دانشگاه صنعتی شیراز

دانشکده مهندسی برق و الکترونیک گروه کنترل

پایان‌نامه کارشناسی ارشد

در رشته مهندسی برق گرایش کنترل

تخمین و دسته‌بندی داده‌ها در شبکه‌های سنسوری با استفاده از عامل‌های متحرک

نگارش:

کاظم حسن پور

استاد راهنما:

دکتر بهروز صفری نژادیان

استاد مشاور:

دکتر مختار شاصادقی

مهرماه ۱۳۹۲

بسم الله الرحمن الرحيم

بسمه تعالی

تخمین و دسته‌بندی داده‌ها در شبکه‌های سنسوری با استفاده از عامل‌های متحرک

پایان‌نامه ارائه شده به عنوان بخشی از فعالیتهای تحصیلی

نگارش:
کاظم حسن پور

برای اخذ درجه کارشناسی ارشد

گروه کنترل دانشکده مهندسی برق و الکترونیک
دانشگاه صنعتی شیراز

ارزیابی پایان‌نامه توسط هیات داوران با درجه: عالی

دکتر بهروز صفری نژادیان استادیار در رشته مهندسی کنترل (استاد راهنما)
دکتر مختار شاصادقی استادیار در رشته مهندسی کنترل (استاد مشاور)
دکتر جعفر زارعی استادیار در رشته مهندسی کنترل (داور)
دکتر طاهره بینازاده استادیار در رشته مهندسی کنترل (داور)

مدیر امور آموزشی و تحصیلات تکمیلی دانشگاه:

حق چاپ محفوظ و مخصوص به دانشگاه صنعتی شیراز است.

تأییدیه‌ی صحت و اصالت نتایج

اینجانب کاظم حسن پور دانشجوی رشته مهندسی برق کنترل مقطع تحصیلی کارشناسی ارشد به شماره دانشجویی ۹۰۱۱۴۰۴۴ تأیید می‌نماید کلیه نتایج این پایان نامه/رساله، بدون هیچگونه دخل و تصرف ، حاصل مستقیم پژوهش صورت گرفته توسط اینجانب است. در مورد اقتباس مستقیم و غیر مستقیم از سایر آثار علمی، اعم از کتاب، مقاله، پایان نامه با رعایت امانت و اخلاق علمی، مشخصات کامل منبع مذکور درج شده است.

در صورت اثبات خلاف مندرجات فوق، به تشخیص مقامات ذی صلاح دانشگاه صنعتی شیراز، مطابق قوانین و مقررات مربوط و آئین نامه های آموزشی، پژوهشی و انضباطی عمل خواهد شد و اینجانب حق هرگونه اعتراض و تجدیدنظر را، نسبت به رأی صادره، از خود ساقط می کند. همچنین، هرگونه مسئولیت ناشی از تخلف نسبت به صحت و اصالت نتایج مندرج در پایان نامه/رساله در برابر اشخاص ذی نفع (اعم از حقیقی و حقوقی) و مراجع ذی صلاح (اعم از اداری و قضایی) متوجه اینجانب خواهد بود و دانشگاه صنعتی شیراز هیچ گونه مسئولیتی دراین زمینه نخواهند داشت.

تبصره ۱- کلیه حقوق مادی این اثر متعلق به دانشگاه صنعتی شیراز است.

تبصره ۲- اینجانب تعهد می نماید بدون اخذ مجوز از دانشگاه صنعتی شیراز دستاوردهای این پایان نامه/ رساله را منتشر نکند و یا در اختیار دیگران قرار ندهد.

نام و نام خانوادگی دانشجو: کاظم حسن پور

تاریخ و امضاء

مجوز بهره‌برداری از پایان‌نامه

کلیه حقوق مادی و معنوی مترتب بر نتایج پایان‌نامه متعلق به دانشگاه و انتشار نتایج نیز تـابع مقرارت دانشگاهی است و با موافقت استاد راهنما به شرح زیر، بلامانع است:

☐ بهره‌برداری از این پایان‌نامه/ رساله برای همگان بلامانع است.

☐ بهره‌برداری از این پایان‌نامه/ رساله با اخذ مجوز از استاد راهنما، بلامانع است.

☐ بهره‌برداری از این پایان‌نامه/ رساله تا تاریخ ..... ممنوع است.

