



Computational Intelligence in Electrical Engineering
Vol. 13, No. 4, 2023
Research Paper

A Drift-Aware Online Learner for Anomaly Detection from Streaming Data

Maryam Amoozegar¹, Behrouz Minaei-Bidgoli¹, Hadi Fanaee², Mansour Rezghi³

¹ Dept. of Computer Engineering, Iran University of Science and Technology, Tehran, Iran

² Center for Applied Intelligent Systems Research (CAISR), Halmstad University, Halmstad, Swede

³ Dept. of Computer Science, Tarbiat Modares University, Tehran, Iran

Abstract:

Streaming data has been evolved in a dynamically changing and evolving environment. Therefore, concept drift or changing the underlying distribution of data over time is considered as an important challenge in processing this type of data. Moreover, concept drift affects the performance of anomaly detection process. The problem of anomaly detection in streaming data is applied to many important applications, for instance, intrusion detection in computer networks or traffic management in the road networks. In recent years, some tensor decomposition based approaches have been presented that track the main pattern or subspace of data in an online manner and adapt the learner with probabilistic changes continuously in all time-intervals by using an implicit strategy. We propose an online approach that detects the concept drift in an explicit manner. Moreover, the learner has been adapted with drift and changes only in their occurrences using informed strategy. Evaluation of the proposed method is performed with real datasets. Analysis of the obtained results confirms the promising performance of the proposed method in terms of learning and detection.

Keywords: anomaly detection, informed and blind adaptation, tensor decomposition, concept drift, robust online subspace tracker.



This is an open access article under the CC BY-NC-ND/4.0/ License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>).



<https://doi.org/10.22108/isee.2021.126591.1443>

مقاله پژوهشی

ارائه یک یادگیرنده برخط رانش آگاه برای تشخیص ناهنجاری در داده‌های جریانی

مریم آموزگار^۱، بهروز مینایی بیدگلی^{۲*}، هادی فنایی^۳، منصور رزقی^۴

۱- دانشجوی دکترا گروه مهندسی نرم افزار، دانشکده مهندسی کامپیوتر- دانشگاه علم و صنعت- تهران-

ایران

Amoozegar_m@comp.iust.ac.ir

۲- دانشیار گروه مهندسی نرم افزار، دانشکده مهندسی کامپیوتر- دانشگاه علم و صنعت- تهران- ایران

b_minai@iust.ac.ir

۳- استادیار مرکز تحقیقات سیستم‌های هوشمند کاربردی - دانشگاه هالمستاد- هالمستاد- سوئد

hadi.fanaee@hh.se.ir

۴- دانشیار گروه علوم کامپیوتر، دانشکده علوم ریاضی - دانشگاه تربیت مدرس - تهران- ایران

rezghi@modares.ac.ir

چکیده: داده‌های جریانی در بستر پویا و در حال تغییر، تکامل می‌یابند؛ بنابراین، رانش مفهوم یا تغییر توزیع اساسی داده‌ها با گذشت زمان، یکی از مهم‌ترین چالش‌های این نوع از داده‌ها است. علاوه بر این، رانش مفهوم بر عملکرد فرآیند تشخیص ناهنجاری نیز تأثیر می‌گذارد. تشخیص ناهنجاری در چنین داده‌هایی در حوزه‌های متعددی از جمله تشخیص نفوذ در شبکه‌های کامپیوتری یا مدیریت ترافیک شبکه راه‌ها کاربرد دارد. در سال‌های اخیر، رویکردهایی مبتنی بر تجزیه تانسور ارائه شده‌اند که به صورت برخط زیرفضا را ردیابی می‌کنند و یادگیرنده را با یک استراتژی ناآگاهانه و به طور ضمنی در همه گام‌های زمانی، در مقابل تغییرات تطبیق می‌دهند. این مقاله، یک رویکرد برخط را پیشنهاد می‌کند که رانش مفهوم را به طور صریح تشخیص می‌دهد و اعلام می‌کند. بدین ترتیب یادگیرنده نیز با یک استراتژی آگاهانه و تنها در گام‌های زمانی لازم با تغییرات و رانش، تطبیق پیدا می‌کند. ارزیابی راهکار پیشنهادی با استفاده از مجموعه داده‌های واقعی انجام شد و تحلیل نتایج به دست آمده، عملکرد روش پیشنهادی را از جنبه‌های یادگیری و تشخیص تأیید می‌کند.

واژه‌های کلیدی: تشخیص ناهنجاری، تطبیق‌پذیری آگاهانه و کورکورانه، تجزیه تانسور، رانش مفهوم

۱- مقدمه

بر پایه یادگیری رفتار طبیعی و استفاده از مدل ساخته شده برای استخراج داده‌های ناهنجار بنا شده‌اند [6]-[4]. در این میان در داده‌های جریانی به دلیل ماهیت پویایشان مسئله کمی پیچیده‌تر خواهد بود؛ زیرا مدل ساخته شده باید به اندازه کافی در مقابل پویایی و تغییرات داده، تطبیق‌پذیر باشد. بدین ترتیب مدل‌های یادگیری برون خط منسوخ و به مدل‌های یادگیری برخط تطبیق‌پذیر نیاز است.

مفهوم تحقیقات بسیار گسترده‌ای در حوزه رانش انجام شده است و در مقالاتی مثل [8]، [7] روش‌های باناظر یا در [9] روش‌های بی‌ناظر مرور شده‌اند. در متون فارسی نیز در [1] و [2] روش‌های باناظر برای تشخیص رانش مفهوم در

یکی از مهم‌ترین مسائل در حوزه داده‌کاوی، تشخیص ناهنجاری در داده‌ها است. [3]-[1] ناهنجاری، هرگونه انحراف از رفتار طبیعی شناخته می‌شود. رویکردهای سستی

^۱ تاریخ ارسال مقاله: ۱۳۹۹/۱۰/۰۷

تاریخ پذیرش مقاله: ۱۴۰۰/۰۴/۲۱

نام نویسنده مسئول: بهروز مینایی بیدگلی

نشانی نویسنده مسئول: ایران - تهران - دانشگاه علم و صنعت ایران - دانشکده مهندسی کامپیوتر

فرآیندهای کسب‌وکار ارائه شده‌اند؛ اولی با تکیه بر شبکه‌های عمیق، روشی خودکار و مستقل از پنجره برای تعدیل دو چالش انتخاب ویژگی و اندازه پنجره لغزان ارائه کرده است. پژوهش دوم نیز یک روش ابتکاری با معرفی یک تابع فاصله اصلاح‌شده، برای شناسایی محل و زمان ایجاد رانش مفهوم ارائه کرده است؛ اما تحقیقات محدودتری در مواجهه رانش مفهوم و ناهنجاری انجام شده‌اند که عمدتاً روش‌های بانظر و شبکه‌های عمیق‌اند [10], [11]. در متون فارسی نیز در [3] راهکاری برای به‌روزرسانی پروفایل مشتریان با لحاظ کردن رانش در رفتار آنها در سیستم‌های کشف تقلب برای کاهش هشدارهای نادرست ارائه شده است. در این پژوهش بر استفاده از الگوریتم‌های فرابتکاری در انتخاب طول مناسب پنجره لغزان تمرکز شده است.

در این میان، رویکردهای برخط مبتنی بر تجزیه، از جمله رویکردهای برجسته در تشخیص ناهنجاری در شبکه‌های در حال تکامل‌اند [15]–[12]. این روش‌ها ویژگی‌های جریان داده را به شکلی انعطاف‌پذیر و کارآمد مدل می‌کنند تا الگوهای پایه و مخفی را در قالب زیرفضای اصلی^۱ یا پنهان، تخمین بزنند و ردیابی کنند. در ادامه، خطای تخمین، مبنای تشخیص ناهنجاری قرار می‌گیرد. نسخه‌های توسعه یافته این روش‌ها که در مقابل نویز استوارند^۲، از ابتدا و در مدل ساخته شده علاوه بر زیرفضای اصلی، مؤلفه‌ای با عنوان مؤلفه خلوت^۳ در نظر می‌گیرند. این مؤلفه در حین ردیابی و تخمین زیرفضای اصلی، انحرافات در داده‌ها را جمع‌آوری می‌کند که از آن برای تشخیص ناهنجاری استفاده می‌شود. همچنین، همان‌گونه که اشاره شد داده‌های جریانی در بستر زمان با تغییرات زیادی مواجه‌اند. تغییر در توزیع داده در طول زمان به پدیده‌ای به نام رانش مفهوم منجر می‌شود. گرچه رویکردهای مبتنی بر تجزیه برای تطبیق‌پذیری با تغییرات ارائه‌اند، پدیده رانش مفهوم به‌تازگی و در قالب رانش در زیر فضای اصلی، معرفی شده است [16].

