند است. در نهایت نتیجه در بخش ششم بیان می شود.

## 2. فرصت ها و چالش ها تلفیق داده ها

مقدار بسیار زیادی داده در طول زمان به طور سریع در محیط IoT تولید می شود. چگونه می توان این حجم زیاد از داده، کامل و بسیار دقیق ساخته شوند، یک مشکل باز است که باید حل شود، زیرا کیفیت اطلاعات نقش مهمی در تصمیم گیری دارد. اطلاعات قابل اعتماد و دقیق بسیار مهم است. این می تواند با تلفیق داده یا تلفیق اطلاعات (اصطلاحاتی که می توانند جایگزین استفاده شوند) حاصل شود. تلفیق داده یک راه موثر برای استفاده بهینه از حجم زیادی از داده ها از منابع مختلف است [11]. تلفیق داده های چند حسگر به دنبال تلفیق اطلاعات از چندین حسگر و منابع برای دستیابی به نتیجه هایی است که از یک تک سنسور یا منبع ممکن نیست [19]. تلفیق اطلاعات از سنسورهایی با ویژگی های فیزیکی مختلف، درک محیط اطراف ما را افزایش می دهد و پایه ای برای برنامه ریزی، تصمیم گیری و کنترل ماشین های مستقل و هوشمند را فراهم می کند.

## A. فرصت های تلفیق داده

تلفیق داده در نمونه اینترنت اشیا (IoT) می تواند در موفقیت خود به خاطر دلایل زیر نقش مهمی داشته باشد: [7]، [17]:

- تلفیق داده باعث می شود اطلاعات هوشمندتر، قطعی، منطقی و دقیق تر باشد که از چندین سنسور و منابع بدست می آید. اطلاعات از هر سنسور در خود ممکن است خیلی حساس نباشد.
- یک آماری موفق از تلفیق با محاسبه N مشاهدات مستقل حاصل می شود؛ می توان پیش بینی کرد که داده ها به طور بهینه ادغام می شوند.
- در IoT، یک چالش بزرگ این است که سنسورهای بسیار کم قدرت را در نظر بگیریم که جایگزینی باتری را در طول عمر خود ندارند. این نیاز برای تأمین انرژی سنسورها را محبوب می کند. این حقیقت ثابت شده است که سنسورهای با دقت بالا می توانند به مصرف مقدار زیادی از انرژی منجر شوند. برای رسیدن به این موضوع، یک مجموعه از سنسورهای با مصرف انرژی بسیار کم می تواند با دقت کم مورد استفاده قرار گیرد. با استفاده از تلفیق داده، اطلاعات بسیار دقیق ایجاد خواهد شد [20].
- تلفیق داده می تواند در رسیدگی به مسائل مربوط به داده های بزرگ IoT مفید باشد، زیرا ما داده بسیاری از سنسورها را به اطلاعات کاملتر تر و دقیق تر تلفیق می کنیم.
- یکی دیگر از مزیت های مهم تلفیق داده این است که به پنهان کردن اطلاعات مهم یا معنایی که عهده دار نتایج تلفیق شده است، کمک می کند. نمونه هایی از این موارد در کاربرد های نظامی، برخی از بخش های مهم پزشکی و در ساختمان های هوشمند هستند.

## B. چالش های تلفیق داده

تلفیق داده با چالش های متعددی روبه رو است که در ادبیات مختلف مورد بررسی قرار می گیرند. برخی از آنها به شرح زیر است:

• نقص داده: داده های سنسور در زمان مبهم است؛ این می تواند نادرست و نامشخص باشد. این رفتار در شبکه های حسگر بی سیم نا معلوم است. این نقص باید به طور موثر با استفاده از الگوریتم های تلفیق داده مورد رسیدگی قرار گیرد.

• ابهام و ناسازگاری: نه تنها عدم دقت عامل ناسازگاری داده است؛ بلکه محیطی که در آن سنسور کار می کند تا حد زیادی تاثیر پذیر است [21]. تشخیص داده های خارج از محدوده، جایگزینی و محاسبه داده ها در محیط IoT حیاتی هستند.

• تداخل طبیعی: ماهیت تداخل داده ها می تواند منجر به نتایج بر خلاف روال معمول شود. مشکل تداخل داده در شواهد استدلال گمان و قاعده تلفیق Dempster's بیشتر دیده می شود. الگوریتم تلفیق داده باید در هنگام پردازش تداخل داده، توجه مهمی داشته باشد (22).

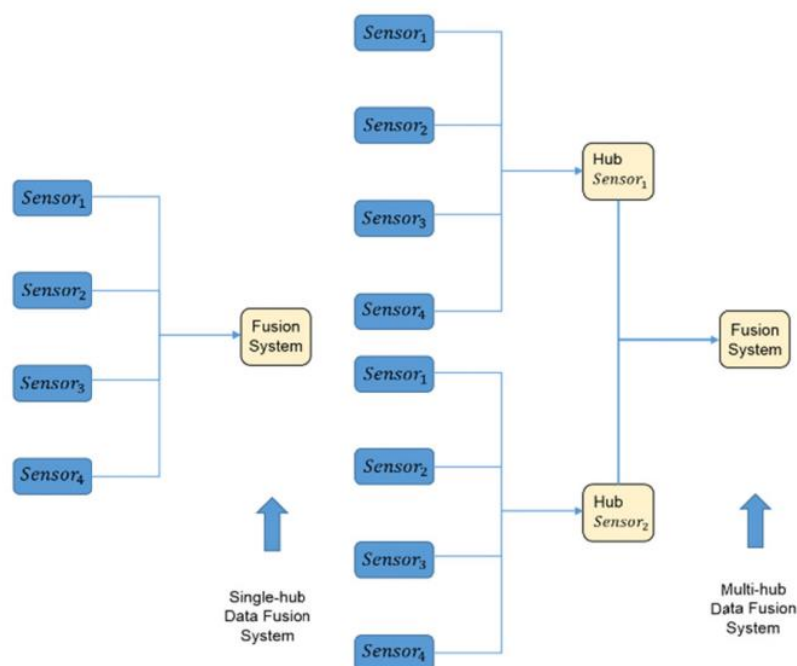
• همبستگی و هم ترازی داده ها: این مشکل بیشتر در شبکه های حسگر بی سیم (WSNs) رایج است و می تواند در یک الگوریتم تلفیقی داده اطمینان بیش از حد و یا کمتر در نتیجه حاصل شود. یک مشکل هم ترازی که همچنین به عنوان یک مشکل ثبت سنسور شناخته می شود زمانی رخ می دهد که داده سنسور از هر فریم محلی سنسور به یک فریم معمول قبل از تلفیق تبدیل شود.

• ویژگی های بی اهمیت: در محیط IoT، برنامه های کاربردی ممکن است شامل چندین صدها و هزاران پارامترهای مختلف حسگر های حس کننده باشد. این مقدار های حس شده در مجموعه های بزرگ مانند شهر های هوشمند و کارخانه های صنعتی شامل داده بی اهمیت و با اهمیت است. پردازش داده های بی اهمیت ممکن است دقت تلفیق داده را تحت تأثیر قرار دهد. بنابراین، قبل از تلفیق داده ها، باید بیشترین ویژگی های مرتبط انتخاب شوند.

• فرآیند تکراری پویا: تلفیق داده ها یک فرایند ایستا در طبیعت نیست با این حال، تکرار پویا نیازمند بازبینی منظم تخمین ها در یک محیط تلفیق است. بدون الگوریتم خبره: با تحقیقات زمان در حوزه تلفیق داده ها پیشرفت ها و الگوریتم های با عملکرد بالا در حال حاضر وجود دارد. با این حال، هنوز هم سخت است که بگوییم یک الگوریتم تلفیق داده بی عیب وجود دارد.

در IoT تلفیق اطلاعات محیط را می توان در حوزه های مختلف برای ارتقاء جنبه ی همه جانبه IoT استفاده کرد. این ها حوزه های نظارت بر محیط زیست، مراقبت های بهداشتی، مدیریت بحران، نظارت، کنترل، ردیابی، هوش جمعی و بسیاری موارد دیگر است. تلفیق داده ها در IoT می تواند در چهار مرحله انجام شود: سطح تصمیم، سطح ویژگی، سطح پیکسل و سطح سیگنال. با توجه به شکل 2، تلفیق داده های IoT را می توان با دو دیدگاه مختلف مشاهده کرد. اولاً می توان آن را به عنوان یک تک hob مشاهده کرد که هر سنسور به طور مستقیم داده ها را به مرکز تلفیق داده انتقال می دهد. دوم، اینکه می تواند با یک فرایند چند hob رخ دهد، که در آن داده ها از طریق سنسورهای مجاور عبور می کنند.

شکل 2: نمودار تلفیق داده های سنسور تک hob و چند hob



مکانیزم چند hob دارای مزایای متعددی است که در بخش IV.A مقاله بحث شده است.

### 3. روش های ریاضی برای تلفیق داده ها

تکنیک های تلفیق داده ها می توانند براساس روش های ریاضی به سه دسته گسترده تر طبقه بندی شوند:



- روش های مبتنی بر احتمال، از جمله تجزیه و تحلیل بیزی، آمار و اپراتورهای بازگشتی.
- تکنیک های هوش مصنوعی (AI) شامل یادگیری ماشین کلاسیک، منطق فازی، شبکه های عصبی مصنوعی (ANN) و ارزیابی ژنتیکی است.

- روش های تلفیق داده ها مبتنی بر نظریه شواهد

در سیستم های تلفیق داده، مدل های مختلف معماری وجود دارد؛ با این حال، برنامه های کاربردی دایما از معماری های تلفیق داده خود استفاده می کنند. برخی از مدل های پایه و اساسی شامل مدل مدیران مشترک آزمایشگاه ها (JDL)، مدل تلفیق ابشاری اصلاح شده، مدل Boyd و مدل کاربردی Dasarathy [23]–[25]. این مدل ها بعداً به لایه های مختلف تقسیم و سپس لایه ها به زیر لایه ها تقسیم می شوند. با این حال، هسته تلفیق داده ها در معماری آن نهفته است. این بدون شک در روش های تلفیق داده ها است که در آن پردازش تلفیق نهایی اتفاق می افتد. در این بخش، ادبیات را بر مبنای دسته بندی های متفاوتی از رویکردهای تلفیق داده ها بررسی و طبقه بندی کردیم. این بخش به سه زیر بخش تقسیم شده است که در آن، ادبیات بر مبنای روش های تلفیق داده احتمالاتی، روش های تلفیق داده ها مبتنی بر هوش مصنوعی و روش های تلفیق داده ها مبتنی بر نظریه شواهد مورد بررسی قرار گرفته است.

## A. الگوریتم های تلفیق داده احتمالاتی

در این بخش، ما برخی از کارهای اخیر مربوط به تکنیک های تلفیق داده احتمالی و الگوریتم ها را بررسی کردیم. تکنیک های احتمالی کلاسیک تر هستند، پیچیدگی کمتری دارند و به طور گسترده برای تلفیق داده استفاده می شوند، اگرچه ممکن است دقتشان نسبت به روش های انتگرال کم تر باشند [19]، [24]. اکثر الگوریتم های معمول برای تلفیق داده برای تلفیق داده ها در هسته هایشان احتمال دارند. با این حال، برخی از چالش هایی که برای تلفیق داده احتمالاتی وجود دارد، شامل: تحقیقات نشان می دهد که سیستم های تلفیق داده احتمالاتی نمیتوانند اطلاعات کامل را که برای تعریف و مجسم کردن حس کردن و عملیات تلفیق داده ها مورد نیاز است، فراهم آورند، سطح بالایی

از پیچیدگی در بررسی منطق غیریکنواخت از طریق احتمال وجود دارد، و برای بررسی عضویت درجه بندی شده در مقایسه با منطق فازی در یک مجموعه، مشکل است [19]. نظریه بیز، زنجیره مارکوف و روش مونت کارلو، برخی از روش های مورد مطالعه اخیر و به طور گسترده استفاده شده در تلفیق داده ها در زمان های اخیر هستند.

الگوریتم های تلفیق داده مبتنی بر احتمال به طور گسترده ای در مسائل ردیابی هدف مورد استفاده قرار می گیرند. نمونه کلاسیک ردیابی تک هدف، الگوریتم تخصیص داده احتمالی (PDA) است. چندین نسخه بهبود یافته PDA در ادبیات داده شده است و در ردیابی تک هدفه بسیار مفید هستند. ما همچنین مشاهده کردیم که وضعیت و پیچیدگی در ردیابی چند هدفه (MTT) تغییر می کند. اعتبار سنجی شیار مشکل است، زیرا شیارها با یکدیگر رقابت می کنند، بنابراین برای MTT یک الگوریتم تخصیص داده احتمالی توام (JPDA) کارآمد تر است. این یک ارزیابی احتمالی تخصیص اندازه گیری به مسیر را استفاده می کند. الگوریتم JPDA یک تقریب تک اسکن زیربهمینه به فیلتر بیزی بهمینه است که در آن تخصیص ها به صورت پیوسته بین شیارهایی که شناخته شده هستند و آخرین مشاهدات ساخته شده اند [19]، [24]، [26]. یک روش بهمینه سازی JPDA در [27] پیشنهاد شده است که به صورت محاسباتی قابل قبول از الگوریتم معمولی JPDA در برنامه های با تراکم درهم ریختگی بالاتر است. با روش بهمینه شده JPDA و چارچوبی ساده در [27] بهتر از برخی از الگوریتم های شناخته شده MTT انجام شده است. هر دو PDA و JPDA از پدیده تعصب رنج می برند. مشاهده می شود که الگوریتم PDA دارای تعصبات بیشتری نسبت به الگوریتم JPDA است. در برش محیط الگوریتم PDA پدیده تعصب دارد و الگوریتم JPDA دارای تعصب رد و همبستگی است [28].

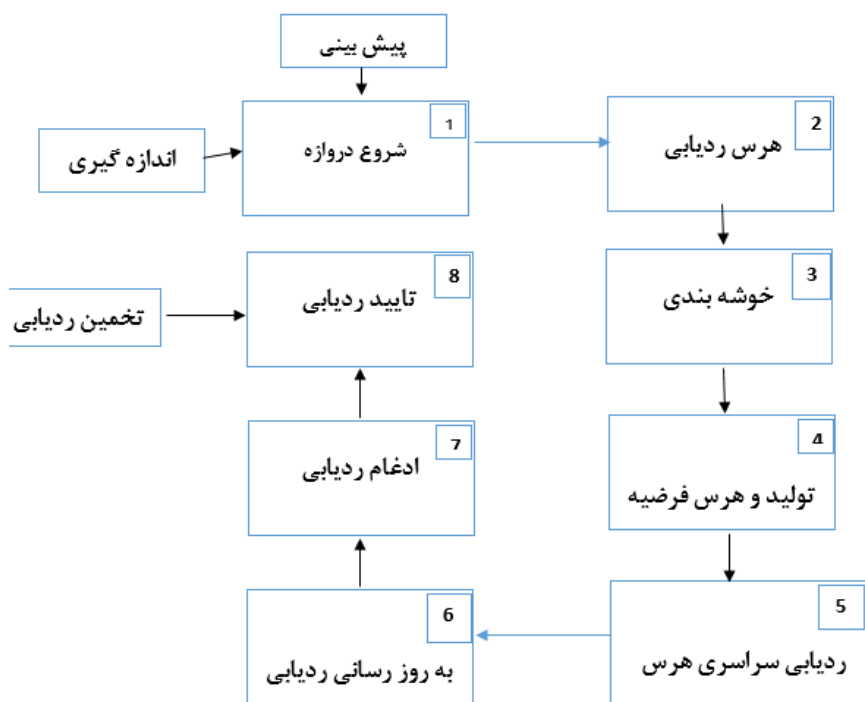
علاوه بر این، یک الگوریتم کارآمدتری نسبت به JPDA شناخته شده به عنوان الگوریتم ردیابی چند فرضیه (MHT) توسعه یافته است. این یک الگوریتم بازگشتی است. در سال های اخیر توجه بیشتر به MHT در برنامه های عملی صورت گرفته است. MHT در نظارت چندین هدف مبهم نقش مهمی دارد، زیرا این امر به تصمیمات معلق اجازه می دهد. بنابراین، تجمع آسان و طبیعی اطلاعات هدف صورت گرفته است. پیاده سازی آسان مسیر جدید با منطق ساده آغاز می شود [29]. در MHT، ردیابی چند هدف با انجام یک ارزیابی احتمالی انجام می شود که عملکرد آن را نسبت

به JPDA افزایش می دهد [19]، [24]. با این حال، الزامات حافظه MHT به صورت نمایی به عنوان افزایش فریم ها به منظور حل تخصیص ها افزایش می یابد [30].

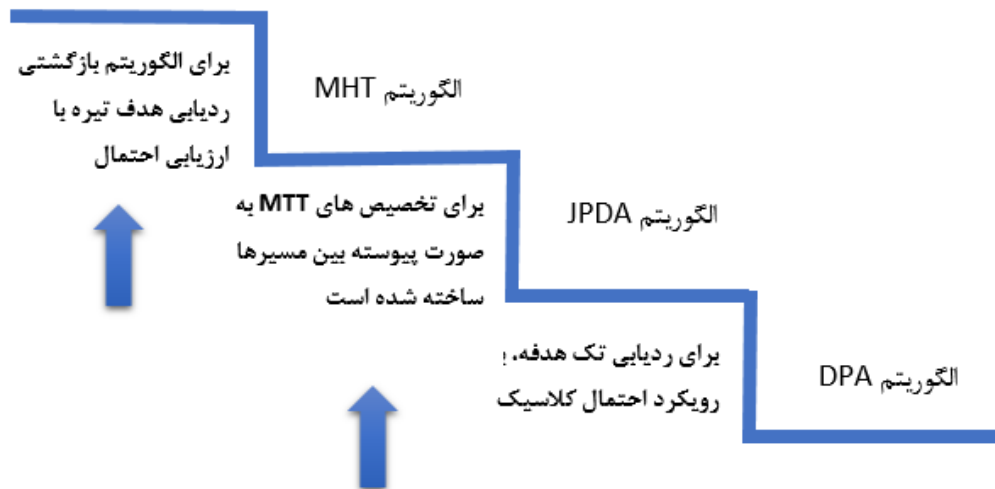
هشت مرحله الگوریتم ردیابی MHT کلاسیک [29] در شکل 3 نشان داده شده است، جایی که در چرخه  $K^{th}$ ، پیچیدگی MHT می تواند به صورت زیر بیان شود:

$$C_{MHT}(k) = \sum_{i=1}^8 C_i(k) \quad (1)$$

جایی که پیچیدگی محاسباتی و پیچیدگی هر مرحله توسط  $C_{MHT}$  و  $C_i$  مشخص می شود. در شکل 4 ما مقایسه الگوریتم های PDA، JPDA و MHT را براساس اهداف اصلی آنها تجسم می کنیم.



شکل 3: هشت مرحله ردیابی چند فرضیه (MHT) [29].



شکل 4: مقایسه تطبیقی برای الگوریتم تخصیص داده احتمالی (PDA)، تخصیص داده احتمالی مشترک (JPDA) و ردیابی چند فرضیه (MHT).

رویکرد تلفیق داده Bayesian رویکردی کلاسیک است، به این عنوان پذیرفته شده و به طور گسترده ای برای تلفیق داده استفاده می شود و این بخشی از هسته روش های مختلف تلفیق داده است [31]. این اطلاعات تلفیقی چند مدله، مبتنی بر نظریه احتمالات است [32]. رویکرد Bayesian از تعریف پیشین، مشخصات آن و محاسبات پسین تشکیل شده است. چندین روش تلفیق داده مبتنی بر رویکرد بیزی در [33] و [34] پیشنهاد شده است. در [33]، برای تشخیص هدف توزیع شده، دو روش پیشنهاد شده است. اولین روش رویکرد بیزی توزیع شده است و روش دوم آزمون نسبت احتمال تعمیم یافته (GLRT) برای WSN است. هر سنسور با توجه به محدودیت های انرژی و پهنای باند، هدف را با یک تک بیت قبل از انتقال به مرکز تلفیق تقریب می زند. رویکرد Bayesian ثابت شده که نسبت به الگوریتم های مبتنی بر GLRT عملکرد بهتر مشخصه عملیاتی گیرنده (ROC) داشته است. مقاله [33] همچنین با مشکلاتی سرو کار داشته که در آن زمان تعداد سنسورها می تواند تصادفی باشد و مکان های مرکز تلفیق ناشناخته است. برای این مشکل، یک قانون تلفیق بر اساس طرح کلی پیشنهاد شده است که توسط آمار اسکن استفاده شده است. در حالی که در [34] یک روش پیشنهاد شده است که نویسندگان از دو ردیاب بیزی با فرمول های مختلف استفاده می کنند این راه حل چالش هایی مانند سایه ها، انسداد ها، تغییرات روشنایی، درهم ریختگی و تغییرات حرکتی را کنترل می کند بنابراین، محاسبه به تدریج احتمال براساس مشاهدات جدید و استفاده از دانش پیشین

احتمال برای استنتاج، با توجه به این واقعیت که رویکرد بیزی در اینجا استفاده می شود، خروجی قوی و با دقت بالا می باشد.

اکثر روش های موجود برای به دست آوردن تخمین های سراسری وابسته به تلفیق داده هستند. هیچ کدام از این روش ها کامل نیستند. دقت الگوریتم ردیابی توسط فرایند نویز رایج تحت تاثیر قرار می گیرد. تعداد کمی از تکنیک های سازگار برای اداره فرایند نویز رایج در مطالعات [35] - [37] استفاده می شود تا معضل فیلترینگ را متوقف کند. اینها با استفاده از مجموعه ای از دو گروه موازی از فیلترها، با ذکر نام، پهنای باند وسیع و فرایند نویز رایج از پهنای باند باریک به دست می آید. با این حال، این روش ها به سناریوهای هدف خاص و هدف های تک مانور وابسته هستند. برای اداره کردن تغییرات دینامیکی هدف در تغییر سریع سناریوها و تنوع الزامات دقت، یک طرح برآورد تلفیق داده سازگار لازم است، و یک راه حل الگوریتم تلفیقی مورد نیاز است. در مقاله [38]، پهنای باند جهت دار فیلتر ماتریس اطلاعات در سطوح خاموش و برای فاز مانور ردیابی هدف انتخاب شده است. پهنای باند وسیع IMF برای غلبه بر چالش های فرایند نویز رایج انتخاب شده است. پیش بینی IMF بستگی به ضریب انتشار دارد که مستقل از مشاهده است. همچنین به جدا کردن و غیر متمرکز کردن ساده تر است. مقاله [39] نیز از تکنیک IMF برای الگوریتم سوئیچینگ جدید استفاده می کند که از نظریه طبقه بندی Bayesian [38] تکامل یافته است. این در یک IMF دو بانده برای ایجاد یک فیلتر سوئیچینگ منطبق (SAF) برای هدف پویای غیر قابل پیش بینی آغاز شده است. الگوریتم طراحی شده توسط فرمول حافظه محو شده و مجموع نوآوری های نرمال شده مربع شده است. این با فرض دو کلاس طبقه بندی در مجموعه های داده ی مستقل آماری به دست می آید. این کار نتایج کارایی بهتری از یک تک IMF برای ردیابی هدف مانور یا غیر مانور را فراهم می آورد.