نام استاد راهنما: دکتر بهروز صفری نژادیان

تاریخ:

امضا:

تقدیم به:

آنان که مهر آسمانی‌شان آرام‌بخش زمینی‌ام است، به استوارترین تکیه‌گاهم، دستان پر مهر پدرم، به سبزترین نگاه زندگیم، چشمان سبز مادرم که هر چه آموختم در مکتب عشق شما آموختم و هر چه بکوشم قطره‌ای از دریای بی‌کران مهربانیتان را سپاس نتوانم بگویم. امروز هستی‌ام به امید شماست و فردا کلید باغ بهشتم رضای شما را آوردی گران سنگ‌تر از این ارزان نداشتم تا به خاک پایتان نثار کنم، باشد که حاصل تلاشم نسیم گونه غبار خستگیتان را بزداید.

بوسه بر دستان پر مهرتان

تشکر و قدردانی:

اکنون که این رساله به پایان رسیده است برخود فرض می‌دانم که از اساتید ارجمند جناب آقای دکتر بهروز صفری‌نژادیان و جناب آقای دکتر مختار شاصادقی کمال تشکر و قدردانی را داشته باشم که با نکته‌های دلاویز و گفته‌های بلند، صحیفه‌های سخن را علم پرور نموده و همواره راهنما و راه گشای نگارنده در اتمام و اکمال پایان‌نامه بوده‌اند..

معلما مقامت ز عرش برتر باد همیشه توسن اندیشه‌ات مظفر باد

چکیده

تخمین و دسته‌بندی داده‌ها در شبکه‌های سنسوری با استفاده از عامل‌های متحرک

نگارش:

کاظم حسن پور

در این پایان‌نامه، از چندین الگوریتم توزیع شده در جهت تخمین متغیرهای حالت یک سیستم خطی و دسته‌بندی داده‌ها در یک شبکه سنسوری با چندین حلقه استفاده شده است. در اینجا، دو روش جدید *MABDKF1* و *MABDKF2* جهت رسیدن به تخمینی بهینه از حالتهای یک سیستم خطی با وجود خطا در سنسورها پیشنهاد می‌شود. در الگوریتم‌های پیشنهادی، عامل‌های متحرک نقش اصلی و اساسی را بازی می‌کنند بطوریکه آنها الگوریتم فیلتر کالمن خطی را بر روی یک سیستم خطی اجراء کرده تا به تخمینی خوب از متغیرهای حالت سیستم برسند. همچنین به کمک دو الگوریتم دیگر با نام‌های *MABDEM1* و *MABDEM2*، پارامترهای یک تابع چگالی احتمال گوسی تخمین زده شده و داده‌های موجود در یک شبکه سنسوری با چندین حلقه دسته‌بندی می‌شود. در این دو الگوریتم نیز از عاملهای متحرک در جهت اجراء الگوریتم *EM* توزیع شده استفاده می‌شود. در ادامه همگرایی این دو الگوریتم مورد بررسی قرار خواهد گرفت. در پایان، عملکرد خوب این الگوریتم‌ها با استفاده از داده‌های شبیه سازی شده نشان داده خواهد شد.

واژه‌های کلیدی: عامل‌های متحرک، مرکز ترکیب، الگوریتم *EM*، شبکه سنسوری و تخمین تابع چگالی.