به‌تازگی مدل‌های برخط مبتنی بر تجزیه ماتریسی که قابلیت مدیریت رانش مفهوم را داشته باشند، به تعداد اندکی ارائه شده‌اند. در [17] روشی با عنوان تجزیه ماتریس

موقتی^۴ برای ردیابی رانش مفهوم در آرایه پنهان هر کاربر در سیستم‌های توصیه‌گر ارائه شده است. این راهکار از یک نسخه اصلاح‌شده کاهش گرادیان تصادفی استفاده می‌کند تا آرایه فضای پنهان کاربر را در هر گام زمانی آموزش دهد. سپس با استفاده از رگرسیون لاسو یک مدل خطی را برای انتقال هر آرایه پنهان آموزش می‌دهد. در واقع این مدل خطی، رانش مفهوم را آموزش و اثر آن را به گام‌های زمانی بعدی انتقال می‌دهد؛ بنابراین، در این راهکار، ردیابی و انتقال اثر رانش مفهوم یا به عبارتی تطبیق ضمنی یادگیرنده با رانش مفهوم مدنظر است و هیچ استراتژی صریحی برای تشخیص و اعلام رانش مفهوم ندارد. علاوه بر این، به بحث تشخیص ناهنجاری نیز پرداخته است.

در [12] نیز روشی مبتنی بر تجزیه ماتریسی برخط با عنوان تجزیه ماتریس فصلی^۵ SMF ارائه می‌شود که با ردیابی تغییرات در فضای پنهان، وجود رانش را در داده‌ها تحلیل می‌کند. مدل ارائه شده در این پژوهش نیز رانش را به صورت ضمنی و تأثیرگذار بر فرآیند یادگیری، مدیریت می‌کند و آگاهی از آن، تنها با تحلیل برون‌خط مؤلفه‌های مدل، امکان‌پذیر است و در واقع استراتژی مشخصی برای تشخیص و اعلام وقوع رانش ندارد.

بحث رانش مفهوم در حوزه روش‌های تجزیه تانسوری برای نخستین بار در [16] مطرح و الگوریتمی برای تشخیص مفاهیم رانش یافته ارائه شد. عملکرد روش ارائه شده بدین شکل است که با رسیدن هر تانسور جدید، ابتدا رتبه^۶ آن، مشخص و به فاکتورهایش تجزیه می‌شود. سپس با مقایسه شباهت مفاهیم جدید با مفاهیم قبلی، مفاهیم جدید، هم‌پوشان یا حذف شده تعیین می‌شوند؛ بنابراین، در این راهکار، رانش مفهوم تنها در قالب ایجاد یا حذف مفهوم پوشش داده شده است و برای تشخیص ناهنجاری نیز راهکاری ندارد.

یکی از مشکلات اساسی در این حوزه، نبود راهکاری با امکان تشخیص هم‌زمان و خودکار رانش و ناهنجاری است. همچنین، در روش‌های فوق، استراتژی صریحی برای تشخیص رانش مفهوم و تطبیق مدل با تغییرات وجود ندارد و مدل به‌طور ناآگاهانه و کورکورانه^۷ و در همه گام‌های زمانی به‌روز می‌شود. توجه به این نکته نیز حائز اهمیت

ورودی و کنترل مکانیسم فراموشی را به عهده دارد. داده ورودی می تواند به صورت تک نمونه یا تجمعی (مبتنی بر پنجره) مدیریت شود. مکانیسم فراموشی نیز به صورت تدریجی یا ناگهانی برای کنترل اثر داده های قدیمی بر مدل اعمال شود. دومین بخش، واحد مدیریت تغییرات، مسئولیت پایش تغییرات و اعلام وقوع رانش را برعهده دارد.

قلب سیستم نیز واحد یادگیری است که سه ویژگی عمده دارد؛ ویژگی نخست، حالت یادگیری را مشخص می کند و مشخص می کند با رسیدن داده جدید، یادگیری از نو آغاز و مدل جدید ساخته شود یا روش های یادگیری افزایشی، برخط یا جریانی استفاده شوند تا به طور تدریجی، فرآیند یادگیری تکمیل شود. ویژگی دوم، تطبیق پذیری را مشخص می کند. واحد یادگیری در این بخش، استراتژی خود را برای تطبیق با تغییرات تعیین می کند. اصولاً دو استراتژی کورکورانه و آگاهانه تعریف شده است. در حالت کورکورانه، یادگیرنده به طور ضمنی و در یک فرآیند پیوسته، تطبیق با تغییرات احتمالی را انجام می دهد؛ بدون اینکه هیچ آگاهی از وقوع رانش داشته باشد. این استراتژی ناآگاهانه و پرهزینه است. در طرف مقابل، سیستم های با استراتژی آگاهانه یا رانش - آگاه، با واحد تشخیص رانش در ارتباط است و فرآیند تطبیق و به روزرسانی یادگیرنده، تنها در صورت به روز رانش انجام می شود. توجه به این نکته نیز حائز اهمیت است که در ادبیات این حوزه، گاهی روش کورکورانه را یادگیری تکاملی و روش آگاهانه را یادگیری خود - تطبیقی می نامند [19]. ویژگی سوم واحد یادگیرنده نیز مشخص می کند یادگیرنده از یک مدل بهره می برد یا مبتنی بر چندین مدل و یادگیری تجمعی بنا شده است.

۲-۲- تجزیه تانسور

تانسور یک آرایه چندبعدی است که مرتبه^۱ آن بیان کننده تعداد ابعاد آن است. تانسور مرتبه یک معادل یک بردار، تانسور مرتبه دو معادل ماتریس و تانسور برای اشاره به تانسورهای با مرتبه سه و بالاتر استفاده می شود. بدین ترتیب، تانسور یک ابزار قدرتمند برای پردازش داده ها از جنبه های مختلف است؛ برای نمونه، تعامل افراد در یک شبکه اجتماعی در بستر زمان در قالب یک تانسور با سه بعد

است که درست است که ناهنجاری، انحراف از رفتار طبیعی است، اما رفتار طبیعی در بستر زمان تغییر می کند؛ بنابراین، این پدیده یعنی همان رانش مفهوم، بر معنای ناهنجاری نیز تأثیر می گذارد. بدین ترتیب تشخیص رانش مفهوم و تفکیک آن از ناهنجاری به کاهش تشخیص هشدار نادرست در فرآیند تشخیص ناهنجاری منجر می شود. همچنین، تشخیص رانش مفهوم و اعلام صریح آن با یک آشکارساز به طراحی یک یادگیرنده رانش - آگاه^۲ منجر می شود. بدین ترتیب یادگیرنده به طور آگاهانه^۳ و تنها در نقاط رانش، به روز و در فرآیند پرهزینه یادگیری صرفه جویی می شود.

با توجه به توضیحات ارائه شده، این مقاله مکانیزمی برای تشخیص صریح رانش مفهوم و به تبع آن، تطبیق آگاهانه یادگیرنده ارائه می کند. همچنین، ناهنجاری را به طور هم زمان و برخط و در هر گام زمانی در کنار وقوع رانش، بررسی و اطلاع رسانی می کند.

در ادامه، مفاهیم پایه در بخش دوم مرور می شوند. روش پیشنهادی در بخش سوم و آزمایشات انجام شده به همراه نتایج به دست آمده در بخش بعدی ارائه می شوند. نتایج و معرفی کارهای آینده در بخش انتهایی ارائه می شوند.

۲- مفاهیم پایه

۲-۱- یادگیری تحت رانش مفهوم

اگر فرآیندی که به تولید داده جریانی منجر می شود با متغیر تصادفی χ مدل شود، رانش مفهوم به صورت «تغییر در احتمال توزیع متغیر تصادفی تحت زمان» [16] تعریف می شود. در یادگیری باناظر، متغیر تصادفی χ به شکل XY (X مجموعه صفات و Y متغیر هدف) و رانش مفهوم به معنی تغییر در توزیع تجمعی X و Y است؛ اما در یادگیری بی ناظر متغیر تصادفی تنها شامل X است و متغیر هدف یا برجستگی وجود ندارد [18]. در روش های مبتنی بر تجزیه تانسور برخط، مفهوم همان ساختار مخفی موجود در داده است [16].