زنجیره مارکوف، مونت کارلو (MCMC) و برخی از روش های تلفیقی نیز در چندین مقاله برای ردیابی هدف بر اساس نظریه احتمالات استفاده می شود. در یکی از این نوع کارها [40] یک تکنیک برای به دست آوردن اطلاعات مفید برای تشخیص و ردیابی انسان با لیزر و RGB-D در رباتیک موبایل استفاده می شود. الگوریتم ردیابی کارامدی بر اساس روش MCMC برای محیط چالش برانگیز ارائه شده است که شامل انسداد مکرر و صحنه های به هم ریخته

است. سیستم تلفیق داده پیشنهادی صحیح است که چندین انسان را در سرویس کلاسیک سناریوهای ربات کشف و ردیابی کند.

در یک کار مشابه دیگر [41]، یک الگوریتم ردیابی تلفیقی پیشنهاد شده است که از نمونه برداری فیلتر Bayesian و MCMC برای مدیریت تعاملات هدف استفاده می کند. در این کار، طولانی مدت اطلاعات تصویر برای رسیدگی به شناسایی چهره گمشده و غیرقابل شناسایی استفاده می شود. الگوریتم از هر دو مشاهدات ایستا و پویا برای ارزیابی های شکست استفاده می کند. عملکرد این الگوریتم بهتر هست از سایر الگوریتم های پیشرفته ای است که به ترتیب از مشاهدات طولانی مدت و مدل مخفی Markov استفاده نمی کنند [41]. رویکردهای تلفیق داده تلفیقی، از مزایای استفاده از هر روش متداول بهره می برند. الگوریتم های تلفیق داده تلفیقی محاسبات پیچیده تری نسبت به روش های غیر تلفیقی دارند [42]. هیچ کدام از این الگوریتم ها هیچ مزیتی بیش از دیگری ندارند. با این حال، روش های تلفیقی کارآمدتر هستند و می توانند با چالش های بیشتری نسبت به رویکرد های غیر تلفیقی برخورد کنند [42].

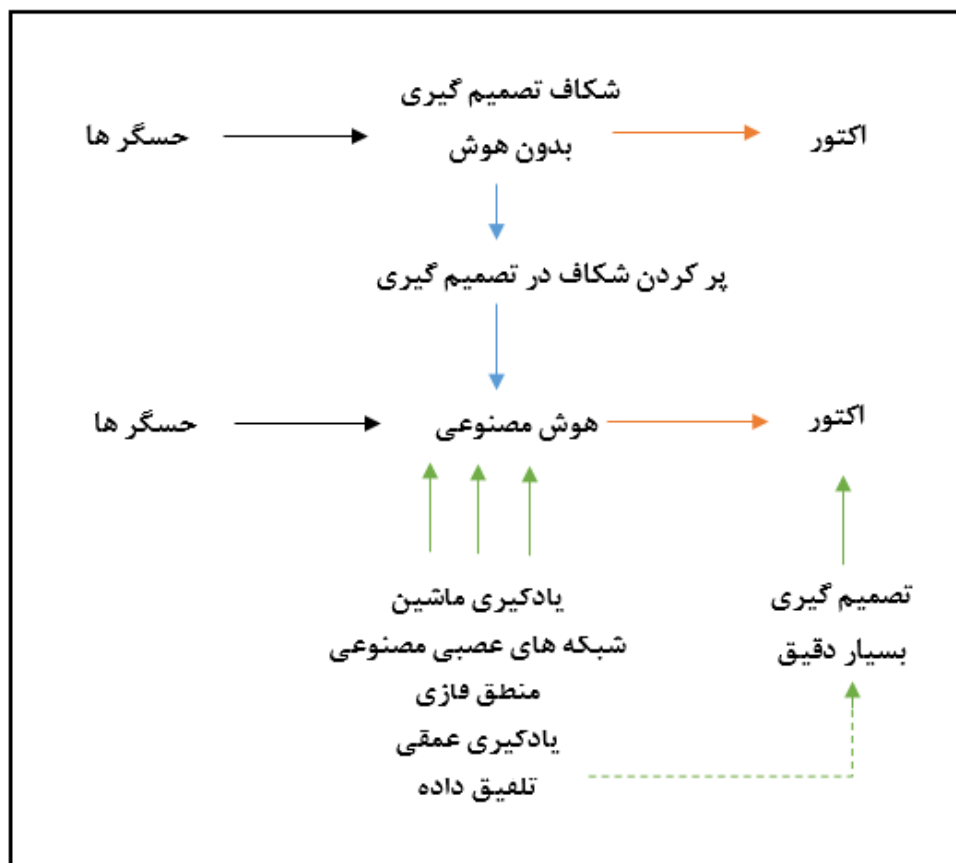
میان افزار IOT یک رابط است که ادغام و تسهیل تعامل بین عناصر مختلف که "چیزها" نام دارد و اینترنت را فراهم می کند. یک قسمت بسیار مهمی از میان افزار IoT پردازش رویداد است. پیش بینی تجزیه و تحلیل تلفیق در [43] با استفاده از میانگین مدل بیزی (BAM) مورد مطالعه قرار گرفته است. برای پردازش رویداد پیچیده فعالانه، پیش بینی تحلیل می تواند نقش یک حمایت کننده را که با تلفیق داده ها به دست می آید، بازی کند. این روش برای برنامه های IoT در مقیاس بزرگ استفاده می شود. این روش با استفاده از استنتاج انتظار-حداکثر (EM) و مدل ماتریس Gaussian برای پیش بینی های Bayesian اولیه استفاده می شود. تقریب MCMC برای میانگین گرفتن Bayesian استفاده می شود و مدل توسعه یافته براساس خوشه بندی محتوی رویداد است. نتایج نشان داد که در مقایسه با سایر رویکردهای سنتی این روش دقت بالایی دارد. با این حال در [43]، BAM و EM برای موازی سازی در تکنیک توسعه یافته همچنان مورد نیاز است. برای پردازش رویداد در IoT، چالش های مختلفی مانند ناهمگونی داده ها و دستگاه ها، حجم زیاد داده ها، زمان رخداد رویداد ناشناخته و پردازش رویداد توزیع شده وجود دارد.

یکی از مشکلات عمده در تلفیق داده های حسگر این است که حسگر ها به طور منظم مشاهدات کاذب ارائه می دهند که پیش بینی را سخت می کند. با توجه به این واقعیت، سیستم های تلفیق داده با این مقادیر نادرست از سنسور ها باید توسط شناسایی و حذف اصلاح شوند. در غیر اینصورت، آنها نتیجه ای با نرخ عدم دقت بالایی خواهند داشت که بر برآورد نهایی تأثیر خواهد گذاشت. برای رسیدگی به مشکل داده های نادرست، مقاله 44 از یک نسخه اصلاح شده رویکرد Bayesian استفاده می کند که به طور خودکار اندازه گیری سنسور و ناسازگاری را تعیین می کند. در این کار، سه تکنیک Bayesian مورد تجزیه و تحلیل قرار گرفته است. اولین تکنیک Bayesian ساده است، دومین تکنیک Bayesian متمرکز است و سومین، تلفیق به ترتیب با یک تکنیک Bayesian تغییر یافته، است. در رویکرد سوم، شناسایی داده های نادرست و بعد از حذف این نوع داده ها، با ارائه بالاترین سطح از نتایج دقیق در میان سه داده می شود. علاوه بر این، دومین بهتر از اولی عمل می کند، زیرا مکانیسمی ساخته شده برای کاهش وزن داده های جعلی است.

## 1) فرصت ها و چالش ها

- پس از بررسی ادبیات مختلف در بخش سوم. A، ما نتیجه گرفتیم که رویکردهای احتمالی برای تلفیق داده ساده هستند، پیچیدگی کمتری دارند و به طور گسترده به عنوان روش های کلاسیک تلفیق داده ها پذیرفته و در نظر گرفته شده اند [19]، [24]. تکنیک Bayesian می تواند دانش پیشین را تا حد زیادی مورد استفاده قرار دهد، بنابراین یکی از بزرگترین مزایای الگوریتم های مبتنی بر تکنیک Bayesian است. معایب کمی از رویکرد احتمالی وجود دارد:
- محدود به ذهنیت قبل است و پیدا کردن مقدارهای قبلی برای آن سخت است، به ویژه در مورد BPT محدود است.
  - نشان دادن اطلاعات کامل برای تعریف و نمایش دادن احساس سخت است.
  - دقت کمتری می تواند در نتیجه داشته باشد.
  - روش های تلفیق داده مبتنی بر احتمال پیچیدگی را با منطق غیریکنواخت افزایش می دهند.
  - روش های تلفیق داده مبتنی بر احتمال، انتخاب خوبی برای مدیریت درجه بندی عضویت نیستند.

- قادر به کنترل عدم اطمینان نیستند.
- خروجی در معرض نویز روند معمول است.
- در PDA و JPDA، زمان محاسباتی به طور نمادین بالا می رود در نتیجه تعداد اهداف، به خصوص در مسائل ردیابی افزایش می یابد.
- با این حال، با گذشت زمان، چند نسخه بهبود یافته از الگوریتم های JPDA در [45] - [48]، MHT در [49] - [51]، Bayesian در [52]، [53]، MCMC و GLRT در [54]، [55] برای غلبه بر اشکالات ذکر شده پیشنهاد شده است. با این وجود باز هم، پنجره ای برای بهبود در این الگوریتم ها برای حل این مشکلات وجود دارد که دروازه های تحقیقات آینده را باز می کند.



شکل 5: تصویری که چگونه هوش مصنوعی قدرت تصمیم گیری اینترنت چیزها (IoT) را افزایش می دهد.

زیرساخت اینترنت اشیاء



## B. الگوریتم های تلفیق داده مبتنی بر هوش مصنوعی

ما باید IoT را به عنوان اینترنت هوشمند "چیزها" تکامل دهیم تا آن را بخوبی به صورت فراگیر بسازیم. شبکه های حگسر و محرک ها در حال تبدیل شدن به ستون فقرات برنامه های مبتنی بر IoT از جمله مراقبت های بهداشتی، جوامع هوشمند، نظامی، سیستم های پیش بینی و سیستم های هشدار دهنده فعالیت های لرزه ای، سیستم های حمل و نقل هوشمند و سیستم های ردیابی هستند. هوش مصنوعی (AI)، محرک ها را قادر می سازد تا بتوانند تصمیمات بسیار دقیق و آگاهانه بر اساس داده های حس شده مانند آنچه که در شکل 5 ذکر شده، بگیرند. این بدان معنی است که هوش مصنوعی میتواند نقش مهمی در پارامتر IoT داشته باشد، به ویژه در حوزه هایی که تصمیم گیری و پیش بینی اهمیت حیاتی دارند. هوش در حال توسعه یک فرآیند تدریجی است که توسط یادگیری ماشین، هوش مصنوعی، منطق فازی، یادگیری عمیق و تلفیق داده ها به دست آمده است که در شکل 5 ذکر شده است. روند در حال توسعه هوش برای جوامع هوشمند مبتنی بر IoT در کارهای مختلفی مورد مطالعه قرار گرفته است [2]، [6]، [15].

نشانه ای از هوش مصنوعی در IoT را می توان در آخرین اتفاقات در دنیای محاسبات مشاهده کرد. Luc Julia، مدیر سابق اپل، که در حال حاضر سرپرست نوآوری های باز سامسونگ است، معماری سامسونگ را برای تعاملات چندرسانه ای (SAMI) ارائه کرد، که بخشی از هوش مصنوعی سامسونگ برای استراتژی مبتنی بر IoT است [56]. اخیراً، گوگل، DeepMind، یک شرکت هوش مصنوعی را در تلاش و تحقیق برای IoT به دست آورد. دیگر دستاوردهای بزرگ اخیر شامل بوستون دینامیک، شرکت رباتیک و Nest Labs است. همه اینها را می توان به عنوان تلاش های جدی گوگل برای تقویت توسعه ی IoT خود مشاهده کرد [57]. به طور مشابه گوگل از چین، Baidu چندین پروژه IoT را به طور خاص با تمرکز بر هوش مصنوعی اعلام می کند. هوش مصنوعی در IoT در حال حاضر به طور گسترده ای برای تلفیق حسگر، پردازش رویداد و محلی سازی استفاده می شود. در ادامه بخش ها، ما بر روی تحقیق و تلاش های توسعه ی تلفیق داده مبتنی بر هوش مصنوعی تمرکز خواهیم کرد.

## 1) یادگیری ماشین تحت نظارت

یادگیری ماشین تحت نظارت (SML) روشی است که الگوریتم ها از مجموعه ای از قوانین شناخته شده به عنوان داده های آموزشی یاد می گیرند. تعدادی از متغیرهای مستقل که توسط یک متغیر پاسخ یا متغیر وابسته ارائه می شوند، که اغلب برچسب نامیده می شوند. پس از آموزش الگوریتم SML که همچنین می تواند به عنوان طبقه بند نامیده شود، پیش بینی متغیرهای پاسخ از مجموعه داده شده متغیر مستقل [59] انجام می شود. تکنیک های مختلف از اتصال چیزهای هوشمند و به دست آوردن بینش کلیدی از داده های حس شده از IoT در [60] - [63] مورد بحث قرار می گیرد. این مطالعات همچنین مکانیسم هایی برای ایجاد محیط های هوشمند برای بهره گیری از مزایای یادگیری ماشین در IoT را توضیح می دهد.

کاربرد عملی تلفیق داده براساس رویکرد Bayesian در سناریو بسیار محدود شده که مقدار داده زیادی در آن دخیل است. این به خاطر این واقعیت است که اطلاعات ذخیره شده در قالب نمونه های کاملاً قابل توجهی قرار دارد، بنابراین این روش تلفیق داده ها برای استفاده در زمان واقعی که در آن مرکز های فیوز دارای پهنای باند محدود هستند، غیر عملی است. در [64]، روش تلفیق داده مبتنی بر SVM برای غلبه بر تأثیر منفی مجموعه داده های بزرگ بر روی رویکرد Bayesian پیشنهاد شده است. نظریه یادگیری آماری که توسط SVM مورد استفاده قرار می گیرد، پشتیبانی از فشرده سازی اطلاعات را با کمک نمایه ها بر اساس هسته بهینه ای پشتیبانی می کند.

سنجش از دور، یکی از حوزه های کاربردی مهم در آینده ای نزدیک است که از زیرساخت های IOT و قابلیت های سنجش آن بهره می گیرد. درک بهتر این سایت بوسیله تلفیق اطلاعات از چندین سنسور به دست می آید. در تلاشهای تحقیق و توسعه تلفیق [65]، الگوریتم تلفیق داده مبتنی بر SVM برای جمع آوری داده چند طیفی و پانکروماتیک برای سنجش از راه دور شهر شکسین، چین پیشنهاد شده است. فرمول بندی عمومی ریاضی [65] در زیر آورده شده است.

$$(x_1, y_1), \dots, (x_\lambda, y_\lambda), x \in R^n, \quad \{y \in +1, -1\} \quad (2)$$

Hyperplane به صورت زیر است:

$$(w \cdot x) - b = 0 \quad (3)$$

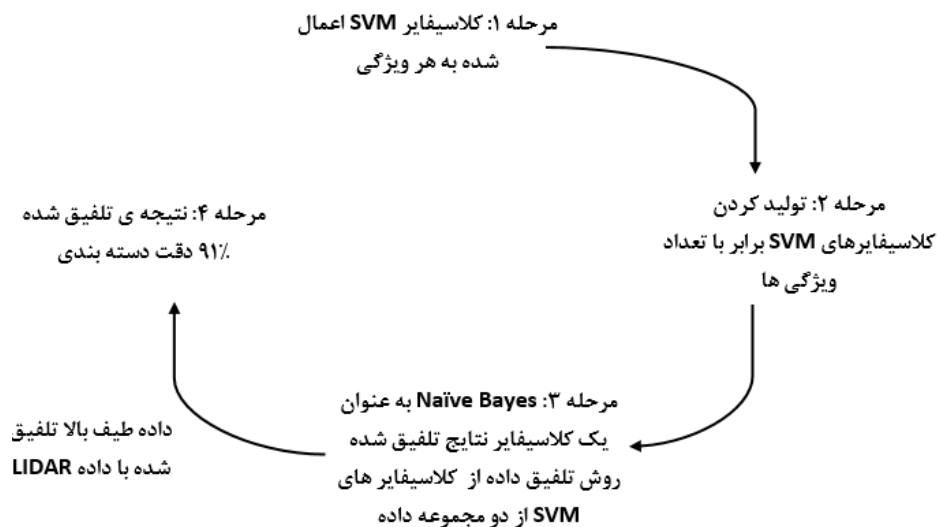
طبقه بندی را میتوان به صورت زیر نشان داد:

$$(w \cdot x) - b \geq 1, \quad \text{if } y_i = +1 \quad (4)$$

$$(w \cdot x) - b \leq -1, \quad \text{if } y_i = -1 \quad (5)$$

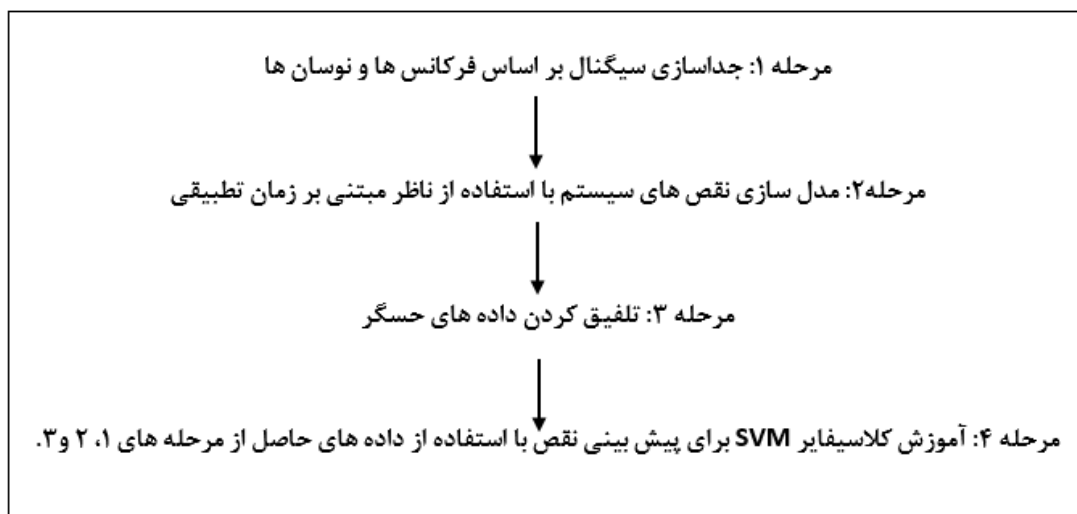
از آنجا که تعداد نمونه توسط  $\lambda$  مشخص می شود، ابعاد داده های ورودی با  $n$  مشخص می شود، " " حاصل ضرب است و  $w$  جهت نرمال hyperplane عملکرد تابع Euclidean است. علاوه بر این خروجی دقت طبقه بندی را 76.5 درصد برای [65] نشان می دهد. به طور مشابه یک کار جدیدتر [66]، که تلفیق های داده طیف بالا و LIDAR با استفاده از رویکرد تلفیق داده تلفیقی انجام شده است. تلفیقی از SVM و کلاسیفایر های Naive Bayes برای انجام تلفیق داده ها در دو مرحله در شکل 6 استفاده می شود. هنگامی که داده های طیف بالا و LIDAR به طور جداگانه در کلاسیفایر پردازش می شوند، دقت 88٪ و 58٪ تولید می کنند. با این حال، نتایج تلفیق داده ها در 91٪ دقت طبقه بندی در [66] انجام شده است. یکی دیگر از روش های تلفیق داده گرا برای داده های طیف بالا و LIDAR در [67] پیشنهاد شده است که از الگوریتم یادگیری ماشین جنگل تصادفی استفاده می کند.

همانطور که IoT نشان می دهد وعده داده است که زندگی روزمره ما را با استفاده از برنامه های مهم ایمنی مانند سیستم های مراقبت بهداشتی هوشمند، سیستم های هشدار دهنده سونامی و فلوت و سیستم های مدیریت خودرو هوشمند، ایمن تر و موفق تر کند. این سیستم های IoT فیزیکی سایبر ما را از موقعیت های ناخواسته و خطرناکی که می تواند ناشی از راه اندازی یک رویداد خاص باشد، آگاه می کند. به عنوان مثال نشت ترمز فلزی در خودرو، این رویداد می تواند یک رویداد خطرناک پس از شکست ناشی از خرابی که باعث آسیب به راننده و مسافران و حتی مرگ شود.



شکل 6: تصویر روش تلفیق داده تلفیقی پیشنهاد شده در [66]

تشخیص نقص نقش مهمی در این جنبه دارد که به طور انتقادی در [68] مورد تجزیه و تحلیل قرار گرفته است. یکی از این کارهایی که در مورد مساله تشخیص نقص برای موتور با استفاده از چهار مرحله تلفیق داده چند سنسور با استفاده از SVM مطرح شده است، در [69] معرفی شده است که بیشتر در شکل 7 توضیح داده شده است. چندین روش تلفیق داده چند حسگر مبتنی بر یادگیری ماشین در [70]، مانند K-نزدیکترین همسایه (KNN)، تجزیه و تحلیل خطی و زاویه سنجی مورد ارزیابی قرار گرفته است. تلفیق داده ها در سطح ویژگی انجام می شود. علاوه بر این، این روش های تلفیق داده ها بر روی مساله مدیریت ترافیک در معماری تلفیق داده های چند سنسور آزمایش می شوند.