فهرست مطالب

۱.فصل اول: مقدمه **۱**

۱-۱- مقدمه ۲

۱-۲- هدف تحقیق و اهمیت آن ۴

۱-۳- بخش‌های پایان‌نامه ۵

۲.فصل دوم: مروری بر تحقیقات انجام شده **۷**

۲-۱- پژوهش‌های انجام شده در زمینه الگوریتم‌های متمرکز و نامتمرکز ۸

۲-۲- تحقیقات انجام شده درباره داده‌کاوی و الگوریتم EM ۸

۲-۳- بیشینه عامل‌های متحرک ۱۰

۲-۴- جایگاه این پایان‌نامه در بین پژوهش‌های انجام شده ۱۱

۳.فصل سوم: عامل‌های متحرک **۱۳**

۳-۱- مقدمه ۱۴

۳-۲- شبکه‌های سنسوری توزیع شده ۱۵

۳-۲-۱- اجزاء تشکیل دهنده یک گره سنسوری توزیع شده ۱۶

۳-۲-۲- انواع توپولوژی‌ها در شبکه‌های سنسوری ۱۷

۳-۳- عامل متحرک ۲۲

۳-۳-۱- خصوصیات عامل‌های نرم‌افزاری ۲۳

۳-۳-۲- کاربردهای عامل‌های متحرک ۲۵

۳-۳-۳- عامل‌های متحرک ۲۷

۳-۳-۴- اجزاء تشکیل دهنده یک عامل متحرک ۲۹

۳-۴- انواع مدلهای محاسباتی ۳۰

۳-۴-۱- مدل محاسباتی Client/Server ۳۱

۳-۴-۲- مدل محاسباتی براساس عامل‌های متحرک ۳۲

۳-۵- سیستمهای چند عاملی ۳۳

۳-۶- ترکیب داده‌ها ۳۶

۴.فصل چهارم: تخمین حالت با عامل‌های متحرک **۴۱**

۴-۱- مقدمه ۴۲

۴-۲- الگوریتم فیلتر کالمن ۴۳

۴-۳- بیان مسئله ۴۵

۴-۴- الگوریتم فیلتر کالمن نامتمرکز با عامل‌های متحرک نوع ۱ ..... ۴۷

۴-۴-۱- مراحل اجرایی الگوریتم فیلتر کالمن نامتمرکز با عامل‌های متحرک نوع ۱ ..... ۴۹

۴-۵- الگوریتم فیلتر کالمن نامتمرکز با عامل‌های متحرک نوع ۲ ..... ۵۰
۴-۵-۱- مراحل اجرایی الگوریتم فیلتر کالمن نامتمرکز با عامل‌های متحرک نوع ۲..... ۵۱
۴-۵-۲- اثبات بهینگی الگوریتم فیلترکالمن نامتمرکز با عامل‌های متحرک نوع ۲..... ۵۲
۴-۶- نتایج شبیه‌سازی ۵۴
۴-۷- نتیجه‌گیری ۶۱

۵.فصل پنجم: دسته‌بندی داده‌ها با عامل‌های متحرک **۶۲**
۵-۱- مقدمه ۶۳
۵-۲- الگوریتم *EM* برای ترکیبات گوسی ۶۴
۵-۳- بیان مسئله ۶۷
۵-۴- الگوریتم *EM* توزیع شده براساس عامل‌های متحرک نوع ۱ ..... ۶۸
۵-۴-۱- مراحل اجرایی الگوریتم EM توزیع شده براساس عامل‌های متحرک نوع ۱..... ۷۱
۵-۵- الگوریتم EM توزیع شده براساس عامل‌های متحرک نوع ۲..... ۷۲
۵-۵-۱- مراحل اجرایی الگوریتم EM توزیع شده براساس عامل‌های متحرک نوع ۲..... ۷۳
۵-۶- اثبات همگرایی الگوریتم‌های پیشنهادی ۷۴
۵-۷- تخمین مرتبه بصورت توزیع شده در الگوریتم‌های پیشنهادی با استفاده از الگوریتم EM ۷۶
۵-۷-۱- تخمین مرتبه بصورت توزیع شده ۷۸
۵-۸- نتایج شبیه سازی ۸۰
۵-۸-۱- دسته‌بندی داده‌ها با معلوم بودن تعداد دسته‌ها..... ۸۰
۵-۸-۲- دسته‌بندی داده‌ها با عدم مشخص بودن تعداد دسته‌ها ۹۰
۵-۹- نتیجه‌گیری ۹۴

۶.فصل ششم: جمع‌بندی و پیشنهادها **۹۵**
۶-۱- نتیجه‌گیری ۹۶
۶-۲- نوآوری‌ها..... ۹۶
۶-۳- پیشنهادات ۹۷