در ادامه، واحدهای سازنده یک سیستم یادگیری بی ناظر تحت رانش مفهوم برگرفته از [8] تشریح می شوند. نخستین بخش، واحد حافظه است که دو وظیفه مدیریت داده

کاربر، کاربر و زمان مدل می‌شود.

روش‌های تجزیه تانسور، ابزار قدرتمندی برای پردازش اینگونه داده‌ها هستند که زیرفضای اصلی و به عبارتی الگوهای معنادار را استخراج می‌کنند. پرکاربردترین تجزیه تانسوری تجزیه CANDECOMP/PARAFAC (CP) است که طبق رابطه زیر تانسور $\chi (L \times W \times T)$ را به ماتریس فاکتورهایش یعنی $B \in \mathbb{R}^{W \times R}$ ، $A \in \mathbb{R}^{L \times R}$ و $C \in \mathbb{R}^{T \times R}$ تجزیه می‌کند [20].

$$\chi \approx \sum_{r=1}^R a_r \circ b_r \circ c_r = [A, B, C] \quad (1)$$

گذر زمان است؛ بنابراین، یک مدل کارآمد باید در مقابل این تغییرات، تطبیق پذیر باشد. نوعی از این تغییرات مربوط به تغییرات ناگهانی یا افزایشی الگوهای اصلی پنهان در داده‌ها یا همان زیرفضای اصلی است. به تازگی در [21] الگوریتمی بر پایه مدل زیر ارائه شده است که تطبیق پذیری بسیار خوبی در مقابل این دسته از تغییرات دارد.

$$\min_{A, B, C, V_\tau, A'_\tau, C'_\tau} \frac{1}{2} \sum_{\tau=1}^t \lambda^{t-\tau} \left[\left\| \begin{matrix} Y_\tau - (A + A'_\tau) \text{diag}(b[\tau]) \\ (C + C'_\tau)^T - V_\tau \end{matrix} \right\|_F^2 \right] + \mu [\tau] (\|A\|_F^2 + \|C\|_F^2) + \lambda_2 (\|A'_\tau\|_1 + \|C'_\tau\|_1) + \frac{\lambda_1}{2} \|V_\tau\|_1 \quad (2)$$

$Y_\tau, V_\tau \in \mathbb{R}^{W \times W}, A, A', C', C \in \mathbb{R}^{W \times R}$

۲-۱- ردیاب برخط و استوار زیرفضا^{۱۱}

در دسته مهمی از روش‌های تشخیص ناهنجاری، داده‌های بعد بالا را به زیرفضای با رتبه پایین^{۱۲} نگاشت می‌کنند؛ به گونه‌ای که در فضای جدید، ناهنجاری‌ها به راحتی تشخیص داده می‌شوند. این مسئله با عنوان یادگیری زیرفضا (یا تخمین رتبه پایین) یا تخمین آنها شناخته می‌شود. یادگیری زیرفضا در محیط نویزی را یادگیری استوار زیرفضا می‌نامند که تخمین به دو مؤلفه زیرفضا و مؤلفه خلوت انجام می‌شود. این مفاهیم در داده جریانی که داده‌ها به صورت یکجا در اختیار نیستند، مفهوم ردیابی برخط زیرفضا را پیدا می‌کنند؛ زیرا فرآیند یادگیری با به روزرسانی و ردیابی زیرفضا در زمان انجام می‌شود. در اینجا نیز در حضور نویز، هم زیرفضا و هم مؤلفه خلوت ردیابی می‌شوند و مسئله ردیابی برخط و استوار زیرفضا شکل می‌گیرد.

۳- روش پیشنهادی

۳-۱- تبیین مسئله

روش‌های مبتنی بر ردیابی برخط و استوار زیرفضا، داده را به دو بخش زیرفضای اصلی و مؤلفه خلوت، تخمین می‌زنند و با تحلیل بخش خلوت، به استخراج ناهنجاری در داده‌ها می‌پردازند. این روش‌ها برای مدل‌سازی داده‌های جریانی بسیار انعطاف پذیرند؛ زیرا ویژگی‌های مختلف مربوط به آنها به راحتی در مدل منعکس می‌شوند. یکی از مهم‌ترین ویژگی داده‌های جریانی، تغییر و پویایی آنها در

بر مبنای مدل فوق، Y_t به عنوان ماتریس مجاورت گراف ورودی در هر گام زمانی به زیرفضای اصلی و تغییراتش به شکل $(A + A'_\tau) \text{diag}(b[\tau]) (C + C'_\tau)^T - V_\tau$ تجزیه می‌شود. یادآوری می‌شود این تجزیه یک فرم توسعه یافته از تجزیه برخط CP است که ماتریس‌های $(A + A'_\tau)$ و $(C + C'_\tau)$ ماتریس‌های فاکتور غیر زمانی و b فاکتور زمان و R رتبه ماتریس داده ورودی است. با اعمال نرم فرینوس و استفاده از مکانیسم کمترین مربعات نمایی وزندار مبتنی بر فاکتور فراموشی λ بر ماتریس‌های A و C ، زیرفضای اصلی با رجوع به سابقه داده در طول زمان ردیابی می‌شود. همچنین، با اعمال نرم یک بر A'_τ و C'_τ به عنوان ماتریس‌های خلوت مربوطه به فاکتورهای A و C انحرافات الگوهای اصلی محاسبه می‌شوند. به این ترتیب، زیرفضای اصلی در یک روند تکاملی و با نگاه به سابقه داده‌ها به روز می‌شود؛ اما انحرافات، در لحظه و برخط به ازای هر گام زمانی محاسبه می‌شوند. ماتریس خلوت V هم در هر گام زمانی انحرافات سراسری را جمع آوری می‌کند که مبنای استخراج ناهنجاری است.

در این رابطه، پارامتر μ و λ_2 برای کنترل دقت تقریب ماتریس‌های مربوط به زیرفضا و پارامتر λ_1 میزان خلوت بودن ماتریس V را کنترل می‌کنند. در مواقعی که نرخ خطا در سیستم کم تخمین زده می‌شود، μ مقدار کمتری خواهد داشت. λ نیز مقداری بین صفر و یک دارد. مقدار یک به معنی استفاده از همه سوابق داده‌ها و معادل مدل برون خط است؛ اما در مقادیر کمتر از یک به صورت نمایی از اهمیت و وزن داده‌های قدیمی برای محاسبه فاکتورهای

جدید کاسته می شود.

مفهوم طراحی شده است. با توجه به توضیحات مدل رابطه (۲) در محاسبه زیرفضای اصلی و استفاده از مکانیسم فراموشی تدریجی، رانش مفهوم ناگهانی و افزایشی به راحتی با این روش تشخیص داده می شود. رانش متناوب تنها در صورتی شناسایی می شود که در داده اخیر در محدوده تطبیقات فاکتور فراموشی به وفور الگوی متناوبی اتفاق افتاده باشد و امکان آموختن آن فراهم باشد که در این مقاله به آن پرداخته نمی شود.