شکل 7: چهار گام تلفیق داده مبتنی بر SVM، [69]

## 2) شبکه های عصبی مصنوعی

شبکه های عصبی مصنوعی (ANNs) توانایی فوق العاده ای برای به دست آوردن مفهوم از داده های پیچیده و مبهم دارند. آنها می توانند الگوها را استخراج و روند جدیدی را در مجموعه های بسیار پیچیده داده ها پیدا کنند. آنها از یادگیری تطبیقی با خود سازماندهی در یک محیط زمان واقعی پشتیبانی می کنند و می توانند به درجه بالایی از تحمل خطا دست یابند. به طور گسترده تر، شبکه های عصبی اساسا برای تلفیق داده ها هستند، زیرا داده دقیق تر و مختصر با آموزش و یادگیری پیچیده می سازند. تلفیق داده های گذشته با استفاده از آموزش و تست شبکه های عصبی مصنوعی واقعا در پیش بینی مقادیر بسیار دقیق، مفید هستند [71].

آخرین الگوریتم های پیش بینی زمان سفر عموما نیاز به حجم زیادی از داده ها برای شناسایی پارامترهای الگوریتمیک دارند. به همین دلیل، این الگوریتم ها کمتر مقرون به صرفه هستند و بسیار زمانبرند. برای حل این مشکل، الگوریتم مبتنی بر شبکه های عصبی مصنوعی برای محاسبه زمان سفر در [72] پیشنهاد شده اند. این یک روش پویا برای پیش بینی سرعت و پیاده سازی داده های تلفیقی برای تلفیق داده حسگر های سرعت در یک لینک بزرگراه است. مجموعه ای دیگر از دو الگوریتم که آینده بدست می آیند الگوریتم های روش محاسبه انتگرال سرعت زمان سفر (SITCM) و گسسته سازی فضایی محاسبه زمان سفر (SDTCM) هستند. این دو الگوریتم در پیش بینی زمان سفر کاربردی تر هستند. خطای میانگین پیش بینی کمتر از 10٪ است. بنابراین، می تواند الزامات استفاده در زمینه را برآورده کند. علاوه بر این، این الگوریتم ها برای پیاده سازی ساده و آسان هستند. با این حال، SITCM به مراتب هموار تر از SDTCM عمل می کند، در حالی که SITCM نرخ دقت کمتری نسبت به SDTCM دارد.

یک حوزه مهم جایی که تلفیق داده توسط شبکه های عصبی مصنوعی به طور موثر مورد استفاده قرار می گیرد، پیش بینی سرعت باد است. یکی از این کارها از شبکه های عصبی مصنوعی برای پیش بینی سرعت باد با پیش بینی فعال شده توسط تلفیق سرعت باد گذشته با یکدیگر استفاده می کند [73]. با توجه به ماهیت بسیار پیچیده داده های سرعت باد، این کار سخت برای پیش بینی مقدار های آینده است. دقت در برآورد خروجی نیروی باد بسیار مهم است، زیرا تولید انرژی باد، متناسب با مکعب سرعت باد است. این مطالعه یک راه حل ارائه می دهد که روند های

مختلف سرعت باد آینده را پیش بینی می کند. این کار با پیشنهاد الگوریتم تلفیق داده با استفاده از چندین شبکه عصبی انجام می شود. شبکه های عصبی مصنوعی توسط مجموعه داده های باد آموزش داده شده و آزمایش می شوند. این الگوریتم حداقل میانگین مطلق خطا را دارند که قابل توجه هستند. برای پیش بینی الگوهای ماهانه سرعت باد، شبکه های انتشار بازگشت به جلو (FFBP) دو لایه با 6 لایه میانه و 30 نورون در لایه های خروجی استفاده می شود. برای پیش بینی ساعتی سرعت باد، شبکه FFBP با 30 و 12 نورون در [73] استفاده می شود. یکی دیگر از مطالعات [74]، از شبکه های عصبی برای پیش بینی سرعت باد استفاده می کند؛ با این حال، تفاوت اصلی بین [73] و [74] این است که در [74] از شبکه عصبی دو لایه ای استفاده شده است. با این وجود، در [73] از چندین شبکه عصبی استفاده می شود که به معنی افزایش پیچیدگی است. بعضی از روش های پیش بینی نیروی باد در کوتاه مدت نیز در [75] و [76] پیشنهاد شده است.

بیشتر شبکه های عصبی مصنوعی از طریق تلفیق داده های آموزشی یاد می گیرند، همچنین در زمینه های دیگر مانند محلی سازی نیز استفاده می شوند. دقت محلی سازی داخلی و خدمات جهت یابی بر اساس شاخص های قدرت سیگنال رادیویی (RSSI) و فرکانس رادیویی که در سمت پایین تر است مشخص می شود. این به دلیل متغیر مقدار های RSSI است. در مقاله [77] روش شبکه های چند عصبی برای حل مسئله RSSI استفاده می شود. محلی سازی داخلی بلوتوث با استفاده از چندین شبکه عصبی با آموزش و استفاده از شبکه های عصبی براساس جهت گیری کاربر به دست می آید. یک معماری بلوتوث بسیار دقیق و با هزینه ای مناسب برای جهت یابی داخلی امکان پذیر است.

### (3) منطق فازی

لطفی زاده در سال 1965 نخستین اصطلاح 'فازی' را به عنوان نظریه مجموعه فازی پیشنهاد کرد و بعدها به ریاضیات فازی آن معروف شد. منطق فازی برای مدیریت مفهوم حقیقت جزئی استفاده می شود که بین کاملاً درست و نادرست محاسبه می شود. منطق فازی کاربردهای متعددی در زندگی واقعی دارد که در [78] مورد بحث قرار گرفته است، مانند هوش مصنوعی، کنترل، نظارت محیط، بازی، خودکارسازی وسایل الکترونیکی، تشخیص پزشکی و تشخیص

رویداد و غیره. یک تکنیک تشخیص رویداد با استفاده از شاخه ی منطقی فازی از هوش مصنوعی در شبکه های حسگر بی سیم خوشه ای (WSN) برای تلفیق داده ها در [79] ارائه شده است. یکی از مهمترین کاربردهای WSN ها نظارت محیط است که می تواند در چشم انداز وسیع جوامع هوشمند مبتنی بر IoT دیده شود. سیستمی در [79] شامل سنسورهای متعددی برای رطوبت، دی اکسید کربن و اندازه گیری دما است. یک سیستم مبتنی بر قوانین فازی، تلفیق سیگنالها را انجام می دهد. سیگنال ها در سر خوشه جمع می شوند و سپس فرایند تلفیق اتفاق می افتد. مزیت تلفیق چند سنسور این است که دقت و قابلیت اطمینان را افزایش می دهد. الزامات مهم جوامع مدرن امروزه سیستم های پارکینگ هوشمند خودرو است که می تواند از زیرساخت های IoT برای درد سر مدیریت پارکینگ اختیاری کمک بگیرد که در [80] و [81] ارائه شده است. الگوریتم تلفیق داده مبتنی بر منطق فازی در [82] برای نظارت بر فضای پارکینگ پیشنهاد شده است. برای این منظور از سنسورهای مغناطیسی استفاده شده و احتمال اشغال از فضای پارکینگ مربوطه محاسبه می شود. الگوریتم تلفیق داده ها در [82] اطلاعات را با دقت بالا از نظارت هدف ها جمع آوری می کند که نتیجه آن تصمیم گیری صحیح و با قابلیت ضد تداخل است. با تمرکز بر ایجادشان جوامع ما با استفاده از زیرساخت های IoT ایمن تر می شود، در [83] یک رویکرد تشخیص و کنترل هوشمند آتش ارائه شده است که با استفاده از تلفیق داده مبتنی بر فازی ارائه شده است.

استفاده از منطق فازی در حال افزایش محبوبیت برای سیستم های ردیابی با توجه به مزایای متعدد ذکر شده در بالا است. به خصوص برای مسائل ردیابی در PDA و JPDA، به عنوان تعدادی از اشیاء به طور موازی افزایش می یابد و افزایش نمایی در زمان محاسبات وجود دارد. با این حال، عقل سلیم در منطق فازی برای غلبه بر این استفاده می شود [84]. چندین روش مبتنی بر فازی برای تلفیق وجود دارد، مانند الگوریتم همبستگی استنتاج فازی [85]، الگوریتم انجمن آستانه دو مسیر فازی و الگوریتم میانگین خوشه بندی فازی (FCMA) [86]، [87].

در [87] یک راه حل تلفیق داده برای مسائل ردیابی مسیر در چند سنسور و چند هدف با ویژگی های چندگانه پیشنهاد شده است. الگوریتم میانگین خوشه بندی فازی (FCMA) برای به حداقل رساندن تعداد مسیر ها و با استفاده از درجه عضویت هر هدف برای تعیین مسیر های تکراری همبسته اظهار دارد. برای اهداف همبستگی و تلفیق، داده

های سنسور با استفاده از وضوح حسگر استفاده می شود؛ همین کار برای شناسایی دقیق ترین سنسورها در سیستم انجام شده است. شبیه سازی مونت کارلو برای نشان دادن این است که طرح جدید، پیچیدگی محاسباتی را به حداقل می رساند و کارایی را با توجه به خوشه بندی اقلیدسی و روش حداقل خطای میانگین مربع Bayesian افزایش دهد. مقاله [88] بر روی سیستم های ردیابی چند سنسوره و چند هدفه توزیع شده مانند موارد مورد بحث در [87] تمرکز دارد.

FCMA نسبت به نویز قوی است؛ با این حال، انتخاب پارامترهای مناسب مشکل است و یک چالش برای فراهم کردن عملکرد بهینه است [87]. کار در [88] بر اساس همبستگی مسیر است. همبستگی بسیار مهم است، زیرا مشخص نیست که آیا مسیر های خاص از حسگرهای مجزا یک هدف مشابه را نشان می دهند. هر گره حسگر اندازه گیری های نویزی را برای موقعیت های متمایز اهداف انجام می دهد. پس از محاسبات هیئت مدیره، داده ها به مرکز تلفیق داده منتقل می شوند. این الگوریتم همبستگی مسیر مبتنی بر فازی دو آستانه یک آستانه تطبیقی را استفاده می کند و خطاهای ارتباطی کمتری را ایجاد می کند که باعث عملکرد بهتر می شود.

یک کلاس تلفیقی از طرح های فازی و فیلتر کلمن نیز به منظور تلفیق داده مناسب استفاده می شود. یکی از این مطالعات [89] فیلتر کالمن تطبیقی مبتنی بر منطق فازی (FLAKF) را برای تلفیق داده های چند حسگر پیشنهاد می کند. اندازه گیری برای نویز ماتریس کوواریانس به صورت انطباقی برای هر FLAKF محلی برای تعریف آمار پروفایل نویز برای داده ها در آینده تنظیم شده است. یک سیستم استنتاج فازی برای سازگاری استفاده می شود. استفاده از منطق فازی کمک می کند تا داده غیر دقیق را اداره کند. در حالی که میزان کردن فیلترهای کالمن، ماتریس کوواریانس را برای بدست آوردن تخمین های دقیق تر تنظیم می کند. بدین ترتیب، روش تلفیقی بدست آمده دقیق تر است. به طور مشابه، ادغام منطق فازی و فیلتر کلمن برای تلفیق داده ها در [90] برای هدایت وسایل نقلیه خودمختار طراحی شده است.

با توجه به تمام ادبیات مورد بررسی در بخش III ب، می توان گفت که رویکردهای تلفیق داده مبتنی بر هوش مصنوعی در استفاده عملی محبوبیت زیادی به دست می آورند. این تکنیک ها دقیق تر از روش های دیگر تلفیق داده ها می



باشند و می توانند برای طیف گسترده ای از مسائل تلفیق استفاده شوند. منطق فازی می تواند به طور قطعی عدم اطمینان داده های حسگر را کنترل کند. شبکه های عصبی مصنوعی برای سیستم های غیرخطی ایده آل هستند و اکتشافات الگوی پیچیده ای قابل دستیابی هستند. با این وجود، شبکه های عصبی نتوانستند توضیح دهند که چگونه آنها تلفیق از مقادیر حسگر ورودی را یاد می گیرند و رویکردهای هوش مصنوعی به صورت محاسباتی پیچیده هستند. علاوه بر این، برخی از چالش ها باقی می ماند، زیرا پیش بینی صحیح 100٪ هنوز هم از طریق تلفیق قابل دستیابی نیست.

#### 4) فرصت ها و چالش ها

در زمان های اخیر، استفاده از روش های تلفیق داده مبتنی بر هوش مصنوعی در حال افزایش است. روش های یادگیری نظارت شده مانند SVM می تواند داده های با ابعاد بزرگ را مدیریت کند. شبکه های عصبی مصنوعی و منطق فازی اجازه می دهد تا ورودی های غیردقیق و متناقض را داشته باشند و از پیچیدگی های دلخواه برخوردار باشند. سیستم های تلفیق داده منطقی فازی برای مدیریت ورودی های غیر دقیق حسگر و عدم اطمینان در مقایسه با روش های احتمالی بهتر است. رویکرد تلفیق داده شبکه های عصبی مصنوعی یکی از بهترین سیستم های تلفیق داده غیر خطی است و اکتشافات الگوی پیچیده امکان پذیر است. سیستم های تلفیق هوش مصنوعی در مقایسه با سیستم های احتمالی بسیار دقیق هستند. با این حال، آنها پیچیده و به صورت محاسباتی گران هستند.

#### C. الگوریتم های تلفیق داده مبتنی بر نظریه شواهد

Dempster's، نخستین نظریه باور در تفکر را پیشنهاد کرد. بعدها، شافر ریاضیات آن را بر اساس یک استدلال مبتنی بر شواهد رسمی کرد [91]. برای مدیریت عدم دقت و عدم اطمینان، نظریه ی توابع باور، یک انتخاب دسته جمعی برای تفکر کردن است. نظریه Dempster-Shafer (DST)، به عنوان تعمیم نظریه احتمال ذهنی Bayesian (BPT) مشاهده می شود. در DST و BPT چند تفاوت وجود دارد. اولاً، BPT وضعیت ناشناخته ای ندارد، در حالی که

در DST، وضعیت ناشناخته می تواند حالت دانش برای ما باشد. ثانیاً، انتساب های قبلی BPT، اگر چه DST از قسمتهای عمده ای برای تخصیص معنی دار برای همه حالات استفاده می کند. ثالثاً، محاسبات کمتری در BPT نسبت به DST مورد نیاز است [24].

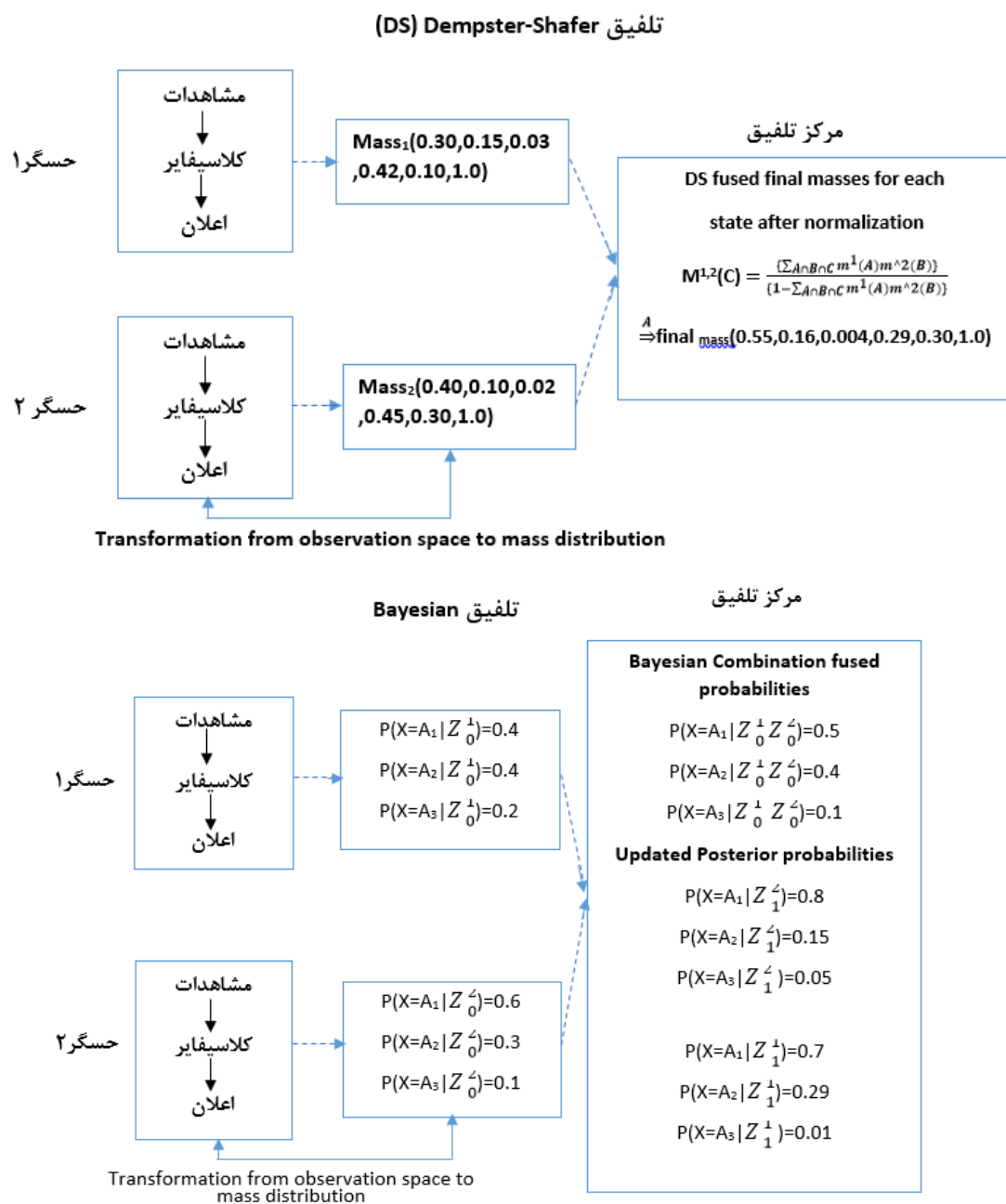
امروزه، رویکرد تلفیق مبتنی بر DST [92] به طور گسترده ای در استفاده است. DST یک روش بسیار کارآمد برای استخراج ویژگی در محیط چند حسگر است [92]. توانایی DST برای اداره کردن داده نامعین و ناقص، آن را در سیستم چند سنسور قابل استفاده می سازد [93]. سیستم های تلفیق مبتنی بر DST در سطح جهانی قابل استفاده هستند و به دلیل تعیین پارامترهای تجربی و استدلالی، از تصاویر ماهواره ای مستقل هستند [94]. فیلد DST به طور گسترده مورد مطالعه قرار گرفته و تحقیقات گسترده روی آن انجام شده است. تشخیص و ردیابی نور (LIDAR) یک روش سنجش از راه دور است که به تعیین فاصله از طریق روشن کردن یک هدف با یک لیزر و بررسی بازتاب نور کمک می کند. LIDAR یک حوزه مهم است در جایی که می تواند به طور قابل ملاحظه ای از تلفیق DST استفاده شود. با این حال، حداقل توجه به حوزه های زیر ضروری است: مشتق شیب از داده های LIDAR، ارزیابی کمی از کلاس ها و تفسیر احتمال پذیری، حداکثر احتمال، پشتیبانی، نقشه های برخورد و عدم قطعیت. برای مقابله با مسائل ذکر شده ی بالا، مقاله [95] یک برنامه DST داده محور را برای تلفیق داده های چند سنسور برای دستیابی به استخراج ویژگی پوشش زمین پیشنهاد کرد. این تکنیک هر دو کارایی هزینه و زمان را دارد.

یک تکنیک تلفیق داده براساس DST برای تلفیق نشانه های نقشه دیجیتال جاده ای و علائم جاده ای شناسایی شده است که در [96] ارائه شده است. یک نوع دیگر از کارهای مشابه [97] که با استفاده از ادغام داده های نقشه و تصویر ویدئویی از یک سیستم تلفیق برای تشخیص وجود علامت ترافیک استفاده می کند، در حالی که [98] از DST و حالت زمینه برای اختصاص دادن اولویت برای استفاده از نقشه های دیجیتال یا سیستم مبتنی بر دوربین استفاده می کند. بنابراین، هیچ کدام از این کارها تشخیص مثبت کاذب سیستم های دوربین مربوطه را نشان نمی دهد. علاوه بر این، شرایط آب و هوایی بد و شرایط نور و محدودیت سرعت جاده ها می تواند عملکرد سیستم تلفیق داده را کاهش دهد و به ندرت مورد توجه قرار می گیرد. همچنین، با توجه به سرعت محاسبات این روش ها، هیچ داده کمی وجود

ندارد. برای مدیریت این مسائل مربوط به اعتماد، در مقاله [99]، پنج الگوریتم تلفیق برای مقابله با مسائل مربوط به تلفیق [96] - [98] مورد بحث قرار گرفته است. چهارتا از آنها بر اساس اولویت اطلاعات هستند و الگوریتم پنجم بر پایه DST کلاسیک است. تمام الگوریتم های پنجگانه کارایی بیشتری داشتند و خروجی های دقیق تری تولید می کردند.

## 1) فرصت ها و چالش ها

ما تجزیه و تحلیل قسمت بالا از ادبیات DST را بر اساس دو زمینه اصلی نتیجه گرفتیم: مزایا و چالش های باز برای DST. DST مسائل پیچیده در تعریف های پیشین را با توجه به عدم قطعیت از بین می برد؛ اینها بر این اساس می توانند مورد به مورد نادیده گرفته شوند. DST یک رویکرد ساده است که به هم جور کردن شواهد را با سطوح انتزاعی متنوع توصیف می کند و شواهد بیشتری را به هم پیوند می دهد. پیچیدگی در محاسبات یک مشکل در DST است. علاوه بر این، نمی توان آن را یک نظریه تصمیم ثابت مطرح شده مانند پذیرش جهانی BPT دانست. در ابتدا، BPT متشکل از دو حالت است که هر دو یا رویداد یا بدون رویداد است؛ با این حال، یک حالت ناشناخته گم شده است. DST شامل سومین حالت، حالت ناشناخته نامیده می شود. تعریف توابع انبار DST (mass)، مهمترین و سخت ترین وظیفه در پیاده سازی آن است. DST به طور گسترده در علوم کامپیوتر و هوش مصنوعی مورد مطالعه قرار گرفته است، اما هرگز توسط متخصصان به طور کامل پذیرفته نشده است [100].