مراجع **۹۸**

فهرست شکل‌ها

شکل ۲-۱ شبکه سنسوری بی‌سیم..... ۱۵
شکل ۲-۲ بلوک دیاگرام یک گره سنسوری..... ۱۶
شکل ۲-۳ توپولوژی ستاره..... ۱۸
شکل ۲-۴ توپولوژی حلقوی..... ۱۸
شکل ۲-۵ توپولوژی مش..... ۱۹
شکل ۲-۶ توپولوژی درختی..... ۲۰
شکل ۲-۷ توپولوژی درختی..... ۲۰
شکل ۲-۸ توپولوژی باس..... ۲۱
شکل ۲-۹ توپولوژی اتصالات کامل..... ۲۱
شکل ۲-۱۰ یک دید انتزاعی از یک عامل..... ۲۳
شکل ۲-۱۱ عامل‌های متحرک در شبکه سنسوری..... ۲۷
شکل ۲-۱۲ ساختار یک عامل متحرک..... ۳۰
شکل ۲-۱۳ مدل محاسباتی بر اساس client/server..... ۳۱
شکل ۲-۱۴ انواع مدلهای محاسباتی بر اساس client/server..... ۳۲
شکل ۲-۱۵ مدل محاسباتی براساس عامل‌های متحرک..... ۳۳
شکل ۲-۱۶ کاربردها و تکنیک های مورد استفاده در ترکیب داده ها..... ۳۹
شکل ۲-۱۷ انواع توپولوژی های مورد استفاده در ترکیب داده ها..... ۴۰
شکل ۳-۱ شبکه ای سنسوری با چندین حلقه و تعدادی عامل متحرک..... ۴۵
شکل ۳-۲ الگوریتم فیلتر کالمن نا متمرکز با عامل متحرک نوع ۱..... ۴۷
شکل ۳-۳ الگوریتم فیلتر کالمن نا متمرکز با عامل‌های متحرک نوع ۲..... ۵۰
شکل ۳-۴ مقایسه کوواریانس خطای فیلترینگ موقعیت $S(t)$ توسط انواع فیلترها..... ۵۵
شکل ۳-۵ مقایسه کوواریانس خطای فیلترینگ سرعت $\dot{S}(t)$ توسط انواع فیلترها..... ۵۶
شکل ۳-۶ مقایسه کوواریانس خطای فیلترینگ شتاب $\ddot{S}(t)$ توسط انواع فیلترها..... ۵۶
شکل ۳-۷ مقایسه ردیابی حالت x_1 توسط فیلتر های متمرکز و نامتمرکز نوع ۱..... ۵۷
شکل ۳-۸ مقایسه ردیابی حالت x_2 توسط فیلتر های متمرکز و نامتمرکز نوع ۱..... ۵۸
شکل ۳-۹ مقایسه ردیابی حالت x_3 توسط فیلتر های متمرکز و نامتمرکز نوع ۱..... ۵۸
شکل ۳-۱۰ مقایسه ردیابی حالت x_1 توسط فیلتر های متمرکز و نامتمرکز نوع ۲..... ۵۹
شکل ۳-۱۱ مقایسه ردیابی حالت x_2 توسط فیلتر های متمرکز و نامتمرکز نوع۲..... ۶۰

شکل ۳-۱۲ مقایسه ردیابی حالت x_3 توسط فیلتر های متمرکز و نامتمرکز نوع۲..... **۶۰**
شکل ۴-۱ شبکه سنسوری با چندین حلقه، تعدادی عامل متحرک و یک مرکز پردازش..... **۶۸**
شکل ۴-۲ شبکه سنسوری با چندین حلقه، تعدادی عامل متحرک و یک مرکز پردازش..... **۸۰**
شکل ۴-۳ تخمین تابع چگالی احتمال با استفاده از الگوریتم *MABDEM1*..... **۸۲**
شکل ۴-۴ تخمین تابع چگالی احتمال با استفاده از الگوریتم *MABDEM2*..... **۸۲**
شکل ۴-۵ تعداد تکرار در الگوریتم‌های پیشنهادی با تعداد داده‌های مختلف در هر گره..... **۸۳**
شکل ۴-۶ نحوه توزیع داده ها..... **۸۴**
شکل ۴-۷ مقادیر میانگین تخمین زده شده با استفاده از الگوریتم *EM* استاندارد..... **۸۵**
شکل ۴-۸ مقادیر میانگین تخمین زده شده با استفاده از الگوریتم*MABDEM1*..... **۸۵**
شکل ۴-۹ مقادیر میانگین تخمین زده شده با استفاده از الگوریتم*MABDEM2* **۸۵**
شکل ۴-۱۰ مقادیر لگاریتم *Likelihood* الگوریتم‌های *MABDEM1* و *EM* استاندارد..... **۸۸**
شکل ۴-۱۱ مقادیر لگاریتم *Likelihood* الگوریتم‌های *MABDEM2* و *EM* استاندارد..... **۸۹**
شکل ۴-۱۲ توابع چگالی گوسی تخمین زده شده با استفاده از الگوریتم‌های پیشنهادی با تخمین تعداد دسته‌ها..... **۸۹**
شکل ۴-۱۳ تخمین تابع چگالی احتمال با استفاده از الگوریتم *MABDEM1* با تخمین تعداد دسته‌ها..... **۹۱**
شکل ۴-۱۴ تخمین تابع چگالی احتمال با استفاده از الگوریتم *MABDEM2* با تخمین تعداد دسته‌ها..... **۹۲**