با توجه به اینکه هر ستون از ماتریس های A' و C' به یک مؤلفه در داده اشاره دارد، اعمال نرم $L_{2,1}$ به صورت $\|X\|_{2,1} = \sum_{j=1}^n \left(\sum_{i=1}^m |x_{ij}|^2 \right)^{1/2}, X \in \mathbb{R}^{n \times m}$ روی این ماتریس ها سنجه مناسبی برای اندازه گیری میزان تغییرات است؛ بنابراین، در هر گام زمانی، این نرم، محاسبه و سری زمانی حاصل از آن به منظور تشخیص قله^۴، پایش می شود. سری های زمانی به دست آمده به صورت برخط با الگوریتم smoothed z-score [22] تحلیل می شوند. این الگوریتم با ورود داده جدید X_t در سری زمانی، میزان انحراف آن نقطه را نسبت به میانگین متحرک به طول m محاسبه می کند. در صورتی که z-score به دست آمده از مقدار حد آستانه بیشتر باشد، یک قله گزارش می شود. میانگین متحرک (s) و انحراف از معیار (σ) طبق روابط زیر محاسبه می شوند:

$$z_t = \frac{x_t - \bar{s}_{t-1}}{\sigma_{s_{t-1}}} \quad (3)$$

$$\bar{s}_t = \frac{1}{m} \sum_{t-m}^{t+m} s_t, \sigma_{s_t} = \sqrt{\frac{\sum_{t-m}^{t+m} (s_t - \bar{s}_t)^2}{m-1}} \quad (4)$$

با توجه به اینکه یک قله ممکن است از تشخیص قله های بعدی جلوگیری کند، قبل از به روزرسانی میانگین متحرک و انحراف معیار، لازم است داده جدید با رابطه زیر فیلتر شود:

$$s_t = I_{nf} x_t + (1 - I_{nf}) s_{t-1} \quad (5)$$

بدین ترتیب، داده جدید به وسیله مقدار خوانده شده فعلی و مقدار فیلتر شده قبلی تعدیل می شود. در این رابطه میزان مصالحه بین دو نقطه داده جاری و قبلی را

مدل فوق یک مسئله بهینه سازی نامحدب است که تخمین فاکتور زمان با استفاده از تخمین گر کمترین مربعات، تخمین ماتریس خلوت V بر اساس عملگر آستانه گیری نرم، تخمین A' و C' در قالب تخمینگر لاسو و تخمین فاکتورهای A و C با استفاده از کمترین مربعات بازگشتی به صورت متناوب انجام می شوند. محاسبات مربوطه به طور کامل در مقاله [21] آمده اند.

با ارجاع به شکل (۱) به عنوان اجزای پایه سیستم مبتنی بر رانش و انتخاب راهکار و مدل مرجع [21] برای حل مسئله این پژوهش، چند نکته مهم درخور ذکر است: نخست اینکه این راهکار بین رانش و ناهنجاری تمایزی قائل نیست و رانش، مفهوم را به عنوان ناهنجاری شناسایی می کند؛ زیرا واحد تشخیص رانش مفهوم ندارد و خطای باقیمانده کل، مبنای تشخیص انحراف است که تمایزی بین تغییرات ناشی از رانش مفهوم و ناهنجاری قائل نمی شود. نکته دوم به دلیل ناآگاهی از بروز رانش، در همه گام های زمانی، تطبیق با تغییرات یا همان یادگیری ماتریس های زیرفضا تکرار می شود. این در حالیست که محاسبه این ماتریس ها بسیار پرهزینه است. بر اساس این، برای رفع موانع فوق، چارچوب پیشنهادی این مقاله با عنوان ردیاب برخط استوار تطبیق پذیر آگاهانه^۴، IAROST، شامل سه واحد یادگیری، تشخیص رانش مفهوم و تشخیص ناهنجاری تشریح می شود.

۳-۲- تشخیص رانش مفهوم

یکی از الزامات یادگیرنده برخط رانش - آگاه در یک محیط تغییرپذیر، وجود مکانیزم تشخیص رانش مفهوم به صورت صریح و مجزا از ناهنجاری است [8]. این مکانیزم اطلاعات مفیدی را در ارتباط با پویایی داده ها فراهم می کند. در این بخش، مکانیزم پیشنهادی تشریح می شود.

با توجه به اینکه دو ماتریس A' و C' در رابطه (۲) انحرافات در زیرفضای اصلی را نشان می دهند، بهترین نشانه برای رانش مفهوم در داده ها هستند. همچنین اندازه بزرگی این ماتریس ها شدت رانش را نشان می دهد. بدین ترتیب، مکانیزم زیر برای پایش این ماتریس ها و اعلام وقوع رانش

یادگیرنده واحد انجام می‌شود یا ترکیبی از چندین یادگیرنده. در این پژوهش، یک یادگیرنده واحد در نظر گرفته شده است که با گذشت زمان تکامل می‌یابد.

۳-۴- تشخیص ناهنجاری

با توجه به اینکه ماتریس V ، بخش خلوت را نشان می‌دهد، اندازه بزرگی آن، سنجه خوبی برای تعیین زمان بروز ناهنجاری است. با استفاده از الگوریتم smoothed z-score، استفاده شده در شکل (۲)، سری زمانی حاصله از نرم دو ماتریس V (سنجه اندازه بزرگی ماتریس به صورت $\|X\|_2 = \left(\sum_{i=1}^m \sum_{j=1}^n |x_{ij}|^2\right)^{1/2}$ ، $X \in \mathbb{R}^{n \times m}$) پایش و با شناسایی یک قله، بروز ناهنجاری، اعلام و محرک لازم به واحد تشخیص ناهنجاری صادر می‌شود تا یال و گره ناهنجار تعیین شوند؛ بدین شکل که ضمن اعمال تابع z-score روی ماتریس V ، درایه‌های با بیش از مقدار آستانه δ ، یال ناهنجار اعلام می‌شوند. همچنین، به منظور تشخیص گره‌های ناهنجار، نرم دو هر ردیف از ماتریس، محاسبه و با اعمال تابع z-score موارد بیش از حد آستانه δ اعلام می‌شوند. با این روال، به صورت برخط، در هر گام زمانی با کمترین هزینه، موارد ناهنجار شناسایی و اعلام می‌شوند. شکل (۳) شبه کد این بخش را نشان می‌دهد.

۳-۵- الگوریتم IAROST

الگوریتم روش پیشنهادی به‌طور کامل در شکل (۴) نمایش داده شده است. با توجه به شکل (۴)، با دریافت داده جدید در گام زمانی t ، متغیرهای $b[t]$ ، $V[t]$ ، $A'[t]$ و $C'[t]$ طی فرآیند یادگیری محاسبه می‌شوند. در ادامه، به منظور تشخیص رانش مفهوم، $A'[t]$ و $C'[t]$ با الگوریتم تشخیص رانش مفهوم پایش می‌شوند. وقتی یک قله شناسایی می‌شود، ضمن اعلام وقوع رانش، متغیر Trigger سیگنال فعال‌سازی بخش تطبیق‌پذیری در واحد یادگیری را برای به‌روزرسانی و تطبیق ماتریس‌های A و C صادر می‌کند. علاوه بر این، اندازه ماتریس V ، با الگوریتم تشخیص قله پایش می‌شود. هنگامی که یک قله شناسایی

تعیین می‌کند. بر اساس این، با اعلام قله توسط این الگوریتم، وقوع رانش مفهوم گزارش می‌شود. شکل (۲) شبه‌کد این الگوریتم را نشان می‌دهد.

۳-۳- یادگیری و تطبیق‌پذیری

بر اساس ویژگی‌های قلب یک سیستم یادگیری مبتنی بر رانش طبق شکل (۱) و با استناد به رابطه (۲)، حالت یادگیری به صورت برخط است. ویژگی دوم مربوط به تطبیق‌پذیری مدل است که نحوه سازگاری مدل در مواجهه با داده‌های جدید را مشخص می‌کند. همان‌طور که قبلاً گفته شد در رویکردهای یادگیری تحت رانش، دو استراتژی عمده تطبیق‌پذیری معرفی شده است، کورکورانه و آگاهانه. استراتژی کورکورانه در واکنش به رانش به صورت پیشگیرانه^{۱۵} و کند عمل می‌کند. علاوه بر این، به دلیل به‌روزرسانی منظم مدل در هر گام زمانی، بسیار هزینه‌بر و زمان‌بر است؛ در حالی که یک استراتژی آگاهانه، به صورت واکنشی^{۱۶} عمل می‌کند و در هر گام زمانی، وقوع رانش را بررسی و در صورت مثبت بودن، فرآیند تطبیق مدل با داده جدید انجام می‌شود. در این حالت، مدل فقط در صورت نیاز به روز می‌شود؛ بنابراین، این استراتژی کارآمدتر است. یکی دیگر از مزایای یک استراتژی آگاهانه، ارائه اطلاعات درباره پویایی فرآیند تولید داده است. این پژوهش، یک استراتژی آگاهانه برای تطبیق‌پذیری ردیاب برخط و استوار زیرفضا در مواجهه با رانش ارائه می‌کند.