شکل 8. تلفیق Bayesian و تلفیق (DS) Dempster-Shafer.

در شکل 8، ما اتصال [42] و [101] را به شکل تجسم BPT و DST، که توضیح مختصری از تلفیق را تشکیل می دهد، ارائه داده ایم.

#### 4. روش های تلفیق داده برای محیط های خاص IoT

چندین چالش مرتبط با حسگر داده وجود دارد. برخی از چالش های کلی عبارتند از:

- محیط توزیع شده WSN ها که زیر مجموعه ای از IoT هستند.
- ماهیت بسیار ناهمگن IoT به علت دستگاه ها و داده های ناهمگن مختلف.
- مسائل غیر خطی و ردیابی مثل: ردیابی چند هدفه (MTT)، هزینه غیر موثر، کاهش خطا، مسائل ناهمگام و مسیر به مسیر (T2T).

در این بخش، ادبیات را بر اساس چالش های ذکر شده راجع به روش های تلفیق داده ها بررسی و طبقه بندی می کنیم. این بخش به چهار قسمت تقسیم می شود: چالش های محیط توزیع شده، چالش های محیط ناهمگن، چالش های محیط غیر خطی و چالش های محیط ردیابی.

#### A. محیط توزیع شده

به لحاظ نظری، تلفیق داده توزیع شده، یک مفهوم نسبتاً توسعه یافته در چند دهه گذشته است و در عمق مورد مطالعه قرار گرفته است [102]. با این حال، پیاده سازی عملی این تکنیک ها هنوز یک وظیفه چالش برانگیز در محیط IoT است. الگوریتم های توزیع شده هیچ کنترل متمرکزی ندارند. آنها همچنین می توانند به عنوان زیر شاخه های الگوریتم های موازی اشاره شوند. عناصر اصلی یک سیستم تلفیق اطلاعات توزیع شده در IoT سنسورها و پردازنده ها هستند. سنسورها مسئول تولید داده ها با مشاهده محیط عملیاتی هستند. پردازنده ها مسئول تلفیق داده ها هستند. در این بخش، الگوریتم های تلفیق داده توزیع شده را در WSN ها بررسی و طبقه بندی می کنیم که می تواند به عنوان زیرمجموعه های IoT دیده شود. در محیط توزیع شده، الگوریتم تخمین اصلی، الگوریتم Kalman Filter (KF) توزیع شده است و بسیار مقیاس پذیر است. در مسائل کاربردی مختلف، الگوریتم های مبتنی بر KF نقش قابل توجهی ایفا می کنند. KF می تواند یک الگوریتم فیلترینگ تصحیح پیش بینی باشد که توسط تکامل یا انتشار حالت و به روز رسانی داده ها به دست می آید. ما می توانیم KF را به عنوان یک الگوریتم تلفیق Bayesian بپذیریم، همانطور که

در [101] ثابت شده است. حوزه های مختلفی که KF توزیع شده پیاده سازی شده است، شامل نظارت بر آب و هوا و محیط زیست، نظارت (مانیتورینگ)، حوزه های ردیابی و پزشکی است.

بیشتر الگوریتم های در این کلاس ماهیتشان ایستا هستند. هر چند، توپولوژی شبکه واقعی محیط پویا است. این تغییر در توپولوژی شبکه ممکن است ناشی از شکست گره باشد که نتیجه آن کاهش انرژی است. مطالعه [103] با چنین نوع شبکه ای سروکار دارد. کار بر اساس شبیه سازی 200 سنسور به صورت تصادفی در یک محیط پویا توزیع می شود. فیلترینگ مرکزی کالمن (CKF) برای ارزیابی KF توزیع شده در کار استفاده می شود. الگوریتم پیشنهادی بسیار نزدیک به CKF رفتار می کند. سنسور مبتنی بر الگوریتم توزیع شده (KF) می تواند تا شش سنسور همسایه اجرا شود. پیاده سازی KF به روش توزیع شده، مزیت به دست آوردن هزینه های محاسباتی ماتریس در مقایسه با CKF را نگه می دارد. مزیت دیگر KF توزیع شده این است که مقیاس پذیری ساده تر است، که در یک محیط پویا مانند IoT مورد نیاز است.

در IoT، تأثیر انرژی نقش مهمی ایفا می کند، زیرا ما صدها سنسور را با یکدیگر ادغام می کنیم. این سیستم باید دارای انرژی کارامدی باشد. در غیر اینصورت، هزینه مصرف انرژی توسط سنسورها زیاد است. یک رویکرد ذرات مبتنی بر فاجعه (CBPA) توسط استقرار تصادفی گره ها بر اساس خوشه های استاتیک با جستجوی کوکو در WSN توزیع شده استفاده می شود. پس از انتخاب سرخوشه ها، داده ها جمع آوری، جمع شده و ارسال می شوند. یک الگوریتم رویکرد ذرات تعمیم یافته استفاده شده و داده ها به ایستگاه پایه فرستاده می شوند. مشکل مصرف انرژی شبکه به دینامیک و سینماتیک تبدیل می شود که توسط الگوریتم مدل ذرات تعمیم یافته (GPMA) تکمیل شده است. نتیجه این رویکرد تلفیقی، منجر به کاهش توان مصرفی انرژی می شود. همچنین با توجه به الگوریتم زیر بهینه، یک درجه بالایی از سازگاری با سطح پایین پیچیدگی [104] نیز به دست می آید. با هدف مشابه، کار [105] با هدف کاهش هزینه های انرژی مرتبط با فرآیند تلفیق داده ها توزیع شده است. این با ارائه یک راه حل تقریبی به نام P2lace به دست می آید. این در دو مرحله انجام می شود؛ در مرحله اول، یک پارتیشن گراف وظیفه اتفاق می افتد، و در مرحله دوم، قرار دادن گراف وظیفه اتفاق می افتد.

راه دیگری برای کاهش مصرف انرژی در سیستم های تلفیق توزیع شده WSN ها این است که برخی از گره های حسگر را در یک حالت خواب در بعضی از زمان ها قرار دهید، در حالی که گره های حسگر باقی مانده فعال هستند. یک الگوریتم تلفیق داده توزیع شده برای گره های حسگر در [106] پیشنهاد شده است. الگوریتم زمان هدایت شده است، که تجمیع داده شبکه را انجام می دهد و با برنامه ریزی گره ها و تخمینی دسته ای انجام می شود. این به دو دلیل انجام می شود: اول، برای رعایت شرایط چرخه اکتساب، و دوم، برای کاهش زمان گره های حسگر در حالت کار است. به طور موازی، تلفیق داده ها براساس برآورد دسته ای معرفی می شود. خروجی تولید شده توسط الگوریتم نشان دهنده کاهش مصرف انرژی شبکه و قابلیت اطمینان بهبود یافته است. این سیستم تلفیق را می توان در برنامه های نظارت سلامت به طور کارآمد استفاده کرد.

مطالعه مشابهی به [106] در [107] پیشنهاد شده است که بر اساس شکل گیری خوشه است و هزینه انرژی تلفیق داده را کاهش می دهد. با این حال، تفاوت عمده در مشارکت این دو کار این است که تکنیک تلفیق داده [107] بر مبنای انتخاب مسیر چندگانه با یک نسبت تحویل بسته در سمت بالاتر است. چندین راه برای دستیابی به کارایی در انرژی در 108 پیشنهاد شده است. دو الگوریتم برنامه ریزی انرژی برای برآورد پیشرونده پیشنهاد شده است. نویسندگان همچنین هزینه انرژی برای برآورد اجماع را محاسبه کردند؛ این تکنیک برای تلفیق توزیع شده در مدل حسگر های چند هاپ بر اساس شبکه های نظیر به نظیر می باشد. مجموعه ای از الگوریتم ها از اصول زیر برای دستیابی به کارایی انرژی تکامل یافته است [108]:

- Multi-hop ها انرژی کارآمد بسیاری دارند زیرا آنها از انتقال داده های مسافت طولانی خودداری می کنند.
- با تلفیق داده های پیشرفته که hop ها از طریق سنسورها پدید می آید، مصرف انرژی کم است.
- توسط انتقال انرژی پیش تعیین شده انرژی کاهش می یابد. این با کمک دانش قبلی از تمام اطلاعات کانال حالت و درخت مسیریابی انجام می شود.

مقیاس پذیری یک کار چالش برانگیز در محیط های توزیع شده مانند IoT است که همچنین ناهمگن و پویا است. در یک محیط مانند IoT، یک مسئله مهم این است که به طور ناگهانی تعدادی از سنسورها می توانند بیدار شوند و

چندین گره را در WSN ها اضافه کنند. الگوریتم تلفیق داده باید برای مقابله با این نوع موقعیت ها مفید باشد. برای حل این مسئله، [109] چندین الگوریتم زیر بهینه را مورد بحث و تحلیل قرار می دهد. این الگوریتم ها عبارتند از:

## 1) فیلتر کانال

این رویکرد ساده تلفیق داده است. فقط ابتدا داده های زائد مرتب شده در نظر گرفته می شوند. هر کانال یک جفت عامل دارد، یک عامل انتقال دهنده و یک عامل دریافت کننده است. اطلاعات زائد توسط عامل انتقال حذف می شوند. اگر چه، داده های انتقال WSN های ad-hoc، گاهی اوقات به انتهای دیگر نمی رسند. بنابراین، عامل دریافت می تواند وظیفه عامل انتقال در WSN های ad hoc پویا را انجام دهد. معادله تلفیق فیلتر کانال به صورت زیر است:

$$p(x) = \frac{p_1(x) p_2(x) / \bar{p}(x)}{\int p_1(x) p_2(x) / \bar{p}(x) dx} \quad (6)$$

در معادله (6)،  $p_1(x)$  و  $p_2(x)$  احتمال تلفیق تابع چگالی و  $\neg p(x)$  تابع چگالی قبلی دریافت شده است. مزیت این الگوریتم این است که نیازی به نگه داشتن حجم زیادی از تاریخ فعالیت های گذشته وجود ندارد. اگر چه یکی از معایب این است که در طول فیلترینگ، اطلاعات وابسته حذف می شود. با این حال، اگر زمان بین پردازش فعلی و هنگامی که افزونگی رخ داده، بیش از حد طولانی باشد، این اثر می تواند به حداقل برسد.

## 2) تلفیق Naïve

این یکی از ساده ترین تکنیک های تلفیق داده است. پیش بینی شده است که وابستگی بین توابع چگالی کم است؛ با این حال، این تکنیک غیر قابل اعتماد است. با توجه به نبود اطلاعات گذشته، بیش از حد اطمینان می تواند رخ دهد. معادله تلفیق naïve می تواند به صورت زیر نوشته شود:

$$p(x) = \frac{p_1(x) p_2(x)}{\int p_1(x) p_2(x) dx} \quad (7)$$



### 3 تلفیق Chernoff

در توزیع وابستگی ناشناخته، روش Chernoff می تواند مورد استفاده قرار گیرد. از لحاظ نظریه، اسکن دو تابع چگالی دلخواه می بایست با استفاده از تلفیق Chernoff در یک روش خطی لگاریتمی تلفیق شوند. با این حال، چگالی تلفیق شده ممکن است منحرف شود. یک نقطه ضعف دیگر این است که محاسبات گسترده ای لازم است. معادله Chernoff به صورت زیر است:

$$p(x) = \frac{p_1^w(x) p_2^{1-w}(x)}{\int p_1(x) p_2(x) / \bar{p}(x) dx} \quad (8)$$

جایی که:  $w$  عضو  $[1,0]$

اگر چه مطالعات متعددی براساس الگوریتم های تلفیق داده در یک محیط توزیع شده وجود دارد، هنوز هم چالش های متعددی وجود دارد. همان طور که در [109] بیان شده، فرایند تلفیق داده ها به صورت ریاضی با نظریه مجموعه ای در معادله داده شده توصیف می شود:

$$\Phi(.|I_i) = \Phi\left(.|\bigcup_{i=1}^n I_i\right) \frac{1}{c} \prod_{i=1}^n S_i^{(-1)^{i-1}} \quad (9)$$

جایی که  $S_i$  تلفیقات احتمالی رویداد  $i$  است، تقسیم متناوب و ضرب احتمالات مشترک از (6)، از فرم وابستگی های مشروط از اطلاعات به اشتراک گذاشته شده در مجموعه داده ها خلاص شود.

### 4 فرصت ها و چالش ها

در نظریه تلفیق داده توزیع شده، حذف اطلاعات تکراری (دو نسخه ای) ساده است. شناسایی اطلاعات تکراری برای سیستم های تلفیق توزیع شده هنگامی که این نظریه به استفاده عملی اعمال می شود دشوار است. در تلفیق توزیع شده، تشخیص منشاء اطلاعات همبسته از رویدادهای تلفیق گذشته و بدست آوردن مقادیر مجموعه داده های قبلی، دشوار است. IoT بسیار ناهمگن است، با توجه به هر دو جنبه دستگاه ها و داده ها، که باعث می شود تلفیق توزیع شده به چالش کشیده شود. با توجه به ماهیت پویای آن، WSN های توزیع شده اغلب شکل و اندازه خود را تغییر می

دهند، بنابراین مقیاس پذیری شبکه یک مسئله نگرانی در سیستم های تلفیق توزیع شده است. ناهمگونی نیز یک چالش برای IoT است. در بخش بعد، ما تجزیه و تحلیل ادبیات الگوریتم های تلفیق داده ناهمگن را بررسی می کنیم.

## B. محیط ناهمگن

محیط IoT همیشه همگن نیست بنابراین، ناهمگونی در این محیط متنوع از لحاظ هر دو دستگاه و داده ها نادر نیست. یکی از مشکلات بزرگ که سیستم های تلفیق ناهمگن با آن مواجه هستند، به علت فضاهای ویژگی مختلف از مجموعه داده ها است. در سیستم های ناهمگن، مجموعه داده ها به طور کلی در چندین فضای ویژگی نشان داده می شوند. این امر، تجزیه و تحلیل روابط بین داده های مختلف را سخت می کند، حتی زمانی که مجموعه داده ها به صورت معنایی مرتبط با یکدیگر هستند.

به عنوان راه حل مشکل فوق، یک چارچوب تعبیه گراف در [110] برای مقابله با مسائل تراز فضایی در سیستم های تلفیق ناهمگن استفاده می شود. چارچوب پیشنهادی هر مجموعه داده را به یک گراف و انتساب فاصله صفر بین جفت های متناظر تبدیل می کند که در نهایت به یک تک گراف منجر می شود. یک مقیاس چند بعدی غیرمتریک پیشنهاد شده است که از ترتیب رتبه استفاده می کند. مزیت استفاده از رتبه بندی این است که این نوع سیستم تلفیق می تواند هم تراز و تغییر شکل را مدیریت کند. این روش ثابت شده کارآمد و بهتر از روش های eigen-maps لاپلاسین محدود، تجزیه tensor و تجزیه Procrustes است.

یک روش مبتنی بر فیلترینگ Bayesian برای سیستم های ناهمگن در [111] مورد بحث قرار گرفته است؛ مدل فضای حالت برای تعیین تخمین سنسورهای اندازه گیری رادیویی و سرعت استفاده می شود. مسئله ردیابی به چندین محدودیت محلی با تعاملات متقابل با گراف های عامل با عبور پیام تقسیم می شود. در طی هر تکرار، پیام ها با اطلاعات قابل اعتماد، کارآمد عبور داده می شوند. این بین مرحله پیش بینی و اصلاح اتفاق می افتد. برای از بین بردن اثرات انتشار خطا به علت واریانس سنسور سرعت، الگوریتم از روش هموارسازی fixed-lag استفاده می کند که بر داده های گذشته و آینده یک نقطه خاص متکی است. الگوریتم کمتر پیچیده است و دقت بالا را نشان می دهد، بنابراین

تمایل دارد تا بار محاسباتی را به حداقل برساند. چندین طرح اضافی در [112] و [113] برای مدیریت داده های زیست پزشکی ناهمگن با استفاده از تکنیک Bayesian پیشنهاد شده است. الگوریتم های مبتنی بر منطق فازی گوناگونی و الگوریتم های تلفیقی بر اساس تلفیقی از منطق فازی و فیلتر کلمن برای سیستم های حسگرهای ناهمگن بهتر و مؤثرتر هستند که پارامترهای یکسانی با پویایی متنوع و آمار نویز را اندازه گیری می کنند [84].

تلفیق شدن داده از مشاهدات ناهمگن وعده می کند تا روابط چند متغیره پیچیده را در میان مجموعه داده ها پیدا کنند. در تحلیل داده های چند متغیره از دو مجموعه متغیرها برای استخراج ویژگی های همبسته، دو روش اساسی وجود دارد: تحلیل همبستگی استاندارد (CCA) [114]، [115] و رگرسیون حداقل مربعات جزئی (PLS) [116]، [117]. PLS زمانی که دو مجموعه ی متغیرها به یکدیگر وابسته هستند یا یک مجموعه از متغیرها توضیح دیگری را در نگه می دارد، امکان پذیر است [114]. از سوی دیگر، CCA، زمانی که دو مجموعه متغیرها به صورت متقارن به یکدیگر وابسته هستند، مناسب تر است. CCA به طور گسترده ای برای تحلیل آماری چند بعدی روشن مورد مطالعه قرار گرفته است. هدف اصلی CCA تعیین تلفیقات خطی هر متغیر در مجموعه داده های داده شده است؛ این کار زمانی انجام می شود که همبستگی در حداکثر خود در میان تلفیقات خطی باشد.

معایب اصلی CCA این است که خروجی تلفیق داده به طور چشمگیری با مجموعه داده های نویزی تضعیف می شود [114]. در مقاله [118] یک روش جدید به نام الگوریتم حذف نویز خارج از محدوده (NORA) پیشنهاد شده است. NORA مجموعه داده های نویزی را در سیستم های تلفیق داده ناهمگن مدیریت می کند که [114] (گونزالز، 2009) فاقد آن است. NORA برای فیلتر ویژگی ها و نقاط داده ی غیر آموزنده قبل از اجرای CCA استفاده می شود. به طور خاص در [118]، NORA برای پیش پردازش عصب شناسی و تصویربرداری طنین مغناطیسی (MRI) قبل از اجرای CCA برای شناسایی ارتباط بین آنها استفاده می شود. چندین روش دیگر برای مدیریت داده های ناهمگن با استفاده از تحلیل Bayesian [119]، DST [120]، روش یادگیری چند معیار [121] و روش تلفیق داده تلفیقی براساس احتمال و DST در [122] پیشنهاد شده است.

## 1) فرصت ها و چالش ها

ناهمگونی در محیط IoT یک مسئله چالش برانگیز است که در طول سیستم ادغام با توجه به منابع متفاوتی از داده ها حل می شود. این منابع داده ها را نمی توان به عنوان روش های لازم برای تبدیل داده های ناهمگن به فضای همگن تلفیق کرد. مجموعه داده های ناهمگن بیشتر عدم قطعیت را اضافه می کنند. روابط چند متغیره در میان مجموعه داده ها پیچیده است. با این حال، تلفیق شدن داده ها از مشاهدات ناهمگن وعده می کند تا روابط چند متغیره پیچیده را در میان مجموعه داده ها پیدا کند.

## C. محیط غیر خطی

سنجش غیر خطی متغیر زمانی نیز چالش های بزرگی را برای تلفیق داده های چند سنسور موجب می شود. غیر خطی بودن می تواند برآورد های دقیق تری را انجام دهد. در [123]، نشان داده شده است که تخمین بهبود یافته در سیستم های غیر خطی حاد در مرکز تلفیق با کمک تلفیق داده های بهینه در راه اندازی چند سنسور تولید می شود. یک الگوریتم بهینه سازی شده با توجه به شرایط کانال هر سنسور برای تخصیص قدرت گره های حسگر مربوطه پیشنهاد شده است. آن را به صورت پویا برای تخصیص قدرت تمام زیر مجموعه گره های حسگر انجام می دهد. این روش بر اساس یک برنامه semi-definite (SOP) [124] است. بنابراین، با قصد کاهش میانگین مربع مقادیر خطا، بهترین تخمین وضعیت موجود را تضمین می کند.

فیلتر کالمن توسعه یافته (EKF) یکی از جامع ترین الگوریتم های در محیط های ردیابی غیر خطی بوده است [125]، [126]. با این حال، از مقادیر دقیق تر از فیلتر کالمن (KF) بدون افت فشار از EKF با کمک آمار تقریبی به دست آمده اند [127]. در عمل، EKF دارای سه عیب است [128]:

- فیلترهای ناپایدار را می توان به دلیل خطی سازی EKF تولید کرد.
- با استفاده از EKF، خطی سازی می تواند تنها در صورت وجود ماتریس Jacobean انجام شود.

• خطی سازی با استفاده از EKF برای پیاده سازی بسیار دشوار است چراکه مشتق ماتریس Jacobean در اغلب موارد بی اهمیت است.

یک تبدیل کسری خطی بدون افت فشار (LFT) که از KF بدون افت فشار کارآمدتر است، پیشنهاد شده است در [129]. LFT سیستم غیرخطی را به یک مدل خطی معادل تبدیل می کند و یک تبدیل بدون تحرک، ساختار غیرخطی را مدیریت می کند. علاوه بر این [123] تکنیک LFT را به یک محیط چند سنسور با استفاده از رویکرد Bayesian گسترش می دهد.