فهرست جدول‌ها

جدول ۴-۱ مقادیر تخمین زده شده با الگوریتم‌های پیشنهادی برای داده‌های یک بعدی..... ۸۱
جدول ۴-۲ مقادیر تخمین زده شده با الگوریتم *MABDEM1* برای داده‌های دو بعدی..... ۸۶
جدول ۴-۳ مقادیر تخمین زده شده با الگوریتم *MABDEM2* برای داده‌های دو بعدی..... ۸۷
جدول ۴-۴ مقادیر تخمین زده شده با الگوریتم *EM* استاندارد برای داده‌های دو بعدی..... ۸۸
جدول ۴-۵ مقادیر تخمین زده شده با الگوریتم‌های پیشنهادی برای داده‌های یک بعدی با تخمین تعداد دسته‌ها..... ۹۱
جدول ۴-۶ مقادیر تخمین زده شده با الگوریتم *MABDEM1* برای داده‌های دو بعدی با تخمین تعداد دسته‌ها ۹۳
جدول ۴-۷ مقادیر تخمین زده شده با الگوریتم *MABDEM2* برای داده‌های دو بعدی با تخمین تعداد دسته‌ها ۹۳

فهرست نشانه‌های اختصاری

بهره کالمن	K_m
تابع چگالی احتمال مربوط به بردار تصادفی $\mathbf{Y}$ که با بردار پارامترهای $\boldsymbol{\theta}$ مشخص می‌گردد	$\mathbf{Y}$
زمان	t
زمان گسسته	k
عملگر امید ریاضی	E
ماتریس کواریانس نویز اندازه‌گیری	R_k
ماتریس کواریانس نویز سیستم	Q_k
ماتریس کواریانس همبستگی بین نویزهای سیستم و اندازه‌گیری	M_k
مجموعه داده‌های تمام گره‌های شبکه	$\boldsymbol{\mathcal{Y}} = \{\boldsymbol{y}_m\}_{m=1}^{M}$
مجموعه داده‌های گره m شامل N_m داده مختلف	$\boldsymbol{\mathcal{Y}}_m = \{\boldsymbol{y}_{m,i}\}_{i=1}^{N_m}$
یک بردار تصادفی d- بعدی با تابع چگالی احتمال $f(y)$	$\mathbf{Y}$
یک بردار تصادفی d- بعدی متناظر با گره m از شبکه	$\mathbf{Y}_m$
یک تابع چگالی احتمال گوسی (نرمال) با بردار میانگین μ و ماتریس کوواریانس $\sum$	$\mathcal{N}(\boldsymbol{y} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma})$
یک تحقق از بردار تصادفی $\mathbf{Y}$	$\boldsymbol{y_i}$
یک تحقق از بردار تصادفی $\mathbf{Y}_m$	$\boldsymbol{y}_{m,i}$
یک مجموعه داده شامل N مقدار تولید شده بر مبنای $f(y)$	$\boldsymbol{\mathcal{Y}} = \{\boldsymbol{y}_i\}_{i=1}^{N}$

فهرست کلمات اختصاری

CPU	Central Processing Unit
DDM	Distributed Data Mining
DEM	Distributed EM
EM	Expectation Maximization
FC	Fusion Center
GPS	Global position system
MA	Mobile agent
MABDEM	Mobile Agent Based Distributed Expectation Maximization
MABDKF	Mobile Agent Based Distributed Kalman Filter
ML	Maximum Likelihood
pdf	Probability density function
WSNs	Wireless Sensor Networks

فصل اول: مقدمه

۱-۱-مقدمه

شبکه‌های سنسوری بی‌سیم[1] بطور گسترده در ساختمانها، صنایع، ادارات و ... بکار گرفته می‌شوند. این شبکه‌های سنسوری از تعدادی گره با سنسورهای مختلف تشکیل شده‌اند که از طریق توپولوژی‌های مختلفی مانند ستاره، مش، حلقه و غیره به هم متصل می‌شوند. چنانچه این گره‌های سنسوری، بصورت توزیع شده در ناحیه‌ای پخش شوند به آنها شبکه‌های سنسوری توزیع شده[2] گفته می‌شوند که از جمله آنها می‌توان به اینترنت، اینترانت و غیره اشاره کرد[۱].