اولین نیاز برای یک سیستم تطبیق‌پذیر آگاهانه، فراهم‌سازی سازوکار پایش و هشدار است تا تنها در زمان وقوع رانش، فرآیند تطبیق‌پذیری را فعال کند. در بخش قبل، پایش اندازه ماتریس‌های A' و C' در واحد تشخیص رانش مفهوم تشریح شد. وقتی وقوع رانش با این واحد اعلام شد، روال به‌روزرسانی ماتریس‌های A و C فراخوانی می‌شود. در این حالت، محاسبه هزینه‌بر و سنگین این ماتریس‌ها که همان زیرفضای اصلی و الگوی اصلی داده‌ها هستند، تنها در زمان وقوع رانش انجام می‌شود. شکل (۳) الگوریتم مربوط به فرآیند یادگیری راهکار پیشنهادی را نشان می‌دهد.

ویژگی سوم یک یادگیرنده در سیستم‌های مبتنی بر رانش این است که فرآیند یادگیری با استفاده از یک

وجود مجموعه داده برچسب‌دار برای ارزیابی است. به‌سختی می‌توان یک مجموعه داده جریانی در حال تکامل از یک شبکه پیدا کرد که با دقت رضایت‌بخشی برچسب‌گذاری شده باشد. یکی از راه‌های متداول، ساخت مجموعه داده انسانی است که علاوه بر هزینه‌بر بودن، به دلیل خطای بالای «هشدار نادرست» چندان کارآمد نخواهد بود. علاوه بر این، به دلیل کم بودن تعداد ناهنجاری در داده‌ها، امکان ارزیابی روش‌ها از جنبه‌های مختلف و درمقابل، انواع مختلف ناهنجاری به‌خوبی فراهم نمی‌شود. یکی از راهکارهای دیگر برای ساخت مجموعه داده آزمایشی، براساس تحقیقات مختلف و استفاده از داده تمیز و تزریق ناهنجاری به آن است.

بدین منظور، مجموعه داده‌های گراف مربوط به سفر تاکسی‌های نیویورک بین ۷۱ ناحیه در ژانویه ۲۰۱۳ انتخاب می‌شوند و در مرحله اول، تابع موجک Daubechies-5 روی سیگنال حاصل از وزن هر یک از یال‌های گراف در طول زمان (تکامل وزن یال در زمان) اعمال می‌شود تا نمودار حاصله هموار شود. در مرحله دوم، برش زمانی از مجموعه داده انتخاب می‌شود که روند تقریباً هموار و یکنواختی دارد. در مرحله سوم، تجزیه CP اعمال می‌شود تا فاکتورهای پایه استخراج شوند. در ادامه، تابع زمان مربوط به هر سیگنال با استفاده از تکنیک‌های برازش منحنی، استخراج و روی فاکتورهای پایه اعمال می‌شود. در انتها نویز گوسی به داده‌ها اعمال می‌شود. به این شکل، ابتدا داده تمیز با پایه داده واقعی تولید می‌شود.

برای تزریق یال ناهنجار به مجموعه فوق، در زمان‌های ۵۰، ۱۲۵ و ۱۵۰ تعدادی از درایه‌های ماتریس، معادل یال‌ها در شبکه، انتخاب می‌شوند و وزن آنها به مقدار دیگری تغییر می‌کند. میزان این تغییر با یک فاکتور ضربی، تنظیم‌شدنی است. با مقادیر مختلف این فاکتور، ناهنجاری‌هایی با میزان انحراف متفاوت به مجموعه داده تزریق می‌شوند. در این تحقیق، فاکتور ضربی بین مقدار ۰,۵ تا ۱,۵ در نوسان بوده است.

برای تزریق گره ناهنجار به شبکه، در گام‌های زمانی ۷۰

شد، Triger، فعال و الگوریتم تشخیص ناهنجاری، فراخوانی می‌شود تا عوامل ناهنجار را شناسایی و اعلام کند. شکل (۵) شبه‌کد کامل این الگوریتم را نشان می‌دهد.

بدین ترتیب، نویز با تحلیل خطای کل حاصل از تخمین، ناهنجاری از طریق تحلیل ماتریس خلوت V در بخش تشخیص ناهنجاری و رانش مفهوم با واحد تشخیص رانش، به‌صورت مجزا از هم تعیین می‌شوند. درنهایت، همان‌گونه که ادعا شد و الگوریتم پیشنهادی نیز تأیید می‌کند، تنها با پس‌پردازش‌های ساده و کم‌هزینه روی متغیرهای مدل در هر گام زمانی، به‌طور خودکار و هم‌زمان ناهنجاری و رانش مفهوم به‌صورت مجزا و برخط تشخیص‌پذیر و تفکیک‌شدنی‌اند.

۴- ارزیابی روش پیشنهادی

در این بخش، ابتدا توانایی یادگیرنده روش پیشنهادی از جنبه همگرایی و تطبیق‌پذیری بررسی می‌شود. سنجه استفاده‌شده برای ارزیابی، سری زمانی ساخته‌شده از محاسبه خطای باقیمانده نرمال‌شده (NRE^V) براساس رابطه زیر است:

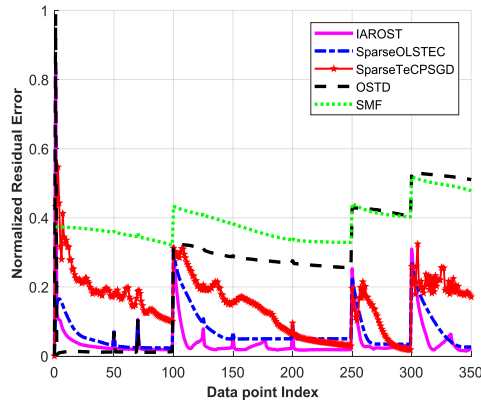
$$NRE_t = \|X_t - Y_t\|_F^2 / \|Y_t\|_F^2 \quad (6)$$

Y_t نمونه داده دریافتی در زمان جاری و X_t تخمین رتبه پایین انجام‌شده با مدل است. هر چقدر مقدار این سنجه بیشتر باشد، نشان‌دهنده این است که مدل در تخمین داده واقعی، خطای بیشتری داشته است. اگر میزان این خطا با نرخ بیشتری بعد از تغییرات کاهش یابد، نشان از سرعت بیشتر در یادگیری و تطبیق‌پذیری مدل با تغییرات دارد. مقدار خطای میانگین در حال اجرا در گام زمانی آخر، نمایش عددی از توان همگرایی مدل است که طبق رابطه زیر محاسبه می‌شود.

$$RAE = \frac{1}{T} \sum_{t=1}^T \|X_t - Y_t\|_F^2 / \|Y_t\|_F^2 \quad (7)$$

یکی از بزرگ‌ترین مشکلات حوزه تشخیص ناهنجاری،

نتایج NRE به‌دست‌آمده از اجرای الگوریتم‌ها در شکل (۶) نشان داده شدند.



شکل (۶): نمودار خطی باقیمانده نرمال شده همه روش‌ها

با توجه به شکل (۶)، کلیه روش‌ها در نقاط مربوط به ناهنجاری (گام‌های زمانی ۵۰، ۷۰، ۱۲۵، ۱۵۰ و ۲۰۰) و رانش مفهوم (نقاط زمانی ۱۰۰، ۲۵۰ و ۳۰۰) دچار خطا شده‌اند؛ به این دلیل که در داده تغییر ایجاد شده است؛ اما نکته درخور توجه، تفاوت بین ناهنجاری و رانش است. خطای به‌وجودآمده در ناهنجاری لحظه‌ای و نقطه‌ای است؛ اما خطای مربوط به نقاط رانش مفهوم، طولانی‌ترند؛ زیرا مدت زمانی طول می‌کشد تا مدل، الگوی جدید را آموزش ببیند؛ اما روش پیشنهادی زودتر از بقیه همگرا شده و زودتر از بقیه تغییر در الگو را آموزش دیده است. در این میان، روش‌های OSTD و SMF کمترین انعطاف را در مقابل رانش مفهوم دارند؛ زیرا نتوانسته‌اند همگرا شوند. همچنین، با استناد به اینکه روش TeCPSGD مبتنی بر کاهش در امتداد گرادیان است، همگرایی مطلوبی در نقاط رانش مفهوم ندارد.

به‌منظور بررسی دقیق‌تر عملکرد روش پیشنهادی، آزمایشات بیشتری برای مجموعه داده‌های جدول (۱) انجام شده‌اند و مقدار خطای میانگین در حال اجرا، طبق رابطه (۷) در جدول (۲) ثبت شد. براساس نتایج، روش پیشنهادی با خطای بسیار کمتری همگرا شده و فرآیند یادگیری را بهتر انجام داده است. با استناد به نتیجه مشاهده‌شده در شکل (۶) و تعمیم آن، خطای کمتر به دلیل همگرایی سریع‌تر روش پیشنهادی در نقاط رانش است.