## 1) فرصت ها و چالش ها

روش های تلفیق داده مبتنی بر KF یک انتخاب محبوب برای محیط غیر خطی است. این ساده است، پیچیدگی کمتر و آسان برای پیاده سازی و به طور گسترده پذیرفته شده است. اما با مشاهدات جعلی عملی نیست. قرار گرفتن در معرض داده خارج از محدوده می تواند منجر به شکست KF به خصوص در محیط های متراکم حسگر مانند IoT شود. این دقیق بودن می تواند در رابطه با هوش مصنوعی و روش های تلفیقی باشد و محاسبات گسترده ای در KF مورد نیاز است اگر سنسور زیادی دخیل باشند. KF برای سیستم های خطی بهتر است و استقلال شرطی درگیر است. روش های تلفیق داده مبتنی بر منطق فازی واقعا در سیستم های غیر خطی و چند متغیره مفید هستند. آنها همچنین به راحتی تغییر می کنند. یک سیستم فازی، بر اساس قوانین، داده های خام بدست آمده از سنسورها را تلفیق می کند. علاوه بر این پیش بینی کننده فازی برای ایجاد تلفیق داده بسیار دقیق برای کاربردهای بسیار حساس، این سیستم تلفیق می تواند با پهنای باند بالا و کارایی بالا کار کند [130]. در سیستم های غیر خطی، مسئله ی به روز رسانی نیز بسیار مهم است. صرف نظر از چند کار، غیر خطی هنوز یک وظیفه پیچیده برای مدیریت در سیستم های تلفیق داده چند سنسور است.

## D. ردیابی هدف

ردیابی هدف یکی از قدیمی ترین حوزه ها است که استفاده از تلفیق داده ها برای آن به کار می آید. تلفیق داده ها در دامنه ردیابی نقش مهمی در برنامه های نظامی، رباتیک، سیستم های بی سیم و حمل و نقل دارد. در [131]، اصول موقعیت یابی مختلف و تعامل بین IoT و اشیاء در سیستم های ردیابی و موقعیت یابی توصیف شده است. برخی از مزایای احتمالی سیستم های تلفیق ردیابی و نظارت عبارتند از انتخاب صحیح هدف، تشخیص تهدید، شناسایی اشیاء متحرک و غیر مجاز در حوزه امنیتی سطح بالا و تصمیم گیری به موقع. در این بخش، چالش های مختلف در الگوریتم های تلفیق داده ردیابی، از جمله ردیابی چند هدفه (MTT)، بهره وری هزینه، کاهش خطا، مسائل غیرهمزمانی و مسیر به مسیر (T2T) را بررسی و طبقه بندی می کنیم.

یک الگوریتم ردیابی چند هدف از یک صحنه ارتباط داده سلسله مراتبی تطبیقی استفاده می کند. این طرح تطبیقی ویژگی های با قابلیت اطمینان بالا در صحنه های مربوط به اهداف داده شده را تعیین می کند. با کمک ویژگی های قابل اطمینان، فضاهای ویژگی سلسله مراتبی ایجاد شده اند و ارتباطات داده مختلف لایه ها صورت می گیرد. الگوریتم به طور مؤثر و کارآمد در سیستم های داخلی و خارجی کار می کند [132]، در حالی که در [133] یک الگوریتم MTT بر اساس مسئله تخصیص حسگر و یک مجموعه ای از سنسورها به صورت پویا شناسایی شده است. بعدها، مسیر تلفیق داده اشیاء ردیابی شده و تعامل انجام می شود. همانطور که هدف اشیاء از زمان به زمان از طریق تکالیف سنسور حرکت می کند، فرموله کردن مسئله در مورد محدودیت بهینه سازی برای به حداکثر رساندن کارایی ردیابی برای اهداف مربوطه انجام شده است. الگوریتم سپس یک زیر جستجوی شیب تکرارشونده را اجرا می کند که بهینه نزدیک برای مسئله برنامه نویسی عدد صحیح است. راه حل بیان شده مقرون به صرفه و مقیاس پذیر است. مقاله [134] همچنین پوشش ضمانت و کارایی انرژی را مورد بحث قرار می دهد. یکی دیگر از تکنیک های تلفیق ردیابی چند شی، استفاده از گشتاور نفس خودرو، تصویر و رادار می باشد که در آن تلفیق داده در سطح بالا قرار می گیرند که نتیجه بسیار قابل اعتماد می باشد. این طرح همچنین مکان های اشیاء ثابت را می دهد و می تواند برآورد عرض انجام دهد.

این الگوریتم ها با استفاده از نرم افزار SASPENCE پیاده سازی می شوند. این رویکرد ثابت شده در شرایط محیطی سخت بسیار قوی است [135].

تلفیق داده ها در برنامه های ردیابی و موقعیت یابی نیز در صنعت خودرو محبوب است. تلفیق داده سیستم موقعیت یابی سراسری (GPS) همراه با داده های سیستم ناوبری (INS) در سیستم های موقعیت یابی محبوب است. اگر چه برنامه های کاربردی INS بسیار دقیق هستند، نصب INS هزینه زیادی دارد و وقت گیر است. برآورد سیستم موقعیت یابی خودرو باید بسیار دقیق، قابل اعتماد و با ارائه تداوم اطلاعات باشد. گیرنده های GPS کم هزینه معمولاً در برنامه های کاربردی سنتی خودرو استفاده می شود. این سیستم ها بسیار دقیق یا قابل اعتماد نیستند و تضمین ارائه تداوم اطلاعات در طول خطای GPS را ارائه نمی دهند.

برای مقابله با اشتباهات در GPS، چندین الگوریتم تلفیق داده مبتنی بر فیلتر Bayesian در ادبیات گذشته مورد بحث قرار گرفته است؛ کارایی برآورد فیلتر Bayesian تا حد زیادی بستگی به انتخاب یک مدل فرآیند دارد. شرایط پویای رانندگی باید توسط فیلتر های Bayesian مورد توجه قرار گیرد. الگوریتم موقعیت یابی تلفیق داده مبتنی بر مدل چندگانه تعاملی (IMM) در [136] پیشنهاد شده است. فیلتر IMM با استفاده از سنسورهای GPS و جاسازی شده با توجه به شرایط مختلف پویای رانندگی تنظیم کنید. مدل خودروی سینماتیک و یک مدل خودرو پویا بخشی جدایی ناپذیر از فیلتر IMM در این مطالعه است. الگوریتم مقرون به صرفه، دقیق و قابل اعتماد در سناریوهای رانندگی پویا است. این به دلیل این واقعیت است که فیلتر IMM با استفاده از  $N$ ،  $KF$  موازی و یک الگوریتم تقریبی است. الگوریتم IMM از گروه فیلترهای مختلف مدل چندگانه، توانایی های محاسباتی برآورد کننده اولین مرتبه pseudo-Bayesian تعمیم یافته (1GPB) انجام شده است. به طور خاص، GPB1 یک روش بسیار قدرتمند برای ارزیابی رفتار است.

$$Cov[X, Y] = \begin{pmatrix} E[(X_1 - E[X_1])(Y_2 - E[Y_2])] & E[(X_1 - E[X_1])(Y_2 - E[Y_2])] \\ E[(X_2 - E[X_2])(Y_1 - E[Y_1])] & E[(X_2 - E[X_2])(Y_2 - E[Y_2])] \\ E[(X_3 - E[X_3])(Y_1 - E[Y_1])] & E[(X_3 - E[X_3])(Y_2 - E[Y_2])] \end{pmatrix} \quad (10)$$

موقعیت یابی اهداف بحرانی یکی از مهمترین کاربردهای تلفیق داده برای ارتش و دیگر حوزه های امنیتی بسیار حساس است. نرخ بالای احتمال تشخیص و نرخ پایین خطا باید برای عملکرد دقیق وجود داشته باشد. تلفیق داده یک رویکرد موثر برای بهبود عملکرد تشخیص با تعامل از طریق حسگرهای با قابلیت اطمینان محدود است. استقرار شبکه سنسور نظامی پرهزینه است؛ بنابراین، ترجیح می دهیم که قرار دادن موقعیت مطلوب سنسور ها برای رسیدن به حداکثر عملکرد استفاده شود. این محاسبات پیچیده، غیر محدب و یک مشکل بهینه سازی غیر خطی است. بر اساس مدل تلفیق داده احتمالاتی، الگوریتم قرار دادن حسگر سریع و کارآمد پیشنهاد شده است [137]. در مقایسه با الگوریتم های دیگر در ادبیات، این کار بهتر انجام می شود.

سیستم های تلفیق داده ردیابی از ارتباط و برآورد استفاده می کنند. سیستم IoT شامل چندین حسگر است و همچنین برای هدف چندگانه طراحی شده است. دو نوع ارتباط گسترده که در این سیستم ها استفاده می شود، ارتباطات اندازه به مسیر (MT) و مسیر به مسیر (T2T) است. تفاوت عمده بین این دو این است که پیاده سازی سطح حسگر در MT انجام می شود و پیاده سازی در سطح مرکز تلفیق داده ها در TT انجام می شود. ارتباط در تلفیق T2T بسیار مهم است. با توجه به مسیرهای نادرست و از دست رفته، خطاهای تصادفی و تعصب حسگر باعث پیچیدگی بیشتری می شود. یک تصمیم و برآورد مشترک (JDE) Bayes به طور موازی با حساسیت حسگر برای به دست آوردن یک JDE ساده شده، بهینه پیاده سازی شده است. با چک کردن ارتباط خطای تعصب حسگر، این طرح باعث بهبود دقت [138] می شود.

یک مسئله پیچیده غیرهمزمانی در تلفیق T2T در طی ردیابی حرکت سریع اشیاء بوجود می آیند؛ هر چند، این برای حرکت آهسته اشیاء ناچیز است. برای حل این مسئله، نویسندگان در [133] یک راه حل پیشنهاد دادند؛ این راه حل در سه مرحله مختلف اجرا می شود. در مرحله اول، برآورد در مرکز تلفیق صورت می گیرد. ضبط زمان واقعی مربوط به زمان ارجاع مرکز تلفیق پس از به دست آوردن داده حسگر در مرحله اول انجام می شود. در مرحله دوم، پیش بینی ها توسط مرکز تلفیق برای شیفتهای داده های دریافت شده مورد استفاده قرار می گیرند، که پس از آن شروع چرخه تلفیق بعدی خواهد شد. این مرحله داده های مورد نیاز را برای تلفیق مسیر به مسیر مجزا زمان واقعی را همگام سازی



می کند. در مرحله سوم و آخرین مرحله، داده های شبه همگام با یک الگوریتم برآوردگر بدون تبعیض خطی حداقل واریانس برای تلفیق تمام داده های حسگر استفاده می شود.

KF ناهمگام اغلب در تلفیق T2T بدون در نظر گرفتن اشکالات شناخته شده خود که در [138] ذکر شده است، استفاده می شود. یک تجزیه و تحلیل مقایسه ای در [139] از سه الگوریتم T2T به خوبی تثبیت شده، انجام شده است: ضرب کوواریانس [140]، اشتراک کوواریانس [19] و اجتماع کوواریانس [141] با KF ناهمگام برای دسترسی به کارایی برای مسئله تلفیق T2T.

تلفیق T2T ضرب کوواریانس نسبت به خطاهای میانگین مربع ریشه (RMSE) دو طرف دیگر بهتر عمل می کند، و نیز زمان اجرای ضرب کوواریانس حداقل است. اشتراک کوواریانس حداقل زمان اجرای دوم را می دهد، در جایی که RMSE تقریباً بیش از KF ناهمگام است. در اجتماع کوواریانس، RMSE تقریباً برابر با ضرب کوواریانس است، اما با یک هزینه محاسباتی بالا. دلیل اینکه KF ناهمگام برای تلفیق T2T حداقل کارایی را با توجه به حقایق زیر به دست می آورد:

- لیست شیء انعطاف پذیر شده KF در هر زمانی که لیست جدید شیء سنسور می رسد، به صورت ناهمگام به روز می شود. این اصولاً برای KF نادرست است، زیرا اشیاء حسگر از هر حسگر به صورت موقتی همبستگی دارند که نتیجه فیلترینگ قبلی است و همین پدیده همبستگی برای اشیاء ردیابی شده مشاهده می شود.
- پیاده سازی KF برای داده های فیلتر شده KF در حال حاضر خواستار تولید تاخیرهای فاز اضافی ناشی از ویژگی پایین گذار KF می باشد.

فرمول سازی ریاضی این روشها در زیر آورده شده است:

### a. معادله ضرب کوواریانس

ماتریس ضرب کوواریانس [140]، [142]، بین  $X$  و  $Y$  یک ماتریس  $MXN$  است که توسط  $Cov[X, Y]$ ، نشان می دهند، جایی که  $X$  و  $Y$  بردارهای تصادفی هستند،  $E$  عملگر انتظار است و  $X=[X_1 X_2 X_3]^T$  و  $Y=[Y_1 Y_2]^T$  [نگاه کنید (10)]، همانطور که در بالای این صفحه نشان داده شده است.

### b. معادله اشتراک کوواریانس

الگوریتم اشتراک کوواریانس [19] برای تلفیق دو یا چند برآورد متغیر حالت در KF با همبستگی ناشناخته استفاده می شود.  $a$  و  $b$  دو آیتم اطلاعات شناخته شده هستند که باید در آیتم اطلاعات  $C$  قرار داده شوند. آیتم  $a$  و  $b$  میانگین / کوواریانس  $\hat{a}$ ،  $A$  و  $\hat{b}$ ،  $B$  است، با اینکه همبستگی متقابل ناشناخته است. میانگین و کوواریانس برای آیتم  $C$  توسط به روزرسانی اشتراک کوواریانس به صورت زیر ارائه شده است:

$$C^{-1} = \omega A^{-1} + (1 + \omega) B^{-1} \quad (11)$$

$$\hat{c} = C(\omega A^{-1} \hat{a} + (1 + \omega) B^{-1} \hat{b}) \quad (12)$$

در اینجا  $\omega$  باید برای کاهش norm محاسبه شود.

### c. معادله اجتماع کوواریانس

روش اجتماع کوواریانس پیشنهاد شده در [141] اجازه می دهد تا دو مسیر را با هم تلفیق شوند، حتی اگر تفاوت تخمین های حالت بیش از کوواریانس ارائه شده توسط حداقل یک مسیر است. یک بردار حالت جدید  $\hat{c}$ ، برای برآورد تلفیق شده به دست آمده  $u$  استفاده می شود. ماتریس کوواریانس تلفیق شده توسط  $P_c$  نشان داده شده است که بیش از هر دو  $P_a$  و  $P_b$  است. برآورد تلفیق شده  $C=\{\hat{c}, P_c\}$ ، مشخص شده توسط:

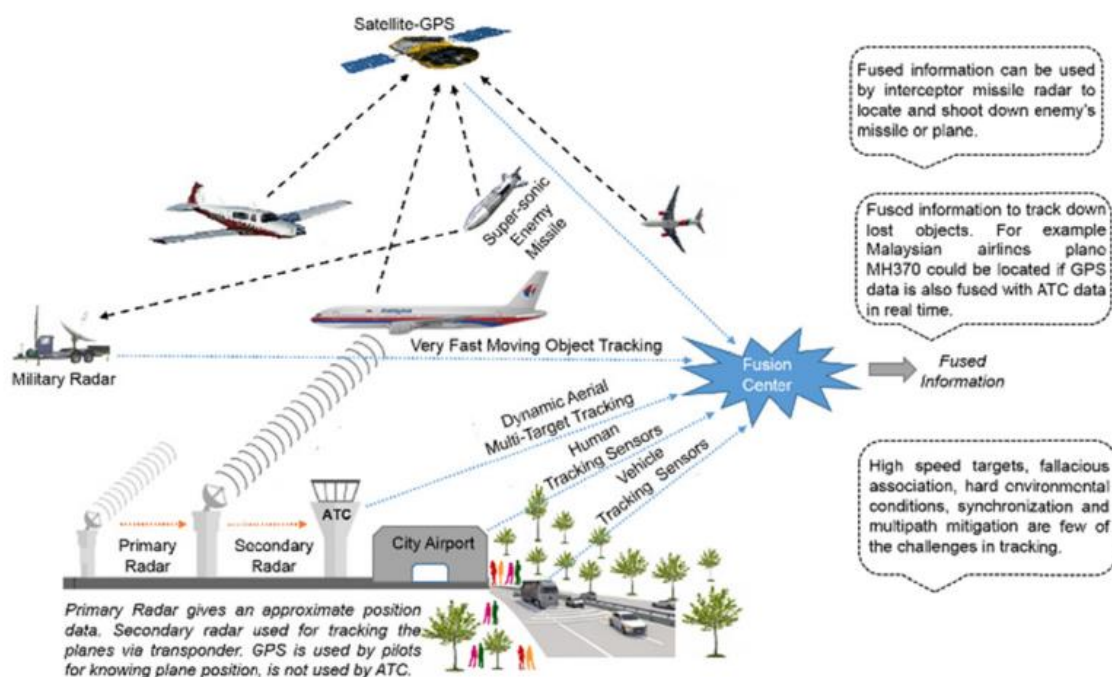
$$U_a = P_a + (\hat{c} - \hat{a}).(\hat{c} - \hat{a})^T$$

و

$$U_b = P_b + (\hat{c} - \hat{b}).(\hat{c} - \hat{b})^T$$

$$P_{ab} = \max(U_a, U_b)$$

$$\hat{e}_c = \arg \min(\det(P_{ab}))$$



تصویر 9. سناریو ردیابی چند هدفه ناهمگن در زمان واقعی

تلفیق T2T با برخی از چالش‌ها مواجه است: (1) در T2T، تلفیق داده ناهمگن در ردیابی زمان واقعی یک کار سخت برای رسیدن به آن است، (2) مسئله وابستگی بین برآورد مسیرها و حالت‌های سنسورهای مختلف در مرکز تلفیق داده‌ها و (3) مشکل افزونگی داده‌ها و فرایند نویز رایج که ممکن است به علت تعصب‌های سنسور و همبستگی باشد [133]، [143].

همچنین الگوریتم‌های تلفیق داده ردیابی با چالش‌هایی مانند اشیاء با حرکت سریع، درهم ریختگی با چگالی بالا و سیستم‌های ناهمزمان مواجه هستند. چالش‌های متعدد دیگری برای الگوریتم‌های تلفیق داده ردیابی با توجه به IoT وجود دارد: (1) کاهش چندگانه در فرکانس رادیویی، (2) هماهنگ‌سازی کوتاه مدت و دقیق و پیش هماهنگ‌سازی

مورد نیاز است؛ و (3) IoT نمی تواند استفاده از سلول های محلی سازی را برای یک گره منفرد به طور انحصاری ارائه دهد به دلیل اینکه امکان پذیر نیست.

در شکل 9 یک اطلاعات گرافیکی برای نشان دادن چندین سناریوی ردیابی نشان داده شده است. در [144]، معایب مربوط به سیستم های ردیابی هواپیمایی تجاری فعلی توضیح داده شده است. اخیراً، سایت خطوط هوایی هواپیمایی مالزی 370MH ممکن است در صورت ردیابی توسط ردیاب های متعدد مانند ATC و ماهواره، بتواند محل خراب شدن را پیدا کند؛ سپس داده تلفیق شده می تواند برای شناسایی محل هواپیماها استفاده شود. این فقط یک مثال زندگی واقعی است که نشان می دهد چگونه تلفیق داده می تواند مفید باشد.

## 1) فرصت ها و چالش ها

با گذشت زمان، ردیابی هدف چندگانه در حال محبوب شدن در زمینه های کاربردی مثل تعامل انسان و کامپیوتر، شناسایی مبتنی بر حرکت، نظارت خودکار، نظارت ترافیک و ناوبری وسایل نقلیه است [137]. کارهای گذشته در ردیابی چند هدفه، تلفیقی از مجموعه ای از ویژگی هایی برای شناسایی اشیاء منحصر به فرد است. با این حال، متمایل به هزینه های محاسباتی بالا با تجمع خطا، به خصوص در محیط IoT است. تنظیمات دوربین های شبکه امروزی با سناریوهای تغییر سریع پیچیده هستند. بنابراین، تکنیک های ثابتی برای انتخاب ویژگی ها امکان پذیر نیست. برخی از چالش های عمده برای الگوریتم های تلفیق داده ردیابی چند شیء در حوزه IoT وجود دارد [132]، [133]، [134]، [145]:

- پیچیدگی پیوند داده ها در سیستم های مبتنی بر IoT بالا است، زیرا تصمیمات مربوط به ارتباطات بحرانی به صورت محلی انجام می شود. این می تواند برای دامنه های محلی بهینه باشد اما نه در سطح جهانی.
- تصمیمات انجمنی نادرست غیرممکن است یا زمانی که آنها اتفاق می افتند، بازگرداندن آنها سخت است.
- مسئله تلفیق ناهمزمان در ردیابی هدف با سرعت بالا، به ویژه در تلفیق T2T اجتناب ناپذیر است.
- انسداد شی و اشکال پیچیده اشیاء.

• از دست دادن اطلاعات در هنگام نمایش D3 به D2.

## 5. تحولات تلفیق داده در حال ظهور

تلاش های اخیر تحقیق و توسعه در زمینه تلفیق داده ها، نشان دهنده تکامل آن از حوزه های مرسوم به حوزه های برنامه کاربردی بیشتر آینده مانند سیستم های اطلاعاتی سرگرمی، شناسایی فعالیت های انسان، وسایل نقلیه متصل و مستقل است. در این بخش، ما در ابتدا به تحولات برنامه های کاربردی اخیر که در آن تلفیق داده ها می توانند اعمال شوند و چگونه یادگیری عمیق می تواند تلفیق داده ها را افزایش دهد، تمرکز می کنیم.

### A. وسایل نقلیه مستقل و متصل

یک وسیله نقلیه مستقل به خودی خود هدایت می شود که هوش خود را دارد تا محیط اطرافش را درک کند و به طور پیوسته و ایمن هدایت شود. الگوریتم های یادگیری ماشین مختلفی برای درک محیط رانندگی با پردازش داده ها از یک نوع سنسور (منبع) مورد استفاده قرار می گیرند مانند: دوربین RGB ، LIDAR ، GPS و غیره. راه دیگری برای دستیابی به درک محیط رانندگی این است که داده ها از چند سنسور برای انجام یک هدف واحد تلفیق شوند. به عنوان مثال: تلفیق تصاویر GPS و دوربین برای پیش بینی فاصله ایمن یک وسیله نسبت به وسیله نقلیه دیگر رانندگی در جاده. استفاده از تلفیق داده ها در خودروهای خودمختار خیلی قدیمی است. با این حال، قبل از سال 2010 کار زیادی در این زمینه انجام نشده است. در سال های اخیر، کار های تحقیقات و توسعه های خودروهای بدون راننده محبوبیت پیدا کرده و علاقه ی فراوانی به فناوری اطلاعات و تیتان های صنعت خودرو برای توسعه ی هوش در رانندگی نشان داده شده است. این راه را برای توسعه فن آوری های تلفیق داده پیشرفته برای رانندگی خودکار آماده کرد.