وقتی حالتهای یک سیستم از طریق چندین سنسور موجود در یک شبکه سنسوری اندازه‌گیری می‌شوند. جهت ترکیب حالتهای این سیستم، از الگوریتم‌های مختلفی از جمله الگوریتم فیلتر کالمن استفاده می‌شوند. این الگوریتم‌ها با ترکیب داده‌های سنسورها به تخمینی بهینه از حالتهای این سیستم می‌رسند که این تخمین بهینه با افزایش تعداد داده‌های اندازه‌گیری شده از سیستم باعث کاهش خطا و افزایش دقت می‌شود.

معمولا ترکیب اطلاعات در شبکه‌های سنسوری با استفاده از الگوریتم‌های متمرکز[3] صورت می‌گیرد. در این الگوریتم‌ها، از توپولوژی ستاره برای انتقال حجم زیادی از داده‌های خام از گره‌ها به مرکز پردازش[4] استفاده می‌شود بطوریکه داده‌های خام از گره‌های سنسوری به مرکز پردازش انتقال داده شده و درآنجا الگوریتم مورد نظر مانند فیلتر کالمن بر روی داده‌های سنسوری اعمال شده و تخمین بهینه حاصل می‌شود. از جمله معایب این روش می‌توان به ایجاد ترافیک در شبکه بدلیل انتقال حجم زیادی از داده‌های خام به مرکز پردازش و از کار افتادن کل شبکه با مختل شدن مرکز پردازش اشاره کرد[۲].

در این پایان‌نامه برای حل این مشکلات، از الگوریتم‌های نامتمرکز[5] با استفاده از عامل‌های متحرک استفاده می‌شود. این الگوریتم‌ها در محیط های نامتمرکز(توزیع شده) مانند اینترنت، اینترانت و شبکه‌های

[1] Wireless sensor networks (WSNs)

[2] Distributed sensor networks (DSNs)

[3] Centralized algorithms

[4] Processing center

[5] Decentralized algorithms

سنسوری و انواع ساختارهای نامتمرکز به کمک عامل‌های متحرک بکار گرفته می‌شود. در این روش توجه‌ای خاص به ترکیب داده‌ها توسط عامل‌های متحرک بصورت توزیع شده به جای جمع‌آوری تمام داده‌ها در یک مرکز پردازش می‌شود. عامل‌های متحرک با حرکت میان گره‌ها، پس از پردازش داده‌های سنسوری آنها، اطلاعات پردازش شده را همراه خود به گره‌های دیگر منتقل می‌کند و از ترافیک شبکه بدلیل استفاده از اطلاعات پردازش شده بجای داده‌های خام، کاسته می‌شود. حال چنانچه یکی از گره‌ها با مشکل روبه رو شود تنها داده‌های سنسوری آن گره غیر قابل استفاده می‌باشد ولی عامل متحرک می‌تواند از اطلاعات گره‌های دیگر استفاده نماید.

دراینجا، جهت پردازش داده‌های سنسورها از الگوریتم فیلتر کالمن استفاده می شود تا بتوان حالتهای یک سیستم را تخمین زد. سپس با اجراء الگوریتم‌های مختلفی به پردازش داده‌های جمع‌آوری شده می‌پردازد. یکی از این الگوریتم‌ها که جهت تخمین داده‌ها استفاده می‌شود الگوریتم فیلتر کالمن می‌باشد. عامل‌های متحرک علاوه بر جمع‌آوری داده‌های سنسوری از گره‌ها، الگوریتم فوق را برای آنها محاسبه کرده و تخمین بهینه حاصل می‌شود. دراینجا مرکز پردازش به معنای واقعی موجود درالگوریتم‌های متمرکز نمی‌باشد بلکه تنها وظیفه تزریق عامل‌های متحرک و دریافت آنها را از شبکه برعهده دارد.