و ۲۰۰ تعداد محدودی از گرهای شبکه انتخاب می‌شوند و وزن ده درصد از یال‌های متصل به آنها متناسب با فاکتور ضربی تغییر داده می‌شود. به‌منظور تغییر زیرفضا، روال زیر در گام‌های زمانی ۱۰۰، ۲۵۰ و ۳۰۰ در پیش گرفته شد. تنها ۲ درصد از عناصر مربوط به ماتریس فاکتورهای اصلی انتخاب شد و مقادیر آنها براساس فاکتور ضربی تغییر داده شدند. سپس ماتریس‌های فاکتور جدید برای بازسازی و تولید داده جدید استفاده شدند. در انتها با مقایسه داده‌ها، همه یال‌های تأثیر گرفته از این فرآیند تعیین شدند و به‌عنوان یال‌های ناهنجاری برچسب خوردند که به‌صورت گروهی باعث تغییر زیرفضا شده‌اند.

روش‌های مقایسه‌پذیر انتخابی برای این بخش باید از بین راهکارهای مبتنی بر تجزیه تانسور یا ماتریس باشند که با دو استراتژی ضمنی یا صریح، رانش را مدیریت می‌کنند. بیشتر روش‌های ارائه‌شده رانش مفهوم را به‌صورت ضمنی مدیریت می‌کنند. همچنین، تحقیقاتی که به‌طور هم‌زمان به تشخیص ناهنجاری و رانش پرداخته باشند، بسیار اندک‌اند. مرتبط‌ترین راهکار [12] SMF است که مدل برخط مبتنی بر تجزیه ماتریسی ارائه داده و امکان رانش مفهوم و تشخیص ناهنجاری را به‌صورت ضمنی و غیرخودکار فراهم ساخته است. روش‌های [23] OLSTEC، [24] OSTD و [15] TeCPSGD روش‌های دیگری‌اند که بحث رانش یا تغییر زیرفضا را به‌صورت ضمنی انجام می‌دهند و برای تجهیز تشخیص ناهنجاری، نسخه استوار آنها در این مقاله طراحی و پیاده‌سازی شد. برای ایجاد شرایط یکسان برای همه الگوریتم‌ها، مقادیر پارامترهای λ_1 و μ به ۰٫۴ و ۰٫۱ تنظیم شده‌اند. گام آموزش یا stepsize در الگوریتم TeCPSGD به مقدار ۰٫۱ براساس دستورالعمل ذکر شده در [25] تنظیم شد. λ_2 در روش پیشنهادی براساس دستورالعمل [23] به $0.04 = \lambda_1 \times 0.1$ و با توجه به دانش اولیه از میزان همواربودن داده‌ها، پارامتر lag به ۱۰ مقداردهی شدند. مقدار δ به مقدار معمول، یعنی عدد ۳ تنظیم شده است؛ البته مقادیر مختلف در محاسبات مربوط به جدول (۳)، ارزیابی شدند.

ناحیه نشان داده شده در این شکل، مطابق با قله‌های تشخیص داده شده در واحد تشخیص رانش مفهوم و فواصل زمانی حرکات پس‌زمینه است. بدین ترتیب، نتایج به‌دست‌آمده توانایی روش پیشنهادی را برای تشخیص رانش مفهوم به‌صورت صریح تأیید می‌کند.

تشخیص ناهنجاری هم‌زمان با رانش و تفکیک این دو از یکدیگر به‌صورت برخط و خودکار، یکی از مهم‌ترین دستاوردهای IAROST است. بر اساس این، انتظار می‌رود IAROST خطای کمتری در تشخیص هشدار نادرست و در نتیجه، عملکرد بهتری نسبت به روش‌های دیگر داشته باشد. بدین منظور، در ادامه، توانایی آن در تشخیص ناهنجاری با هم‌تایش [21] BAROST ارزیابی می‌شود که مدل مشابه اما رویکرد کورکورانه دارد.

با توجه به نادر بودن نمونه‌های ناهنجار (کلاس مثبت) در مقایسه با نمونه‌های عادی (کلاس منفی)، مجموعه داده بسیار نامتوازن است؛ بنابراین، لازم است معیاری برای ارزیابی انتخاب شود که به کلاس نادر اهمیت بیشتری دهد. دقت^{۱۹} و فراخوانی^{۲۰} به‌عنوان دو سنجه عملکرد مناسب و F-measure به‌عنوان میانگین هارمونیک آنها به‌طور معمول استفاده می‌شوند. در ادامه، مقادیر F-measure مربوط به حد آستانه‌های مختلف، در جدول (۳) نشان داده شده‌اند. مقادیر بیشتر این سنجه در روش IAROST، کارایی بیشتر آن را در تشخیص ناهنجاری تأیید می‌کند.

جدول (۳): سنجه F-measure در مقادیر مختلف حد آستانه در تشخیص ناهنجاری با دو روش IAROST و BAROST

مقدار حد آستانه	IAROST	BAROST
۱	۰,۸۰۱۶	۰,۶۶۴۵
۲	۰,۴۷۸	۰,۶۴۸۴
۳	۰,۶۷۳۱	۰,۵۹۳۵
۴	۰,۶۲۴۶	۰,۵۵۱۳
۵	۰,۶۲۴۱	۰,۵۵۴۷
۶	۰,۶۲۲۳	۰,۵۶۶۳
۷	۰,۶۳۲۲	۰,۵۷۸۷
۸	۰,۶۱۱	۰,۵۷۵۴
۹	۰,۶۰۶۲	۰,۵۷۵
۱۰	۰,۹۴۳۷	۰,۴۶۶۱

در ادامه، برای نشان‌دادن عملکرد تشخیص رانش مفهوم به‌صورت ملموس و شهودی، روش پیشنهادی روی داده ویدئویی به‌عنوان یک داده جریانی دنیای واقعی پیاده‌سازی شد و نتایج تجزیه و تحلیل شدند.

در حوزه پردازش ویدئو، شاخه‌ای با عنوان جداسازی پس‌زمینه^{۱۸} وجود دارد که در یک جریان ویدئو، اشیای متحرک (پیش‌زمینه) را از پس‌زمینه جدا می‌کند. اشیای پیش‌زمینه معمولاً شامل تعداد کمی از پیکسل‌ها در یک فریم فیلم‌اند که به‌عنوان ناهنجاری در نظر گرفته می‌شوند [26].

یک داده ویدئویی در قالب جریانی از فریم‌ها و به‌عنوان یک توالی طولانی از تصویر، در قالب یک تانسور، مدل و به‌صورت برخط و فریم به فریم پردازش می‌شود. به این ترتیب، برخی از رویکردهای برخط مبتنی بر ردیابی زیرفضایی استوار ارائه شده‌اند که پس‌زمینه را به‌عنوان زیرفضای اصلی یا داده رتبه پایین و اجسام متحرک را به‌عنوان یک بخش خلوت، مدل و ناهنجاری اعلام می‌کنند [24], [27].

با تغییر پس‌زمینه در ویدئو، زیرفضای اصلی در طول زمان تغییر می‌کند؛ بنابراین، یک ردیاب برخط و استوار زیرفضا لازم است تا تغییرات پس‌زمینه را مدیریت کند. به‌تازگی روشی برای تعدیل و تطبیق یادگیرنده با تغییرات پس‌زمینه معرفی شده است؛ اما این روش، توانایی تشخیص و اعلام رانش را ندارد.