تلفیق چندین ورودی به یک تک خروجی یک مسئله پیچیده است، اما به نظر می رسد نتیجه خیلی دقیق تر از تحلیل داده های یک سنسور است که به وسیله روند ادبیات به دست آمده است. به عنوان مثال در [146] نویسندگان از دوربین های تلفیقی و LIDAR برای درک صحنه رانندگی با برچسب گذاری بخش های تصویر استفاده می کنند که

در [147] از دوربین و لیزر برای ایجاد شیء نقشه های شبکه استفاده می کنند. در [148] و [149]، داده های تک منبع برای شناسایی عابران پیاده استفاده می شود. با در نظر گرفتن همین مسئله، با وجود استفاده از چندین منبع داده، یک روش نمونه برداری بالای عمقی مبتنی بر smoothing برای تشخیص عابر پیاده در [150] ارائه شده است که داده های مربوط به دوربین و LIDAR را تلفیق می کند. به طور مشابه در [151] نویسندگان از دانش شی کلاس برای شناسایی عابران پیاده، موانع ماشین و دوچرخه سواران استفاده می شود. کلاسیفایر پرسپترون چند لایه (MLP) در [152] برای شناسایی، توضیح و ردیابی اشیاء متحرک مستقل مورد استفاده قرار می گیرد. تلفیقی از stereovision، سرعت سنج و داده های LIDAR استفاده می شود و به MLP در [152] تغذیه می شود. هان و همکاران، از تصویر دوربین ها و دورسنجی چرخ برای استخراج موانع ایستا استفاده می کنند [153] در حالی که در [154] DST به منظور تلفیق داده سنسورها و شناسایی موانع با استفاده از دوربین، رادار لیزر، GPS، سنسور ادراک پیاده سازی شده است. به عنوان دانش وضعیت جاده برای رانندگی ایمن و صاف حیاتی است، در [155] داده های دوربین و LIDAR با یکدیگر برای انجام برآورد راه باریک تلفیق شده است. هیچ کدام از این ادبیات ذکر شده در بالا از مزایای و معایب واضحتری نسبت به دیگری برخوردار نیستند با توجه به این واقعیت که شرایط محیط رانندگی، مدل یادگیری و داده های مورد استفاده برای آزمایش، متفاوت هستند.

با توجه به سرعت پیشرفت تکنولوژیکی در رانندگی مستقل و گسترش موفقیت آمیز زیرساخت های IoT برای ایجاد جوامع هوشمند آینده در اقتصادهای توسعه یافته در سراسر جهان، بالا بردن قدرت برای افزایش "هوش برای رانندگی" به سطح جدیدتر در آینده نزدیک است. این را می توان با اتصال وسایل نقلیه خودمختار با هم با استفاده از زیرساخت های IOT بدست آورد و به آنها اجازه دهید صحبت، داده و هوشمندی خود را به اشتراک بگذارند. سقوط مرگبار اخیر گوگل [156] و تسلا [157] ماشین های خود راننده، نقطه عطفی را در مورد بلوغ هوش مصنوعی برای رانندگی مطرح کردند که توسط تکنولوژی titans توسعه یافته است. روش های تلفیق داده هایی که در بخش 1-5 مورد بحث قرار می گیرند تلفیقی از داده ها در داخل صفحه سنسور وسیله نقلیه خودمختار است، هیچ داده خارجی به غیر از داده های GPS استفاده نمی شود. به طور کامل از دانش ما، تقریباً هیچ تمرکزی در توسعه روش های تلفیق داده ها برای

یک وسیله نقلیه مستقل نیست به طوری که می تواند اطلاعاتی را که از سایر وسایل نقلیه مستقل بدست می آید که با استفاده از زیرساخت های IOT به آن متصل هستند را تلفیق کند.

با این حال، برای توسعه روش های تلفیق داده ها در آینده برای وسایل نقلیه مستقل متصل، می تواند موجب یک مجموعه کاملاً متفاوت از چالش ها شود. ما بر این باوریم که این امر درک محیط رانندگی برای رانندگی مستقل را بهتر خواهد کرد.

## **B. یادگیری عمیق برای تلفیق داده**

یادگیری عمیق (DL) شاخه ی از یادگیری ماشین توجه مهمی به خود جلب می کند. گارتنر یادگیری عمیقی را به عنوان یکی از 10 تحولات تکنولوژیکی که تاثیر مهمی بر برنامه ریزی استراتژیک اکثر سازمان ها در سال 2016 دارد، طبقه بندی کرد [158]. یادگیری عمیق نماینده یک مدل یادگیری است که سیستم عصبی انسان را تقلید می کند. داده های خام را به عنوان ورودی می گیرد و به طور خودکار نمایندگی هایی را که برای انجام پیش بینی لازم است، کشف می کند. یادگیری عمیق تلاش می کند تا انتزاع سطح بالای داده ها را مدل کند. مدل یادگیری عمیق می تواند چند لایه بین ورودی و خروجی داشته باشد که به آن برای فکر کردن کمک می کند. واقعیت جالب در مورد یادگیری عمیق این است که لایه های ویژگی ها از داده ها به طور خودکار آموخته می شود. LeCun مدیر تحقیق هوش مصنوعی در فیس بوک در نشریه بررسی خاصیت معروف خود در یادگیری عمیق [159] اظهار داشت که با یادگیری عمیق می تواند بسیاری از موفقیت های نزدیک در آینده به دلیل دو عامل مهم را ببینید: (1) آن نیاز به مهندسی بسیار کم در دست دارد و (2) به طور ذاتی از افزایش در مقدار منابع محاسباتی و داده های موجود بهره مند است. در یک بررسی دیگر، وانگ و راج، تکامل یادگیری عمیق را از مرحله اولیه تاکنون بررسی می کنند [160]. یادگیری عمیق در [159] و [160] در مورد موضوع مشابه، اما با اهداف مختلف مورد بررسی قرار گرفته است. مقاله [159] به طور انتقادی از تحولات حاضر در حال پیشرفت در یادگیری عمیق با حوزه کاربردی آن توضیح می دهد. علاوه بر ارائه فرصت های آینده در یادگیری عمیق نظر به اینکه در [160] نویسندگان بحث گذشته و فعلی یادگیری

عمیق و نحوه تبدیل از ANN های کلاسیک به یادگیری عمیق است. معماری های عمیق و فرمول بندی ریاضی مدل های یادگیری عمیق به طور جامع در [161] و [162] مورد بحث قرار می گیرد که می توان برای اطلاعات بیشتر به آنها اشاره کرد.

روش های یادگیری ماشین، کلاسیک مانند SVM، جنگل تصادفی و غیره به طور گسترده ای برای تلفیق اطلاعات از منابع مختلف استفاده می شود. با این حال، با یک سرعت بسیار آهسته شاخه یادگیری عمیق از یادگیری ماشین، در حال تبدیل شدن به حوزه تلفیق داده است. مدل های DL به طور موفقیت آمیزی برای یادگیری چندمدله در [163] - [165] مورد استفاده قرار گرفته است که شامل روش های پیچیده یادگیری است. در [163]، ویژگی های صوتی و تصویری برای استخراج اطلاعات مربوطه همبستگی دارند، در حالی که در [164] نویسندگان مدل یادگیری عمیق برای تلفیق عمیق چند مدله رویدادهای گسسته را پیشنهاد دادند. علاوه بر این در [165]، شبکه های عصبی پیچشی کاملاً عمیق برای تلفیق چند مدله داده های مشاهدات زمین مورد استفاده قرار می گیرند. [166] برای نظارت بر خطای پیچ های توپ که برای شناسایی الگوهای قطعی مربوط به سلامت پیچ توپ تعیین می شود، یک روش تلفیق داده مبتنی بر شبکه های باور عمیق (DBNs) است. یکی دیگر از کارها [167]، از DBN ها برای تلفیق داده استفاده می کند. با استفاده از داده سنسور دید دینامیک 128x128 و 64-کانال AER-EAR silicon cochlea به عنوان ورودی استفاده می شود. روش تلفیق داده پیشنهادی می تواند ارقام را حتی با داشتن حواس پرتی نیز شناسایی کند [167]. تعداد کمی از روش های تلفیق داده مبتنی بر یادگیری عمیق که برای دامنه های کاربردی مختلف در ادبیات زیر ارائه شده اند، عبارتند از: تلفیق داده ها برای تشخیص فعالیت در [168]، تلفیق داده ها برای ترافیک شبکه [169]، تلفیق داده های صوتی و تصویری در [170] و تشخیص عابر پیاده در [171].

DL شاخه ای از یادگیری ماشین دارای چندین مسئله است. در مرحله اول، آنها منابع مصرف می کنند این به این معنی است که آنها نیاز به پردازش قدرت و حافظه قابل توجه دارند. اما در روزهای اخیر هزینه سخت افزار کاهش می یابد و این روند در آینده نیز کاهش می یابد، بنابراین جنبه های مصرف منابع یادگیری عمیق را می توان نادیده گرفت. با این حال، در مورد سیستم تلفیق داده هایی که در داخل یک دستگاه تلفن همراه به جای یک سیستم قدرت مستقر



شده است، در این شرایط استفاده از یادگیری عمیق قابل بحث است. ثانياً، روش های یادگیری عمیق نسبت به نمونه های متضاد در مرحله آموزش آسیب پذیرند. سوم اینکه روش های یادگیری عمیق نیاز به تن دادن به داده ها برای آموزش دادن دارند اما انسان می تواند محیط را فقط از یک نمونه یا نمونه های کمی درک کند. علیرغم این موضوعات، یادگیری عمیق در زمینه های مختلف کاربرد مانند تشخیص گفتار، تجزیه و تحلیل تصویر، رانندگی مستقل و تشخیص الگو به اثبات رسیده است.

### C. دستگاه های موبایل، شهرها و جوامع هوشمند

با افزایش تعداد روزافزون، مردم به دنبال زندگی بهتر از حوزه روستایی به حوزه شهری مهاجرت می کنند. این در نهایت به بار اضافی برای زیرساخت های شهری و خدمات می انجامد. این به این معنی است که منابع ضروری تر مانند آب، برق، تجهیزات پزشکی و حمل و نقل و غیره ترسناک می شوند. حتی خدمات اصلی در شهرهای بزرگ در سراسر جهان در معرض آشفتگی هستند. نیاز فوری به مدیریت بسیار کارآمد منابع موجود و پیش بینی بسیار دقیق نیازهای آینده برای رشد جمعیت شهری وجود دارد. چگونه به این هدف برسیم؟ پاسخ در فرض شهر هوشمند است که به نظر می رسد جذاب است [172]، [173]. "مفهوم شهر هوشمند می تواند به جامعه هوشمند گسترش یابد؛ یعنی یک جامعه مبتنی بر دانش دیجیتالی فعال، آگاه و کارآمد به سوی پایداری اجتماعی، محیطی و اقتصادی" [6]. شهر هوشمند از زیرساخت های IOT استفاده می کند که هزاران حسگر محیط ما را حس می کنند. علاوه بر این، داده های حس شده که در اندازه و در قالب های مختلف بزرگ هستند، به محرکها کمک می کنند تا اقدامات مفید انجام دهند. همانطور که قبلاً در بخش های قبلی با تلفیق داده ها از چندین سنسور بحث شده است، دقت بهبود یافته و نتیجه گیری را می توان از یک منبع داده به دست آورد. برای شهر های هوشمند، تلفیق داده ها به علت ویژگی های آن به منظور ایجاد نتیجه گیری مبتنی بر منابع داده ای ناهمگن با دقت بالا و همچنین ایجاد پایگاه دانش شهری حیاتی است.

تلفیق داده برای برنامه های مختلف شهرهای هوشمند مانند مدیریت ترافیک، سیستم های هشدار، تشخیص رویداد، مراقبت های بهداشتی، مدیریت تامین انرژی و کنترل آلودگی مورد مطالعه قرار گرفته است. داده های پروژه اروپایی هوشمند سانتاندر برای همبستگی الگوهای ترافیک در شهر سانتاندر در رابطه با درجه حرارت تلفیق شده است [174].

در حالی که در [175]، نرم افزار ترافیک جاده هوشمند با تلفیق داده ها از سنسورهای کنار جاده ای در مرکز تلفیق در یک روش انرژی کارآمد مقرون به صرفه پیشنهاد شده است. در شهر هوشمند، سنسورهای مختلف می توانند ویژگی های مختلف یک رویداد خاص را ثبت کنند. به عنوان مثال، سنسور های ایستا (ثابت) می توانند ویژگی هایی مانند "از کجا و چگونه" را ضبط کنند، اما "چه کسی" را نتوانست توصیف کند. در حالی که سنسورهای پوشیدنی میتوانند ویژگی "چه کسی" را ضبط کنند. یک الگوریتم تلفیق داده در [176] برای تلفیق چنین جریانهای داده ای که رویداد شبیه را حس کند، اما ویژگی های متمایز آن را شناسایی کند، معرفی شده است. چند کار دیگر که با استفاده از برنامه های کاربردی مختلف شهرهای هوشمند با استفاده از پیاده سازی تلفیق داده ها هدف قرار می گیرد، مانند مدیریت آب [177] - [179]، داده های بزرگ اجتماعی [180]، منبع تغذیه و مدیریت هوشمند [181] - [183]، مدیریت هوشمندانه ترافیک [184] - [186]، و مراقبت های بهداشتی هوشمند [187] - [189] معرفی شده اند تا از زیرساخت های IoT استفاده کنند. کارهای دیگر مربوط به خدمات متنوع در شهرها یا جوامع هوشمند شامل سیستم های مدیریت اضطراری [190]، [191]، [208]، طرح های پیشنهادی مبتنی بر IoT برای بهبود واقعیت فرهنگ مجازی مبتنی بر شبیه سازی رویدادهای ترافیک [192]، سیستم های تحرک خود مختار [193] [194]، تدارکات شهری [195] - [198]، خدمات مبتنی بر مکان [199]، تحلیل عملکرد چندرسانه ای در شبکه های شهرهای هوشمند [200]، [201]، [203]، یافتن منابع جرم و جنایت [204]، انعطاف پذیری جامعه [198] شبکه های ad hoc وسیله نقلیه که می تواند برای تحرک و حمل و نقل داده (مانند مه) مورد استفاده قرار گیرد پیشنهاد شده در [191]، محاسبات سبز برای تلفن های موبایل [206]، ابرقابل ها [207]، هوش مصنوعی [208]، تحرک هوشمندانه با آگاهی اجتماعی [209]، و خدمات مبتنی بر مکان با حفظ حریم خصوصی داده [199]، اینترنت چیزهای فرهنگی و پیشنهادات مشابه [210] - [212] و برنامه ریزی شهری [8].

استفاده روزافزون از دستگاه های همراه مانند تلفن های هوشمند، ساعت های هوشمند و تبلت ها و دیگر موارد در حال افزایش است، برنامه های مختلف در این دستگاه ها نیاز به دسترسی به اطلاعات زیادی و از نوع متفاوتی برای درک بهتر زمینه های ما دارند. این به نوبه خود، افزایش تقاضا برای سیستم عامل هایی دارد که امکانات حسگر تلفیق داده است. این قدرت نسل بعدی دستگاه های همراه هوشمند برای جوامع هوشمند آینده است. یک تلفن هوشمند شامل شتاب سنج،ژیروسکوپ، قطب نمای دیجیتال، سنسور نور محیط و سنسور مجاورت است که توانایی ارائه بهتر زمینه هشیاری دارد. Pires و همکاران، در کار خود بررسی جامع روش های تلفیق داده را که برای سنسورهای جاسازی شده در دستگاه های تلفن همراه برای به تشخیص فعالیت های انسانی مورد استفاده قرار گرفته است، مورد بررسی قرار داد [15]. علاوه بر این در [213] و [214]، چندین روش تلفیق را بر اساس دستگاه های تلفن همراه پیشنهاد کردند تا فعالیت های مختلف روزانه زندگی را تشخیص دهند. کارهای بسیار کمی بر روی روشهای تلفیق داده انجام شده است که به خصوص در استفاده روزانه از دستگاه های تلفن همراه متمرکز شده است. ما معتقدیم که در آینده استفاده از دستگاه های تلفن همراه به مراتب بیشتر خواهد بود، از این رو برنامه کاربردی آگاهانه بهتر مورد نیاز خواهد شد، که در آن داده تلفیق داده می تواند نقش مهمی ایفا کند.

امروزه، پیاده سازی کامل و مقیاس بزرگ شهرهای هوشمند در سراسر جهان به علت محدودیت های مالی و تکنولوژیکی در نخستین دوره رشدشان است. چالش های واقعی زمانی پیش می آید که در آینده نزدیک برنامه های مورد بحث در بالا شهرهای هوشمند با یکدیگر برای تبدیل محیط اجتماعی و تکنولوژیکی امروز فعلی ما به جوامع پر جنب و جوش هوشمند جمع بندی شوند.

## 6. نتیجه

اینترنت اشیاء (IoT) قرار است به یکی از پیشرفت های کلیدی فناوری زمان ما تبدیل شود؛ به شرطی که بتوانیم از پتانسیل کامل آن استفاده کنیم. با توجه به جریان عظیمی از اشیاء متنوعی که به تدریج در حال ظهور هستند انتظار می رود که تعداد اشیاء متصل به IoT تا سال 2020 به 50 میلیارد برسد. IoT ازین جهت انتظار میرود یک تولید

کننده اصلی داده های بزرگ باشد. به اشتراک گذاری و تعامل داده ها و سایر منابع می تواند کلیدی برای ایجاد محیط های پایدار فراگیر، مانند شهرهای هوشمند و جوامع باشد. تلفیق و تحلیل به موقع داده های بزرگ که از IoT و سایر منابع به دست آمده است، برای ایجاد تصمیم گیری دقیق، قابل اعتماد و مدیریت بسیار کارآمد محیط های فراگیر می تواند یک چالش بزرگ در آینده باشد. هوش محاسباتی در این چالش نقش کلیدی ایفا می کند. تعدادی از مطالعات بر روی تلفیق داده ها وجود دارد. با این حال، این مطالعات عمدتاً بر روی حوزه ی برنامه ها یا طبقه بندی های خاصی متمرکز هستند.

در این مقاله ما به بررسی ادبیات تلفیق داده ها برای IoT با تمرکز خاص بر روش های ریاضی و محیط های خاص IoT پرداخته ایم. روش های ریاضی مورد بحث شامل روش های احتمالاتی، هوش مصنوعی و نظریه اعتقاد است. محیط مورد بحث شامل محیط های توزیع شده، ناهمگن، غیر خطی و ردیابی شیء می باشند. فرصت ها و چالش ها برای هر یک از روش های ریاضی و محیط ها مورد بحث قرار گرفت. حوزه های در حال توسعه که به طور ذاتی از تلفیق داده ها و IoT، وسایل نقلیه مستقل، یادگیری عمیق برای تلفیق داده ها و شهرهای هوشمند بهره می برند، مورد بحث قرار گرفت. فرصت ها و چالش های تلفیق داده ها به طور کلی و همچنین مشخص شده برای IoT ارائه شد. به نظر ما، در حال حاضر چنین مطالعه ی وجود ندارد.

## REFERENCES

- [1] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A survey," *Comput. Netw.*, vol. 54, no. 15, p. 2787–2805, Oct. 2010.
- [2] F. Alam, R. Mehmood, I. Katib, and A. Albeshri, "Analysis of eight data mining algorithms for smarter Internet of Things (IoT)," *Procedia Comput. Sci.*, vol. 98, pp. 437–442, Dec. 2016.
- [3] Overview of the Internet of Things, Rec. ITU-T Y.2060, 2012.
- [4] The National Intelligence Council, National Intelligence Council, and SRI Consulting Business Intelligence, *Disruptive Civil Technologies: Six Technologies With Potential Impacts on US Interests Out to 2025*. SRI Consulting Business Intelligence, Apr. 2008, p. 48.
- [5] D. Evans, "The Internet of Things—How the next evolution of the Internet is changing everything," CISCO, San Jose, CA, USA, White Paper, Apr. 2011.
- [6] R. Mehmood, F. Alam, N. N. Albogami, I. Katib, A. Albeshri, and S. M. Altowaijri, "UTiLearn: A personalised ubiquitous teaching and learning system for smart societies," *IEEE Access*, vol. 5, pp. 2615–2635, 2017.
- [7] H. B. Mitchell, *Multi-Sensor Data Fusion: An Introduction*. Berlin, Germany: Springer, 2007.
- [8] MIT Technology Review. (2013). 2013: The Year of the Internet of Things. [Online]. Available: <https://www.technologyreview.com/s/509546/2013-the-year-of-the-internet-of-things/>
- [9] S. Li, L. Da Xu, and S. Zhao, "The Internet of Things: A survey," *Inf. Syst. Frontiers*, vol. 17, no. 2, pp. 243–259, 2015.
- [10] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A survey on enabling technologies, protocols, and applications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 2347–2376, 4th Quart., 2015.
- [11] J. Dong, D. Zhuang, Y. Huang, and J. Fu, "Advances in multi-sensor data fusion: Algorithms and applications," *Sensors*, vol. 9, no. 10, pp. 7771–7784, Sep. 2009.
- [12] F. Castanedo, "A review of data fusion techniques," *Sci. World J.*, vol. 2013, Sep. 2013, Art. no. 704504.
- [13] B. Khaleghi, A. Khamis, F. O. Karray, and S. N. Razavi, "Multisensor data fusion: A review of the state-of-the-art," *Inf. Fusion*, vol. 14, no. 1, pp. 28–44, 2013.
- [14] N.-E. El Faouzi, H. Leung, and A. Kurian, "Data fusion in intelligent transportation systems: Progress and challenges—A survey," *Inf. Fusion*, vol. 12, no. 1, pp. 4–10, 2011.
- [15] I. M. Pires, N. M. Garcia, N. Pombo, and F. Flórez-Revuelta, "From data acquisition to data fusion: A comprehensive review and a roadmap for the identification of activities of daily living using mobile devices," *Sensors*, vol. 16, no. 2, p. 184, 2016.
- [16] X. Qin and Y. Gu, "Data fusion in the Internet of Things," *Procedia Eng.*, vol. 15, pp. 3023–3026, Dec. 2011.
- [17] S. Gite and H. Agrawal, "On context awareness for multisensor data fusion in IoT," in *Proc. 2nd Int. Conf. Comput. Commun. Technol.*, vol. 381. 2015, pp. 85–93.
- [18] M. Wang, C. Perera, P. P. Jayaraman, M. Zhang, P. Strazdins, and R. Ranjan, "City data fusion: Sensor data fusion in the Internet of Things," *Int. J. Distrib. Syst. Technol.*, vol. 7, no. 1, pp. 15–36, 2016.
- [19] D. L. Hall and J. Llinas, *Handbook of Multisensor Data Fusion*. New York, NY, USA: CRC Press, 2001, p. 529.
- [20] S. Vural and E. Ekici, "On multihop distances in wireless sensor networks with random node locations," *IEEE Trans. Mobile Comput.*, vol. 9, no. 4, pp. 540–552, Apr. 2010.
- [21] M. Kumar, D. P. Garg, and R. A. Zachery, "A generalized approach for inconsistency detection in data fusion from multiple sensors," in *Proc. Amer. Control Conf.*, 2006, pp. 2078–2083.
- [22] P. Smets, "Analyzing the combination of conflicting belief functions," *Inf. Fusion*, vol. 8, no. 4, pp. 387–412, 2007.
- [23] *Data Fusion System Engineering Guidelines*. USAF Space Warfare Center (SWC), 1997.
- [24] J. R. Raol, *Multi-Sensor Data Fusion With MATLAB*, 1st ed. Boca Raton, FL, USA: CRC Press, 2009.