امروزه داده‌کاوی توزیع شده[1] در محیط‌های توزیع شده‌ای مانند اینترنت، شبکه‌های سنسوری با توپولوژی‌های مختلف و ... بسیار مورد توجه قرار گرفته‌اند. در داده‌کاوی توزیع شده توجه‌ای خاص به تحلیل داده‌ها بصورت توزیع شده بجای متمرکز می‌شود. در بسیاری از کاربردهای داده‌کاوی، ما بدنبال بدست آوردن پارامترهای یک مدل سراسری برای داده‌های بدست آمده با استفاده از الگوریتم‌های یادگیری هستیم[۴-۳].

توزیع آماری ترکیبی، یک ابزار ریاضی قوی و انعطاف پذیر جهت مدلسازی و دسته‌بندی داده‌های مشاهده شده در پدیده‌های تصادفی می‌باشد. در این روش، عموما داده‌ها بوسیله یک مدل ترکیبی آماری تولید شده و داده‌های ایجاد شده به دسته‌هایی متناظر با مولفه‌های این مدل ترکیبی دسته‌بندی می‌شوند[۵].

امروزه جهت مدلسازی داده‌های یک بعدی و چند بعدی از ابزارهای آماری قوی و انعطاف پذیری

[1] Distributed Data Mining (DDM)

همچون مدلهای ترکیبی (مانند مدلهای ترکیبی گوسی) استفاده می‌شود. در اینجا از ترکیبات گوسی بعنوان متداول ترین مدلهای ترکیبی در بیشتر شبیه سازی‌ها استفاده شده است. به دلیل سادگی و مشتق پذیر بودن ترکیبات گوسی، پارامترهای آنرا می‌توان با استفاده از الگوریتم EM[1] تخمین زده و همچنین می‌توان در بسیاری از فرآیندها به کمک مدلهای ترکیبی گوسی مدلسازی را انجام داد[۶].

در ادامه، از یک مدل آماری ترکیب گوسی بعنوان تابع چگالی احتمال استفاده شده تا داده‌های سنسوری در یک شبکه سنسوری توزیع شده با چندین حلقه و تعدادی عامل‌های متحرک تولید نماید. بطورکلی تابع چگالی در نظر گرفته شده بایستی علاوه برکاربردی بودن بتواند محدوده وسیعی از داده‌های واقعی در فرآیندهای آماری مختلف را در بر بگیرد. در اینجا با گوسی در نظرگرفتن مولفه‌های این مدل ترکیبی، می‌توان حجم کمی از داده‌های آماری (که ما آنرا چکیده‌های آماری می‌نامیم) را توسط عامل‌های متحرک به گره‌های دیگر انتقال داده و در عوض پارامترهای مولفه را تخمین بزنند.

در اینجا موضوع اصلی بکارگیری عامل‌های متحرک جهت اجراء الگوریتم EM برای تخمین پارامترهای یک مدل ترکیبی گوسی در جهت دسته‌بندی داده‌های سنسوری در یک شبکه چند حلقه‌ای و همچنین تخمین حالتهای یک سیستم خطی می‌باشد.

۱-۲-هدف تحقیق و اهمیت آن

از مهمترین اهداف این پایان‌نامه می‌توان به نقش عامل‌های متحرک در تخمین حالتهای یک سیستم خطی و دسته‌بندی داده‌ها در یک شبکه سنسوری چند حلقه‌ای اشاره کرد. در اینجا ابتدا به بررسی استفاده از عامل‌های متحرک در اجراء الگوریتم فیلترکالمن، جهت تخمین حالتهای یک سیستم خطی با وجود خطا در اندازه‌گیری های آن پرداخته شده و نشان می‌دهیم که به کمک عامل‌های متحرک می‌توان به تخمینی دقیق از حالتهای سیستم با حداقل کوواریانس رسیده و به تخمینی با اطمینان بالا و مقاوم در برابر وجود خطا دست پیدا کنیم. در ادامه، به دنبال انجام اهداف دیگر از این پایان‌نامه، می‌توان به چگونگی استفاده از عامل‌های متحرک در دسته‌بندی داده‌ها اشاره کرد بطوریکه عاملهای متحرک با تخمین

[1] Expectation Maximization (EM)