رانش مفهوم در یک جریان ویدئویی به‌عنوان یک نقطه زمانی است که در آن پس‌زمینه تغییر می‌کند. در ادامه، الگوریتم IAROST روی یک جریان ویدئویی با پس‌زمینه پویا به نام "Airport hall" از مجموعه داده [28] "I2R" اعمال می‌شود تا عملکرد الگوریتم از جنبه یادگیری و تشخیص رانش مفهوم ارزیابی شود. این ویدئو شامل ۵۰۰ فریم با اندازه ۲۸۸ × ۳۵۲ است. رانش مفهوم یا تغییر پس‌زمینه در سه بازه زمانی شامل فریم‌های ۳۸ تا ۱۱۳، فریم‌های ۱۹۰ تا ۲۶۵ و فریم‌های ۳۴۲ تا ۴۱۷ رخ می‌دهد. روش پیشنهادی برای تشخیص رانش مفهوم بر مجموعه داده اعمال شد. شکل (۷) به‌ترتیب سری زمانی دو شاخص رانش مربوط به ماتریس‌های A' و C' را نشان می‌دهد. سه

۵- نتیجه‌گیری

در این مقاله، راهکاری مبتنی بر ردیابی برخط و استوار زیرفضا ارائه شده است که به‌طور هم‌زمان و خودکار رانش مفهوم و ناهنجاری را در داده جریانی تشخیص می‌دهد. در این مقاله، یک آشکارساز تشخیص رانش مفهوم، طراحی شده است که علاوه بر اعلام وقوع رانش، اطلاعات مفیدی درباره پویایی داده‌ها در اختیار قرار می‌دهد. مهم‌ترین دستاورد تفکیک رانش مفهوم و ناهنجاری از یکدیگر، کاهش درخور توجه میزان تشخیص هشدار نادرست روش پیشنهادی در تشخیص ناهنجاری است. نتایج به‌دست‌آمده از آزمایش‌ها روی یک داده ویدئویی، توانمندی روش پیشنهادی را در تشخیص رانش مفهوم تأیید می‌کند.

علاوه بر این، وجود آشکارساز رانش به طراحی مکانیزمی آگاهانه برای تطبیق‌پذیری یادگیرنده با رانش مفهوم منجر شده است؛ بنابراین، تنها در مواقع ضرورت، فرآیند پرهزینه یادگیری الگوهای اصلی انجام می‌شود. این مقاله در کارهای آتی می‌تواند از ابعاد مختلف، از جمله امکان تشخیص ناهنجاری‌های ساخته‌یافته‌تر در داده‌ها یا حل مسئله بهینه‌سازی با روش‌های دیگر به‌منظور افزایش سرعت یادگیرنده توسعه یابد.

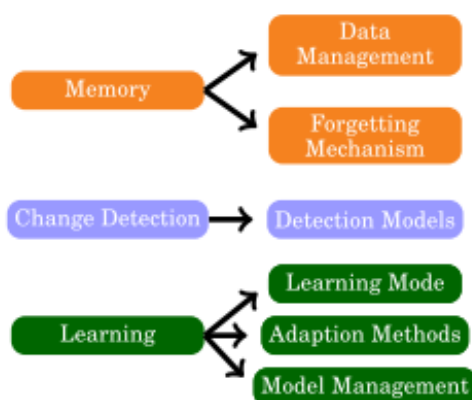
علاوه بر این، سنجه کشف هشدار نادرست یا FPD براساس رابطه زیر محاسبه می‌شود تا به‌طور صریح‌تر و واضح‌تر امکان ارزیابی و مقایسه دو روش را فراهم کند.

$$FPD = FP/TP + FP \quad (۸)$$

مقدار کمتر برای FPD نشان‌دهنده مقدار کمتر هشدار نادرست و کارایی بهتر است. مقادیر این سنجه برای دو روش IAROST و BAROST در جدول (۴) به‌ازای سطوح مختلف حدآستانه نمایش داده شده‌اند. همان‌طور که نتایج نشان می‌دهد IAROST عملکرد بهتری داشته است؛ بنابراین، تشخیص هم‌زمان رانش و ناهنجاری و تفکیک این دو از یکدیگر، به‌طور مستقیم کارایی سیستم را در تشخیص ناهنجاری افزایش می‌دهد.

جدول (۴): مقدار FPD در حد آستانه های متفاوت

مقدار حد آستانه	IAROST	BAROST
۱	۰,۳۲۲۱	۰,۴۹۷۵
۱,۵	۰,۳۱۹۸	۰,۴۷۷۱
۲	۰,۳۲۲۸	۰,۴۷۳۸
۲,۵	۰,۳۳۹۱	۰,۴۸۴۲
۳	۰,۳۳۹۳	۰,۴۷۵۱



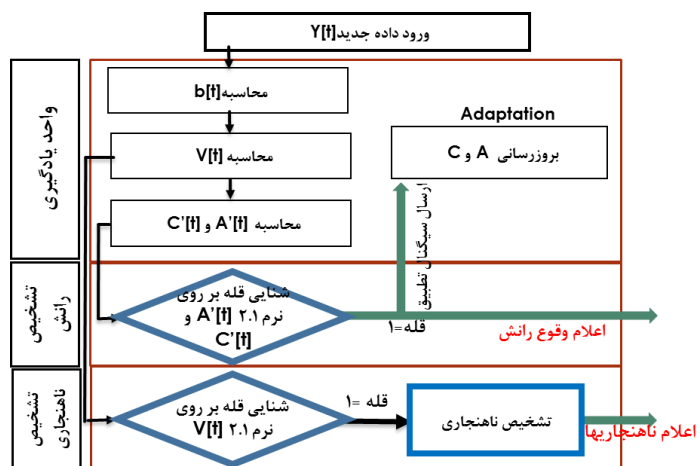
شکل (۱): واحدهای اصلی یک سیستم یادگیری تحت رانش مفهوم [8]

Peak Detection Procedure
<p>Input <i>signal</i></p> <p><i>Trigger</i>=0</p> <p>If $t > \text{lag}$</p> <p>Calculating z_t for <i>signal</i> according to equation (3)</p> <p>If obtained $z_t > \text{threshold}$</p> <p style="padding-left: 20px;"><i>Trigger</i>=1</p> <p>End if</p> <p>Calculate s_t according to equation (5)</p> <p>Calculating new mean and variance according to equation (4)</p> <p>End if</p> <p>Output: <i>Trigger</i></p>

شکل (۲): الگوریتم شناسایی قله

Anomaly Detection Procedure
<p>Input V_t, δ</p> <p>Apply Zscore on vectorized V_t</p> <p>For each element of V_t</p> <p style="padding-left: 20px;">if $Zscore(V[t]_{l,w}) > \delta$</p> <p style="padding-left: 40px;">the connected edge between l and w is returned as anomalous</p> <p style="padding-left: 20px;">end if</p> <p>end for</p> <p>Output: list of anomalous edges</p>

شکل (۳): روال تشخیص ناهنجاری



شکل (۴): نمودار مربوط به الگوریتم روش پیشنهادی

Informed Adaptive Robust Online Subspace Tracking
<p>Input: $\{Y_t\}_{t=1}^{\infty}, \mu, \lambda_1, \lambda_2$</p> <p>Initialize: $\{A[0], b[0], C[0], A'[0], C'[0]\}, V[0]=0, (RA_v[0])^{-1} = (RC_w[0])^{-1} = \gamma I_R, \gamma > 0$</p> <p>For $t=1, 2, \dots$ Do</p> <p>Calculating $b[t], V[t], A[t]', C[t]'$ according to [21]</p>
Drift Detection and adaptation
<p>SignalA[t]= $\ell_{2,1}$-norm of A'</p> <p>SignalC[t]= $\ell_{2,1}$-norm of C'</p> <p>Apply peak detection procedure on SignalA (Monitoring A') and return triggerA</p> <p>Apply peak detection procedure on SignalC (Monitoring C') and return triggerC</p> <p>If triggerA or triggerC=1</p> <p style="padding-left: 20px;">Current time point is returned as drift</p> <p style="padding-left: 20px;">Adaptation is performed by updating the main subspace A and C</p> <p style="padding-left: 40px;">%(calculating A and C according to appendix)</p> <p>End if</p>
Anomaly Detection
<p>SignalV[t]= ℓ_2-norm of V</p> <p>Apply peak detection procedure on SignalV (Monitoring V)</p> <p>If trigger=1</p> <p style="padding-left: 20px;">Call anomaly detection procedure</p> <p>End if</p> <p>End for</p>