- [25] D. P. Mandic et al., "Data fusion for modern engineering applications: An overview," in *Artificial Neural Networks: Formal Models and Their Applications—ICANN (Lecture Notes in Computer Science)*. Heidelberg, Germany: Springer, 2005, p. 715.
- [26] T. E. Fortmann, Y. Bar-Shalom, and M. Scheffe, "Multi-target tracking using joint probabilistic data association," in *Proc. 19th IEEE Conf. Decision Control Including Symp. Adapt. Process.*, Dec. 1980, pp. 807–812.
- [27] S. H. Rezatofighi, A. Milan, Z. Zhang, Q. Shi, A. Dick, and I. Reid, "Joint probabilistic data association revisited," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2015, pp. 3047–3055.
- [28] Z. Ding and L. Hong, "Bias phenomenon and compensation for PDA/JPDA algorithms," *Math. Comput. Model.*, vol. 27, no. 12, pp. 1–16, 1998.
- [29] S. Cong and L. Hong, "Computational complexity analysis for multiple hypothesis tracking," *Math. Comput. Model.*, vol. 29, no. 9, pp. 1–16, 1999.
- [30] S. Streltsov, "Airborne learning algorithms: A case of a ground target multiple hypothesis tracker," in *Proc. Int. Conf. Integr. Knowl. Intensive Multi-Agent Syst.*, Sep./Oct. 2003, pp. 190–195.
- [31] H. Holzapfel, K. Nickel, and R. Stiefelhagen, "Implementation and evaluation of a constraint-based multimodal fusion system for speech and 3D pointing gestures," in *Proc. 6th Int. Conf. Multimodal Interfaces (ICMI)*, 2004, pp. 175–182.
- [32] R. C. Luo, C.-C. Yih, and K. L. Su, "Multisensor fusion and integration: Approaches, applications, and future research directions," *IEEE Sensors J.*, vol. 2, no. 2, pp. 107–119, Apr. 2002.
- [33] M. Guerriero, L. Svensson, and P. Willett, "Bayesian data fusion for distributed target detection in sensor networks," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3417–3421, Jun. 2010.
- [34] T. A. Biresaw, A. Cavallaro, and C. S. Regazzoni, "Tracker-level fusion for robust Bayesian visual tracking," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 5, pp. 776–789, May 2015.
- [35] L.-W. Fong, "An adaptive filter for multi-sensor track fusion," in *Proc. Int. Conf. Signal Process. (ICSP)*, 2008, pp. 231–235.
- [36] L.-W. Fong, "Multi-sensor track fusion via multiple-model adaptive filter," in *Proc. IEEE Conf. Decision Control*, Dec. 2009, pp. 2327–2332.
- [37] L.-W. Fong, "Adaptive information matrix filtering fusion with nonlinear classifier," in *Proc. SICE Annu. Conf.*, 2010, pp. 2214–2219.
- [38] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York, NY, USA: Wiley, 2001.
- [39] L.-W. Fong, "Switching adaptive filter design using Bayesian classification approach for multi-sensor data fusion," in *Proc. 8th Asian Control Conf. (ASCC)*, 2011, pp. 1310–1315.
- [40] A. Noyvirt and R. Qiu, "Human detection and tracking in an assistive living service robot through multimodal data fusion," in *Proc. IEEE 10th Int. Conf. Ind. Inform.*, Jul. 2012, pp. 1176–1181.
- [41] S. Duffner and J.-M. Odobez, "Track creation and deletion framework for long-term online multiface tracking," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 272–285, Jan. 2013.
- [42] D. L. Hall and S. A. H. McMullen, *Mathematical Techniques in MultiSensor Data Fusion*, 1st ed. Norwood, MA, USA: Artech House, 2004.
- [43] X. Zhu, F. Kui, and Y. Wang, "Predictive analytics by using Bayesian model averaging for large-scale Internet of Things," *Int. J. Distrib. Sensor Netw.*, vol. 2013, no. 3, pp. 1–10, 2013.
- [44] M. Kumar, D. P. Garg, and R. A. Zachery, "A method for judicious fusion of inconsistent multiple sensor data," *IEEE Sensors J.*, vol. 7, no. 5, pp. 723–733, May 2007.
- [45] Y. Guo, X. Cai, and J. Ma, "Improved JPDA for fast fault detection," in *Proc. 30th Chin. Control Conf. (CCC)*, Jul. 2011, pp. 4167–4169.
- [46] J. Liu, C. Han, and Y. Hu, "Multiple target tracking using particle filter based multi-scan joint probabilistic data association filter," in *Proc. 30th Chin. Control Conf. (CCC)*, Jul. 2011, pp. 3134–3139.
- [47] J. Liu, C. Han, F. Han, and Y. Hu, "Multiple maneuvering target tracking by improved particle filter based on multiscan JPDA," *Math. Problems Eng.*, vol. 2012, Nov. 2012, Art. no. 372161.

- [48] G. Liu and G. Tian, "Distributed vision network for multiple target tracking using a dynamic hybrid consensus filter," in Proc. IEEE Int. Conf. Autom. Sci. Eng., Nov. 2016, pp. 805–808.
- [49] J. Guo and R. Zhang, "Novel implementation of track-oriented multiple hypothesis tracking algorithm," Chin. J. Electron., vol. 21, no. 4, pp. 770–774, 2012.
- [50] J. Hou, Y. Yang, Y. Chen, and L. Xu, "Multiple hypothesis tracker in the presence of RGPO/RGPI using amplitude information," in Proc. 19th Int. Conf. Inf. Fusion, 2016, pp. 703–710.
- [51] S. Coraluppi, C. Carthel, W. Kreamer, and A. Willsky, "Recent advances in multi-INT track fusion," in Proc. IEEE Aerosp. Conf., Mar. 2016, pp. 1–13.
- [52] W. A. Abdulhafiz and A. Khamis, "Bayesian approach with pre- and postfiltering to handle data uncertainty and inconsistency in mobile robot local positioning," J. Intell. Syst., vol. 23, no. 2, pp. 133–154, 2014.
- [53] F. Bourgault et al., "Scalable Bayesian human-robot cooperation in mobile sensor networks," in Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS), Sep. 2008, pp. 2342–2349.
- [54] W. C. Chen and N. V. Z. Shuley, "Robust target identification using a modified generalized likelihood ratio test," IEEE Trans. Antennas Propag., vol. 62, no. 1, pp. 264–273, Jan. 2014.
- [55] J. Liu, H. Li, and B. Himed, "Two target detection algorithms for passive multistatic radar," IEEE Trans. Signal Process., vol. 62, no. 22, pp. 5930–5939, Nov. 2014.
- [56] B. Wang. (2013). Next big future: X-Apple Siri Director Samsung's AI for the Internet of Things. Next Big Future Inc., accessed on Dec. 22, 2016. [Online]. Available: <http://www.nextbigfuture.com/2013/11/x-apple-siridirector-is-taking.html>
- [57] P. Mishra. (2015). Google is making a land grab for the Internet of Things. Tech Crunch, accessed on Dec. 20, 2016. [Online]. Available: <https://techcrunch.com/2014/01/27/google-is-making-a-land-grabfor-the-internet-of-things/>
- [58] BI-Intelligence. (2016). The Google of China announces new AI and Internet of Things initiatives. Business Insider, accessed on Dec. 20, 2016. [Online]. Available: <http://www.businessinsider.com/baidu-announcesnew-ai-and-internet-of-things-initiatives-2016-9>
- [59] S. B. Kotsiantis, "Supervised machine learning: A review of classification techniques," Informatica, vol. 31, no. 1, pp. 249–268, 2007.
- [60] F. Chen, P. Deng, J. Wan, D. Zhang, A. V. Vasilakos, and X. Rong, "Data mining for the Internet of Things: Literature review and challenges," Int. J. Distrib. Sensor Netw., vol. 11, no. 8, pp. 249–268, 2015.
- [61] C. C. Aggarwal, N. Ashish, and A. Sheth, "The Internet of Things: A survey from the data-centric perspective," in Managing and Mining Sensor Data. New York, NY, USA: Springer, 2013, pp. 383–428.
- [62] S. Bin, L. Yuan, and W. Xiaoyi, "Research on data mining models for the Internet of Things," in Proc. Int. Conf. Image Anal. Signal Process., 2010, pp. 127–132.
- [63] C.-W. Tsai, C.-F. Lai, M.-C. Chiang, and L. T. Yang, "Data mining for Internet of Things: A survey," IEEE Commun. Surveys Tuts., vol. 16, no. 1, pp. 77–97, 1st Quart., 2014.
- [64] S. Challa, M. Palaniswami, and A. Shilton, "Distributed data fusion using support vector machines," in Proc. 5th Int. Conf. Inf. Fusion (FUSION) (IEEE Cat.No.02EX5997), vol. 1. Jul. 2002 pp. 881–885.
- [65] S. Zhao, "Remote sensing data fusion using support vector machine," in Proc. IEEE Int. Geosci. Remote Sens. Symp., vol. 4. Sep. 2004, pp. 2575–2578.
- [66] B. Bigdeli, F. Samadzadegan, and P. Reinartz, "A decision fusion method based on multiple support vector machine system for fusion of hyperspectral and LIDAR data," Int. J. Image Data Fusion, vol. 5, no. 3, pp. 196–209, 2014.
- [67] A. Merentitis, C. Debes, R. Heremans, and N. Frangiadakis, "Automatic fusion and classification of hyperspectral and LiDAR data using random forests," in Proc. IEEE Int. Geosci. Remote Sens. Symp., Jul. 2014, pp. 1245–1248.
- [68] E. T. McGee and J. D. McGregor, "Using dynamic adaptive systems in safety-critical domains," in Proc. IEEE/ACM 11th Int. Symp. Softw. Eng. Adapt. Self-Manag. Syst., May 2016, pp. 115–121.
- [69] T. P. Banerjee and S. Das, "Multi-sensor data fusion using support vector machine for motor fault detection," Inf. Sci., vol. 217, pp. 96–107, Dec. 2012.

- [70] A. Starzacher and B. Rinner, "Evaluating KNN, LDA and QDA classification for embedded online feature fusion," in Proc. Int. Conf. Intell. Sensors, Sens. Netw. Inf. Process. (ISSNIP), 2008, pp. 85–90.
- [71] N. El Gayar, F. Schwenker, and C. Y. Suen, Eds., *Artificial Neural Networks in Pattern Recognition (Lecture Notes in Artificial Intelligence)*, vol. 8774. Switzerland: Springer International Publishing, 2014.
- [72] D. Chen, K. Zhang, and T. Liao, "Practical travel time prediction algorithms based on neural network and data fusion for urban expressway," in Proc. 6th Int. Conf. Natural Comput. (ICNC), vol. 4. 2010, pp. 1754–1758.
- [73] H. B. Azad, S. Mekhilef, and V. G. Ganapathy, "Long-term wind speed forecasting and general pattern recognition using neural networks," *IEEE Trans. Sustain. Energy*, vol. 5, no. 2, pp. 546–553, Apr. 2014.
- [74] F. Zhang, K. Cho, J. Choi, Y.-M. Lee, and K. Y. Lee, "A study on wind speed prediction using artificial neural network at Jeju Island in Korea II," in Proc. IEEE 54th Int. Midwest Symp. Circuits Syst., Aug. 2011, pp. 1–4.
- [75] P. M. Fonte, G. X. Silva, and J. C. Quadrado, "Wind speed prediction using artificial neural networks," *WSEAS Trans. Syst.*, vol. 4, no. 4, pp. 379–384, 2005.
- [76] Z. Liu, W. Gao, Y.-H. Wan, and E. Muljadi, "Wind power plant prediction by using neural networks," in Proc. IEEE Energy Convers. Congr. Expo., Sep. 2012, pp. 3154–3160.
- [77] M. Altini, D. Brunelli, E. Farella, and L. Benini, "Bluetooth indoor localization with multiple neural networks," in Proc. IEEE 5th Int. Symp. Wireless Pervasive Comput., May 2010, pp. 295–300.
- [78] H. Singh et al., "Real-life applications of fuzzy logic," *Adv. Fuzzy Syst.*, vol. 2013, Apr. 2013, Art. no. 581879.
- [79] P. Manjunatha, A. K. Verma, and A. Srividya, "Multi-sensor data fusion in cluster based wireless sensor networks using fuzzy logic method," in Proc. IEEE Region 10 3rd Int. Conf. Ind. Inf. Syst., Dec. 2008, pp. 1–6.
- [80] C. Lee, Y. Han, S. Jeon, D. Seo, and I. Jung, "Smart parking system for Internet of Things," in Proc. IEEE Int. Conf. Consum. Electron. (ICCE), Jan. 2016, pp. 263–264.
- [81] A. Khanna and R. Anand, "IoT based smart parking system," in Proc. Int. Conf. Internet Things Appl. (IOTA), 2016, pp. 266–270.
- [82] Z. Jian, C. Hongbing, S. Jie, and L. Haitao, "Data fusion for magnetic sensor based on fuzzy logic theory," in Proc. 4th Int. Conf. Intell. Comput. Technol. Autom., Mar. 2011, pp. 87–92.
- [83] I. Mobin, M. Abid-Ar-Rafi, N. Islam, and R. Hasan, "An intelligent fire detection and mitigation system safe from fire (SFF)," *Int. J. Comput. Appl.*, vol. 133, no. 6, pp. 1–7, 2016.
- [84] A. M. Aziz, M. Tummala, and R. Cristi, "Fuzzy logic data correlation approach in multisensor-multitarget tracking systems," *Signal Process.*, vol. 76, no. 2, pp. 195–209, 1999.
- [85] F. En, X. Wei-Xin, L. Zong-Xiang, and L. Peng-Fei, "A novel fuzzy track-to-track association algorithm in distributed sensor network," *J. Signal Process.*, vol. 27, no. 10, pp. 1561–1565, 2011.
- [86] A. M. Aziz, "A new fuzzy clustering approach for data association and track fusion in multisensor-multitarget environment," in Proc. Aersp. Conf., 2011, pp. 1–10.
- [87] G.-L. Ran and H.-H. Wu, "Fuzzy fusion approach for object tracking," in Proc. 2nd WRI Global Congr. Intell. Syst. (GCIS), vol. 3. 2010, pp. 219–222. [
- [88] W. Du, H. Ning, Y. Wei, and J. Wang, "Fuzzy double-threshold track association algorithm using adaptive threshold in distributed multisensor-multitarget tracking systems," in Proc. IEEE Int. Conf. Green Comput. Commun. IEEE Internet Things IEEE Cyber, Phys. Soc. Comput. (GreenCom-iThings-CPSCOM), Aug. 2013, pp. 1133–1137.
- [89] N. Mort and P. J. Escamilla-Ambrosio, "Hybrid Kalman filter-fuzzy logic adaptive multisensor data fusion architectures," in Proc. 42nd IEEE Conf. Decision Control, Dec. 2003, pp. 5215–5220.
- [90] V. Subramanian, T. F. Burks, and W. E. Dixon, "Sensor fusion using fuzzy logic enhanced Kalman filter for autonomous vehicle guidance in citrus groves," *Trans. ASABE*, vol. 52, no. 5, pp. 1411–1422, 2009.
- [91] G. Shafer, *A Mathematical Theory of Evidence*, 1st ed. Princeton, NJ, USA: Princeton Univ. Press, 1976.
- [92] A. Bellenger and S. Gatepaille. (2011). "Uncertainty in ontologies: Dempster–Shafer theory for data fusion applications." [Online]. Available: <https://arxiv.org/abs/1106.3876>
- [93] F. Rottensteiner, J. Trinder, S. Clode, and K. Kubik, "Using the Dempster–Shafer method for the fusion of LIDAR data and multi-spectral images for building detection," *Inf. Fusion*, vol. 6, no. 4, pp. 283–300, 2005.



- [94] D. Jiang, D. Zhuang, Y. Huang, and J. Fu, "Survey of multispectral image fusion techniques in remote sensing applications," in *Image Fusion and Its Applications*. Croatia: INTECH, 2011.
- [95] V. Saeidi, B. Pradhan, M. O. Idrees, and Z. A. Latif, "Fusion of airborne LiDAR with multispectral SPOT 5 image for enhancement of feature extraction using Dempster-Shafer theory," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6017–6025, Oct. 2014.
- [96] H. Janssen and W. Niehsen, "Vehicle surround sensing based on information fusion of monocular video and digital map," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2004, pp. 244–249.
- [97] S. K. Gehrig, S. Wagner, and U. Franke, "System architecture for an intersection assistant fusing image, map, and GPS information," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2003, pp. 144–149.
- [98] D. Nienhuser, T. Gump, and J. M. Zollner, "A situation context aware Dempster-Shafer fusion of digital maps and a road sign recognition system," in *Proc. IEEE Intell. Veh. Symp.*, Jun. 2009, pp. 1401–1406.
- [99] H. Jamshidi, T. Lukaszewicz, A. Kashi, A. Berghuvud, H.-J. Zepernick, and S. Khatibi, "Fusion of digital map traffic signs and camera-detected signs," in *Proc. 5th Int. Conf. Signal Process. Commun. Syst. (ICSPCS)*, 2011, pp. 1–7.
- [100] R. Martin, J. Zhang, and C. Liu, "Dempster-Shafer theory and statistical inference with weak beliefs," *Statist. Sci.*, vol. 25, no. 1, pp. 72–87, 2010.
- [101] S. Challa and D. Koks, "Bayesian and Dempster-Shafer fusion," *Sadhana*, vol. 29, no. 2, pp. 145–176, Apr. 2004.
- [102] T. W. Martin and K. Chang, "A data fusion formulation for decentralized estimation predictions under communications uncertainty," in *Proc. 9th Int. Conf. Inf. Fusion (FUSION)*, 2006, pp. 1–7.
- [103] A. Abdelgawad, "Distributed data fusion algorithm for wireless sensor network," in *Proc. 11th IEEE Int. Conf. Netw., Sens. Control*, Apr. 2014, pp. 334–337.
- [104] M. Dhivya, M. Sundarambal, and L. N. Anand, "Energy efficient computation of data fusion in wireless sensor networks using cuckoo based particle approach (CBPA)," *Int. J. Commun. Netw. Syst. Sci.*, vol. 4, pp. 249–255, Apr. 2011.
- [105] Z. Lu, S. L. Tan, and J. Biswas, "Function placement of data fusion for active networks paradigm in wireless sensor networks," in *Proc.- 8th IEEE Int. Conf. Mobile Ad-Hoc Sens. Syst. (MASS)*, Oct. 2011, pp. 152–154.
- [106] K. Zhang, C. Li, and W. Zhang, "Wireless sensor data fusion algorithm based on the sensor scheduling and batch estimate," *Int. J. Future Comput. Commun.*, vol. 2, no. 4, p. 333, Aug. 2013.
- [107] V. Sasikala and C. Ch, "Energy efficient multipath data fusion technique for wireless sensor networks," *ACEEE Int. J. Netw. Secur.*, vol. 3, no. 2, pp. 34–41, 2012.
- [108] Y. Huang, *Energy Efficient Distributed Data Fusion in Multihop Wireless Sensor Networks*. Riverside, CA, USA: Univ. California Riverside, 2010.
- [109] K. C. Chang, C. Y. Chong, and S. Mori, "Analytical and computational evaluation of scalable distributed fusion algorithms," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 4, pp. 2022–2034, Oct. 2010.
- [110] J. Choo, S. Bohn, G. C. Nakamura, A. M. White, and H. Park, "Heterogeneous data fusion via space alignment using nonmetric multidimensional scaling," in *Proc. SIAM Int. Conf. Data Mining*, 2012, pp. 177–188.
- [111] Y. S. Chiou and F. Tsai, "A reduced-complexity data-fusion algorithm using belief propagation for location tracking in heterogeneous observations," *IEEE Trans. Cybern.*, vol. 44, no. 6, pp. 922–935, Jun. 2014.
- [112] E. F. Lock and D. B. Dunson, "Bayesian consensus clustering," *Bioinformatics*, vol. 29, no. 20, pp. 2610–2616, 2013.
- [113] P. Ray, L. Zheng, J. Lucas, and L. Carin, "Bayesian joint analysis of heterogeneous genomics data," *Bioinformatics*, vol. 30, no. 10, pp. 1370–1376, 2014.
- [114] I. González, S. Déjean, P. G. P. Martin, O. Gonçalves, P. Besse, and A. Baccini, "Highlighting relationships between heterogeneous biological data through graphical displays based on regularized canonical correlation analysis," *J. Biol. Syst.*, vol. 17, no. 2, pp. 173–199, 2009.
- [115] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, nos. 3–4, pp. 321–377, 1936.

- [116] K.-A. L. Cao, P. G. Martin, C. Robert-Granié, and P. Besse, "Sparse canonical methods for biological data integration: Application to a crossplatform study," *BMC Bioinformatics*, vol. 10, no. 1, p. 34, 2009.
- [117] H. Wold, *Estimation of Principal Components and Related Models by Iterative Least Squares*. New York, NY, USA: Academic, 1966.
- [118] N. Alaydie, F. Fotouhi, C. K. Reddy, and H. Soltanian-Zadeh, "Noise and outlier filtering in heterogeneous medical data sources," in *Proc. Workshop Database Expert Syst. Appl.*, 2010, pp. 115–119.
- [119] L. Fricoteaux, I. M. Thouvenin, and J. Olive, "Heterogeneous data fusion for an adaptive training in informed virtual environment," in *Proc. IEEE Int. Conf. Virtual Environ. Human-Comput. Inter. Meas. Syst. (VECIMS)*, Sep. 2011, pp. 113–118.
- [120] J. Zhao, Y. Liu, Y. Cheng, Y. Qiang, and X. Zhang, "Multisensor data fusion for wildfire warning," in *Proc. 10th Int. Conf. Mobile Ad-Hoc Sens. Netw.*, Dec. 2014, pp. 46–53.
- [121] H. Zhang, T. S. Huang, N. M. Nasrabadi, and Y. Zhang, "Heterogeneous multi-metric learning for multi-sensor fusion," in *Proc. 14th Int. Conf. Inf. Fusion*, 2011, pp. 1–8.
- [122] P. Pong, M. Morelande, and S. Challa, "Heterogeneous fusion with a combined evidential, probability and OWA methods for target classification," in *Proc. 13th Conf. Inf. Fusion*, Jul. 2010, pp. 1–12.
- [123] U. Rashid, H. D. Tuan, P. Apkarian, and H. H. Kha, "Multisensor data fusion in nonlinear Bayesian filtering," in *Proc. 4th Int. Conf. Commun. Electron.*, Aug. 2012, pp. 351–354.
- [124] U. Rashid, H. D. Tuan, H. H. Kha, and H. H. Nguyen, "Semi-definite programming for distributed tracking of dynamic objects by nonlinear sensor network," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, May 2011, pp. 3532–3535.
- [125] R. Zhan and J. Wan, "Iterated unscented Kalman filter for passive target tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 3, pp. 1155–1163, Jul. 2007.
- [126] S. J. Julier and J. K. Uhlmann, "A new extension of the Kalman filter to nonlinear systems," *Proc. AeroSense 11th Int. Symp. Aerospace/Defence Sens., Simulation Control*, 1997, pp. 1–2.
- [127] S. Julier, J. Uhlmann, and H. F. Durrant-Whyte, "A new method for the nonlinear transformation of means and covariances in filters and estimators," *IEEE Trans. Autom. Control*, vol. 45, no. 3, pp. 477–482, Mar. 2000.
- [128] D. L. Hall and J. Llinas, "An introduction to multisensor data fusion," *Proc. IEEE*, vol. 85, no. 1, pp. 6–23, Jan. 1997.
- [129] S. A. Pasha, H. D. Tuan, and B. N. Vo, "Nonlinear Bayesian filtering using the unscented linear fractional transformation model," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 477–489, Feb. 2010.
- [130] M. A. A. Akhoundi and E. Valavi. (2010). "Multi-sensor fuzzy data fusion using sensors with different characteristics." [Online]. Available: <https://arxiv.org/abs/1010.6096>
- [131] C. Wang, H. Liu, and Y. Gao, "Scene-adaptive hierarchical data association for multiple objects tracking," *IEEE Signal Process. Lett.*, vol. 21, no. 6, pp. 697–701, Jun. 2014.
- [132] Y. Fu, Q. Ling, and Z. Tian, "Distributed sensor allocation for multitarget tracking in wireless sensor networks," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 4, pp. 3538–3553, Oct. 2012.
- [133] F. Tango, E. Richter, U. Scheunert, and G. Wanielik, "Advanced multiple objects tracking by fusing radar and image sensor data—Application on a case study," in *Proc. 11th Int. Conf. Inf. Fusion*, Jul. 2008, pp. 1–7.
- [134] H. Talebi and A. M. A. Hemmatyar, "Asynchronous track-to-track fusion by direct estimation of time of sample in sensor networks," *IEEE Sens. J.*, vol. 14, no. 1, pp. 210–217, Jan. 2014.
- [135] J. Kichun, C. Keounyup, and S. Myounggho, "Interacting multiple model filter-based sensor fusion of GPS with in-vehicle sensors for real-time vehicle positioning," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 1, pp. 329–343, Mar. 2012.
- [136] X. Chang et al., "Sensor placement algorithms for fusion-based surveillance networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 22, no. 8, pp. 1407–1414, Aug. 2011.
- [137] Y. Han, H. Zhu, and C. Han, "Track-to-track association in the presence of sensor bias and the relative bias estimation," in *Proc. 15th Int. Conf. Inf. Fusion (FUSION)*, 2012, pp. 1044–1050.

- [138] K. Weiss, D. Stueker, and A. Kirchner, "Target modeling and dynamic classification for adaptive sensor data fusion," in Proc. IEEE Intell. Veh. Symp., Jun. 2003, pp. 132–137.
- [139] S. Matzka and R. Altendorfer, "A comparison of track-to-track fusion algorithms for automotive sensor fusion," in Proc. IEEE Int. Conf. Multisens. Fusion Integr. Intell. Syst. (MFI), vol. 35. Aug. 2008, pp. 69–81.
- [140] Y. Bar-Shalom, "On the track-to-track correlation problem," IEEE Trans. Autom. Control, vol. 26, no. 2, pp. 571–572, Apr. 1981.
- [141] J. K. Uhlmann, "Covariance consistency methods for fault-tolerant distributed data fusion," Inf. Fusion, vol. 4, no. 3, pp. 201–215, Sep. 2003.
- [142] Statlect. (2014). Cross Covariance Matrix. [Online]. Available: <http://www.statlect.com>
- [143] M. Aeberhard, S. Schlichtharle, N. Kaempchen, and T. Bertram, "Track-to-track fusion with asynchronous sensors using information matrix fusion for surround environment perception," IEEE Trans. Intell. Transp. Syst., vol. 13, no. 4, pp. 1717–1726, Dec. 2012.
- [144] BBC Asia. (2014). How do you Track a Plane?. [Online]. Available: <http://www.bbc.com/news/world-asia-pacific-26544554>
- [145] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," ACM Comput. Surv., vol. 38, no. 4, p. 13, 2006.
- [146] P. Xu, F. Davoine, H. Zhao, and T. Dencœux, "Multimodal information fusion for urban scene understanding," Mach. Vis. Appl., vol. 27, no. 3, pp. 331–349, 2014.
- [147] D. Nuss, M. Thom, A. Danzer, and K. Dietmayer, "Fusion of laser and monocular camera data in object grid maps for vehicle environment perception," in Proc. 17th Int. Conf. Inf. Fusion, Jul. 2014, pp. 1–8.
- [148] A. Tsukada and A. Background, "Road structure based scene understanding for Intelligent Vehicle Systems," in Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst., Oct. 2010, pp. 5557–5562.
- [149] Q. Hu, P. Wang, C. Shen, and F. Porikli, "Pushing the limits of deep CNNs for pedestrian detection," in Proc. Comput. Vis. Pattern Recognit., 2016, pp. 123–144.
- [150] C. Premebida, J. Batista, and U. Nunes, "Pedestrian detection combining RGB and dense LIDAR data," in Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS), Sep. 2014, pp. 4112–4117.
- [151] H. Cho, Y. Seo, B. V. K. V. Kumar, and R. R. Rajkumar, "A multi-sensor fusion system for moving object detection and tracking in urban driving environments," in Proc. IEEE Int. Conf. Robot. Autom., Jun. 2014, pp. 1836–1843.
- [152] N. Chumerin and M. M. van Hulle, "Cue and sensor fusion for independent moving objects detection and description in driving scenes," in Signal Processing Techniques for Knowledge Extraction and Information Fusion. Springer, LLC, 2008, pp. 161–180.
- [153] C. Häne, T. Sattler, and M. Pollefeys, "Obstacle detection for selfdriving cars using only monocular cameras and wheel odometry," in Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst., Oct. 2015, pp. 5101–5108.
- [154] Y. Zhao, J. Li, L. Li, M. Zhang, and L. Guo, "Environmental perception and sensor data fusion for unmanned ground vehicle," Math. Probl. Eng., vol. 2013, Oct. 2013, Art. no. 903951.
- [155] A. S. Huang, "Lane estimation for autonomous vehicles using vision and LIDAR," M.S. thesis, Massachusetts Inst. Technol., Feb. 2010, pp. 1–114.
- [156] D. Z. Morris, "Google self-driving car in serious crash in Mountain View," Fortune, 2016.
- [157] S. Levin and N. Woolf, "Tesla driver killed while using autopilot was watching Harry Potter, witness says," Guardian, 2016.
- [158] H. P. Levy. (2015). Top 10 Technology Trends Signal the Digital Mesh, Gartner. [Online]. Available: <http://www.gartner.com/smarterwithgartner/top-ten-technology-trends-signal-the-digital-mesh/>
- [159] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, May 2015.
- [160] H. Wang and B. Raj, "A survey: Time travel in deep learning space: An introduction to deep learning models and how deep learning models evolved from the initial ideas," Cornell Univ. Library, pp. 1–43, Nov. 2015.

- [161] L. Deng, "A tutorial survey of architectures, algorithms, and applications for deep learning," *APSIPA Trans. Signal Inf. Process.*, vol. 3, no. e2, pp. 1–29, Jan. 2014.
- [162] Y. Bengio, "Learning deep architectures for AI," *Found. Trends Mach. Learn.*, vol. 2, no. 1, pp. 1–127, 2009.
- [163] J. Ngiam, A. Khosla, M. Kim, J. Nam, H. Lee, and A. Y. Ng, "Multimodal deep learning," in *Proc. 28th Int. Conf. Mach. Learn.*, 2011, pp. 689–696.
- [164] H. Martínez and G. Yannakakis, "Deep multimodal fusion: Combining discrete events and continuous signals," in *Proc. 16th Int. Conf. Multimodal Interact.*, 2014, pp. 34–41.
- [165] N. Audebert, B. Le Saux, and S. Lefèvre. (Sep. 2016). "Semantic segmentation of earth observation data using multimodal and multi-scale deep networks." [Online]. Available: <https://arxiv.org/abs/1609.06846>
- [166] L. Zhang and H. Gao, "A deep learning-based multi-sensor data fusion method for degradation monitoring of ball screws," in *Proc. Prognostics Syst. Health Manage. Conf. (PHM-Chengdu)*, Oct. 2016, pp. 1–6.
- [167] P. O'Connor, D. Neil, S. C. Liu, T. Delbruck, and M. Pfeiffer, "Real-time classification and sensor fusion with a spiking deep belief network," *Front. Neurosci.*, vol. 7, no. 7, pp. 1–13, Oct. 2013.
- [168] E. Park, X. Han, T. L. Berg, and A. C. Berg, "Combining multiple sources of knowledge in deep CNNs for action recognition," *J. Chem. Inf. Model.*, vol. 53, no. 9, pp. 1689–1699, 2013.
- [169] X. Tao, D. Kong, Y. Wei, and Y. Wang, "A big network traffic data fusion approach based on fisher and deep auto-encoder," *Information*, vol. 7, no. 2, p. 20, 2016.
- [170] D. Neil and S.-C. Liu, "Effective sensor fusion with event-based sensors and deep network architectures," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2016, pp. 2282–2285.
- [171] J. Wagner, V. Fischer, M. Herman, and S. Behnke, "Multispectral pedestrian detection using deep fusion convolutional neural networks," in *Proc. 24th Eur. Symp. Artif. Neural Netw., Comput. Intell. Mach. Learn.*, Apr. 2016, pp. 509–514.
- [172] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of Things for smart cities," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 22–32, Feb. 2014.
- [173] H. Arasteh et al., "IoT-based smart cities: A survey," in *Proc. IEEE 16th Int. Conf. Environ. Elect. Eng. (EEEIC)*, Jun. 2016, pp. 1–6.
- [174] A. J. Jara, D. Genoud, and Y. Bocchi, "Big data for smart cities with KNIME a real experience in the SmartSantander testbed," *Softw.-Pract. Exp.*, vol. 45, no. 8, pp. 1145–1160, Aug. 2015.
- [175] K. Pollhammer, T. Novak, P. Raich, W. Kastner, A. Treytl, and G. Kovacs, "Open traffic data platform for scenario-based control," in *Proc. 42nd Annu. Conf. IEEE Ind. Electron. Soc. (IECON)*, Oct. 2016, pp. 4677–4682.
- [176] A. Kalyanaraman and K. Whitehouse, "An event-based data fusion algorithm for smart cities," in *Proc. Adjunct ACM Int. Joint Conf. Pervasive Ubiquitous Comput., Int. Symp. Wearable Comput.*, 2015, pp. 1575–1582.
- [177] J. Rivera-Mejía, C. Seáñez-Hernández, A. Hernández-López, and A. López-Pérez, "Intelligent sensor with data fusion to improve the care and management of water," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf.*, May 2011, pp. 26–30.
- [178] J. Rivera-Mejía, J. Rivera-Jacquez, C. Modesto-Acosta, and M. Tena-Vega, "Auto-compensated smart sensor with F-ISE to continuous monitoring in public water supplies," in *Proc. IEEE Int. Instrum. Meas. Technol. Conf. (I2MTC)*, May 2012, pp. 2628–2633.
- [179] N.-B. Chang and S. Imen, "Multi-sensor acquisition, data fusion, criteria mining and alarm triggering for decision support in urban water infrastructure systems," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2015, pp. 539–544.
- [180] G. Bello-Orgaz, J. J. Jung, and D. Camacho, "Social big data: Recent achievements and new challenges," *Inf. Fusion*, vol. 28, pp. 45–59, Mar. 2016.
- [181] X. Z. Wang, X. L. Bi, Z. Q. Ge, and L. Li, "Deep data fusion model for risk perception and coordinated control of smart grid," in *Proc. Int. Conf. Estim. Detect. Inf. Fusion (ICEDIF)*, Jan. 2015, pp. 110–113.

- [182] S. Li and X. Wang, "Cooperative change detection for voltage quality monitoring in smart grids," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 1, pp. 86–99, Jan. 2016.
- [183] S. P. Talebi, S. Kanna, Y. Xia, and D. P. Mandic, "A distributed quaternion Kalman filter with applications to fly-by-wire systems," in *Proc. IEEE Int. Conf. Digit. Signal Process. (DSP)*, Oct. 2016, pp. 30–34.
- [184] X. Sevillano, E. Màrmol, and V. Fernandez-Arguedas, "Towards smart traffic management systems: Vacant on-street parking spot detection based on video analytics," in *Proc. 17th Int. Conf. Inf. Fusion (FUSION)*, Jul. 2014, pp. 1–8.
- [185] Y. Wang, B. Liang, W. Zheng, L. Huang, and H. Liu, "The development of a smart taxicab scheduling system: A multi-source data fusion perspective," in *Proc. IEEE 16th Int. Conf. Data Mining (ICDM)*, Dec. 2016, pp. 1275–1280.
- [186] L. Yu, D. Shao, and H. Wu, "Next generation of journey planner in a smart city," *Proc. IEEE Int. Conf. Data Mining Workshop (ICDMW)*, Nov. 2015, pp. 422–429.
- [187] P. J. F. White, B. W. Podaima, and M. R. Friesen, "Algorithms for smartphone and tablet image analysis for healthcare applications," *IEEE Access*, vol. 2, pp. 831–840, Aug. 2014.
- [188] D. De, P. Bharti, S. K. Das, and S. Chellappan, "Multimodal wearable sensing for fine-grained activity recognition in healthcare," *IEEE Internet Comput.*, vol. 19, no. 5, pp. 26–35, Sep./Oct. 2015.
- [189] Y. Chen, W. Gao, S. Wang, and S. Jiao, "Demo abstract: WearableHUB: An open pervasive wearable data fusion platform for personal health management," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops (PerCom Workshops)*, Mar. 2015, pp. 217–219.
- [190] Z. Alazawi, O. Alani, M. B. Abdljabar, S. Altowaijri, and R. Mehmood, "A smart disaster management system for future cities," in *Proc. ACM Int. Workshop Wireless Mobile Technol. Smart Cities (WiMobCity)*, 2014, pp. 1–10.
- [191] Z. Alazawi, M. B. Abdljabar, S. Altowaijri, A. M. Vegni, and R. Mehmood, *ICDMS: An Intelligent Cloud Based Disaster Management System for Vehicular Networks (Lecture Notes in Computer Science)*, vol. 7266. Vilnius, Lithuania: Springer, 2012.
- [192] G. Ayres and R. Mehmood, "On discovering road traffic information using virtual reality simulations," in *Proc. 11th Int. Conf. Comput. Modelling Simulation*, Mar. 2009, pp. 411–416.
- [193] J. Schlingensiepen, F. Nemtanu, R. Mehmood, and L. McCluskey, "Autonomic transport management systems—Enabler for smart cities, personalized medicine, participation and industry grid/industry 4.0," in *Intelligent Transportation Systems—Problems and Perspectives*, A. Sladkowski and W. Pamula, Eds. London, U.K.: Springer, 2016, pp. 3–35.
- [194] J. Schlingensiepen, R. Mehmood, and F. C. Nemtanu, "Framework for an autonomic transport system in smart cities," *Cybern. Inf. Technol.*, vol. 15, no. 5, pp. 50–62, 2015.
- [195] R. Mehmood and G. Graham, "Big data logistics: A health-care transport capacity sharing model," *Procedia Comput. Sci.*, vol. 64, pp. 1107–1114, 2015.
- [196] R. Mehmood and J. A. Lu, "Computational Markovian analysis of large systems," *J. Manuf. Technol. Manag.*, vol. 22, no. 6, pp. 804–817, 2011.
- [197] R. Mehmood, R. Meriton, G. Graham, P. Hennelly, and M. Kumar, "Exploring the influence of big data on city transport operations: A Markovian approach," *Int. J. Oper. Prod. Manag.*, vol. 37, no. 1, pp. 75–104, 2016.
- [198] G. Graham, R. Mehmood, and E. Coles, "Exploring future cityscapes through urban logistics prototyping: A technical viewpoint," *Supply Chain Manage., Int. J.*, vol. 20, no. 3, pp. 341–352, 2015.
- [199] G. Ayres and R. Mehmood, *LocPriS: A Security and Privacy Preserving Location Based Services Development Framework (Lecture Notes in Artificial Intelligence)*, vol. 6279. Heidelberg, Germany: Springer, 2010.
- [200] R. Mehmood, R. Alturki, and S. Zeadally, "Multimedia applications over metropolitan area networks (MANs)," *J. Netw. Comput. Appl.*, vol. 34, no. 5, pp. 1518–1529, Sep. 2011.
- [201] R. Mehmood and R. Alturki, "A scalable multimedia qos architecture for ad hoc networks," *Multimed. Tools Appl.*, vol. 54, no. 3, pp. 551–568, 2011.

- [202] R. Mehmood, M. A. Faisal, and S. Altowaijri, "Future Networked Healthcare Systems: A Review and Case Study," in *Big Data: Concepts, Methodologies, Tools, Application*, I. Management Association. Hershey, PA, USA: IGI Global, 2016, pp. 2429–2457.
- [203] R. Alturki, K. Nwizege, R. Mehmood, and M. Faisal, "End to end wireless multimedia service modelling over a metropolitan area network," in *Proc. 11th Int. Conf. Comput. Modelling Simulation (UKSim)*, Mar. 2009, pp. 532–537.
- [204] G. Graham and R. Mehmood, "The strategic prototype 'crime-sourcing' and the science/science fiction behind it," *Technol. Forecast. Soc. Change*, vol. 84, pp. 86–92, 2014.
- [205] R. Mehmood and M. Nekovee, "Vehicular AD HOC and grid networks: Discussion, design and evaluation," in *Proc. 14th World Congr. Intell. Transp. Syst. (ITS)*, vol. 2. 2007, pp. 1555–1562.
- [206] L. A. Tawalbeh, A. Basalamah, R. Mehmood, and H. Tawalbeh, "Greener and smarter phones for future cities: Characterizing the impact of GPS signal strength on power consumption," *IEEE Access*, vol. 4, pp. 858–868, 2016.
- [207] L. A. Tawalbeh, R. Mehmood, E. Benkhelifa, and H. Song, "Mobile cloud computing model and big data analysis for healthcare applications," *IEEE Access*, vol. 4, pp. 6171–6180, 2016.
- [208] F. Piccialli and J. E. Jung, "Understanding customer experience diffusion on social networking services by big data analytics," *Mobile Netw. Appl.*, pp. 1–8, Dec. 2016.
- [209] M. Büscher et al., "Intelligent mobility systems: Some socio-technical challenges and opportunities," in *Communications Infrastructure. Systems and Applications in Europe (Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering)*, vol. 16. Heidelberg, Germany: Springer, 2009, pp. 140–152.
- [210] A. Chianese, F. Marulli, F. Piccialli, P. Benedusi, and J. E. Jung, "An associative engines based approach supporting collaborative analytics in the Internet of cultural things," *Future Generat. Comput. Syst.*, vol. 66, pp. 187–198, Jan. 2017.
- [211] A. Chianese and F. Piccialli, *Improving User Experience of Cultural Environment Through IoT: The Beauty or the Truth Case Study*. Cham, Switzerland: Springer, 2015, pp. 11–20.
- [212] S. Cuomo, P. De Michele, F. Piccialli, A. Galletti, and J. E. Jung, "IoT-based collaborative reputation system for associating visitors and artworks in a cultural scenario," *Expert Syst. Appl.*, vol. 79, pp. 101–111, Aug. 2017.
- [213] I. M. Pires, N. Garcia, and F. Flórez-Revuelta, "Multi-sensor data fusion techniques for the identification of activities of daily living using mobile devices," in *Proc. ECML PKDD*, 2015, pp. 1–10.
- [214] I. M. Pires, N. M. Garcia, N. Pombo, and F. Flórez-Revuelta, "Identification of activities of daily living using sensors available in off-the-shelf mobile devices: Research and hypothesis," *Adv. Intell. Syst. Comput.*, vol. 476, pp. 121–130, May 2016.