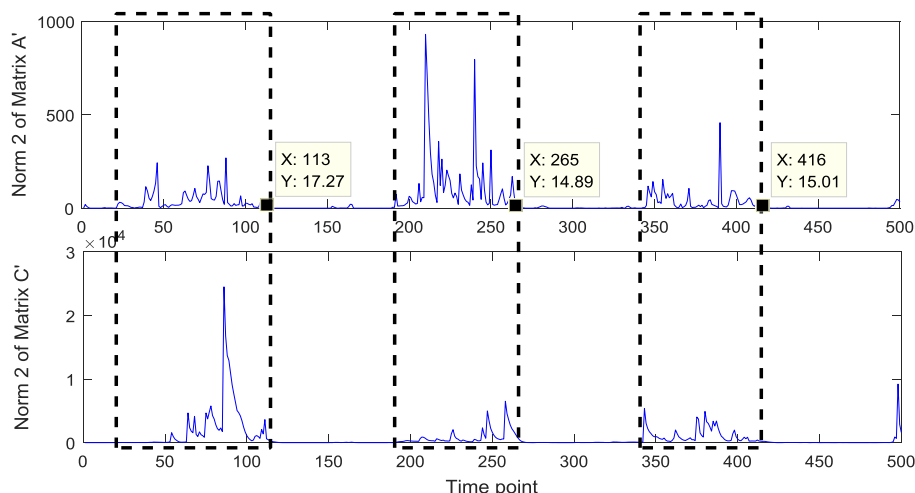
شکل (۵): شبه‌الگوریتم روش پیشنهادی

جدول (۱): فهرست مجموعه داده‌های استفاده شده

عنوان	تعداد گره‌ها	شرح
NYC Taxi [29]	۷۱	سفرهای بین ۷۱ ناحیه شهری نیویورک توسط تاکسی‌های زرد در ژانویه ۲۰۱۳
Boston Bike-Sharing [6]	۹۵	سفرهای مربوط به دوچرخه‌های اشتراکی در شهر بوستون در بازه زمانی ۲۸ جولای ۲۰۱۱ تا آخر سپتامبر ۲۰۱۲
Enron [30]	۱۴۸	تعاملات کاربران با پست الکترونیکی در بازه زمانی بین سالهای ۱۹۹۸ تا ۲۰۰۲
Trade [31]	۲۰۰	حجم تبادلات تجاری سالانه بین کشورها در بازه زمانی ۱۸۷۰ تا ۲۰۰۹ میلادی
Immigration [32]	۲۲۲	مهاجرت بین کشورها در یک دوره پنج ساله
Car [33]	۷۷۱	سفرهای مربوط به روزهای کاری و وسائل نقلیه ظرفیت بالا (HOV) در پاسادنا و حومه

جدول (۲): خطای میانگین در حال اجرای مربوط به روش‌های مختلف روی کلیه مجموعه داده‌ها

عنوان	NYC Taxi [29]	Boston Bike-Sharing	Enron	Trade	Immigration	Car
IAROST	۰.۰۳۵	۰.۱	۰.۰۲۰	۰.۰۲۵	۰.۰۲۵	۰.۰۲۶
SparseOLSTEC	۰.۰۶۷	۰.۱۴	۰.۰۶۰	۰.۰۴۶	۰.۰۴۱	۰.۰۵۱
TeCPSGD	۰.۱۵۴	۰.۱۸	۰.۱۶۸	۰.۲۱۶	۰.۱۳۸	۰.۱۲۶
OSTD	۰.۲۶۲	۰.۲۱۳	۰.۰۶۱	۰.۰۵۰	۰.۰۷۴	۰.۰۰۶
SMF	۰.۳۹۱	۰.۳۱	۰.۰۷۵	۰.۰۶۷	۰.۱۴۳	۰.۰۶۳



شکل (۷): تشخیص رانش مفهوم در قالب جابه‌جایی پس‌زمینه در مجموعه داده "Airport hall"

Management of Data - CoDS-COMAD '18, 2018, pp. 78–87.

- [11] D. Zambon, C. Alippi, and L. Livi, "Concept Drift and Anomaly Detection in Graph Streams," *IEEE Trans. Neural Networks Learn. Syst.*, Vol. 29, No. 11, pp. 5592–5605, 2018.
- [12] B. Hooi, K. Shin, S. Liu, and C. Faloutsos, "SMF: Drift-aware matrix factorization with seasonal patterns," in *SIAM International Conference on Data Mining, SDM 2019*, 2019, pp. 621–629.
- [13] H. Kasai, W. Kellerer, and M. Kleinsteuber, "Network Volume Anomaly Detection and Identification in Large-Scale Networks Based on Online Time-Structured Traffic Tensor Tracking," *IEEE Trans. Netw. Serv. Manag.*, Vol. 13, No. 3, pp. 636–650, 2016.
- [14] M. Mardani and G. B. Giannakis, "Estimating Traffic and Anomaly Maps via Network Tomography," *IEEE/ACM Trans. Netw.*, Vol. 24, No. 3, pp. 1533–1547, 2016.
- [15] M. Mardani, G. Mateos, and G. B. Giannakis, "Dynamic anomalography: Tracking network anomalies via sparsity and low rank," *IEEE J. Sel. Top. Signal Process.*, Vol. 7, No. 1, pp. 50–66, 2013.
- [16] R. Pasricha, E. Gujral, and E. E. Papalexakis, "Identifying and alleviating concept drift in streaming tensor decomposition," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, Vol. 11052 LNAI, pp. 327–343, Apr. 2019.
- [17] Y. Y. Lo, W. Liao, C. S. Chang, and Y. C. Lee, "Temporal Matrix Factorization for Tracking Concept Drift in Individual User Preferences," *IEEE Trans. Comput. Soc. Syst.*, Vol. 5, No. 1, pp. 156–168, Mar. 2018.
- [18] G. I. Webb, R. Hyde, H. Cao, H. L. Nguyen, and F. Petitjean, "Characterizing concept drift," *Data Min. Knowl. Discov.*, Vol. 30, No. 4, pp. 964–994, Jul. 2016.
- [19] M. Sayed-Mouchaweh, *Learning from Data Streams in Dynamic Environments*. Cham: Springer International Publishing, 2016.

مراجع

- [1] F. Khojasteh, M. Kahani, and B. Behkamal, "Concept drift detection in business process logs using deep learning," *Signal Data Process.*, Vol. 17, No. 4, pp. 33–48, 2021.
- [2] K. S. Yaghoubi M, Sebti A, "Identifying Concept Drift in Business Processes by Analyzing Type of Surveys in Event History," *J. Inf. Commun. Technol.*, Vol. 12, No. 43, 1399.
- [3] H. H. Kordi Ardestan F, Siyadati S, "An improved concept-based drift solution to update customer profiles in fraud detection systems," in *8th Annual Conference on Electronic Banking and Payment Systems*, 1397.
- [4] J. Sun, D. Tao, and C. Faloutsos, "Beyond streams and graphs," in *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '06*, 2006, p. 374.
- [5] H. F. Tork, M. Oliveira, J. Gama, S. Malinowski, and R. Morla, "Event and anomaly detection using Tucker3 decomposition," in *CEUR Workshop Proceedings*, 2012, Vol. 960, pp. 8–12.
- [6] H. Fanaee-T and J. Gama, "Event detection from traffic tensors: A hybrid model," *Neurocomputing*, Vol. 203, No. C, pp. 22–33, Aug. 2016.
- [7] J. Lu, A. Liu, F. Dong, F. Gu, J. Gama, and G. Zhang, "Learning under Concept Drift: A Review," 2019.
- [8] J. Gama, I. Žliobaitė, A. Bifet, M. Pechenizkiy, and A. Bouchachia, "A survey on concept drift adaptation," *ACM Comput. Surv.*, Vol. 46, No. 4, pp. 1–37, 2014.
- [9] R. N. Gemaque, A. F. J. Costa, R. Giusti, and E. M. Santos, "An overview of unsupervised drift detection methods," *WIREs Data Min. Knowl. Discov.*, Jul. 2020.
- [10] S. Saurav *et al.*, "Online anomaly detection with concept drift adaptation using recurrent neural networks," in *Proceedings of the ACM India Joint International Conference on Data Science and*

- Process.*, Vol. 63, No. 10, pp. 2663–2677, 2015.
- [26] X. Ding, L. He, and L. Carin, “Bayesian robust principal component analysis,” *IEEE Trans. Image Process.*, Vol. 20, No. 12, pp. 3419–3430, 2011.
- [27] J. Y. Wei, J. F. Zhao, Y. Y. Zhao, and Z. C. Zhao, “Unsupervised anomaly detection for traffic surveillance based on background modeling,” 2018.
- [28] L. Li, W. Huang, I. Y. H. Gu, and Q. Tian, “Statistical modeling of complex backgrounds for foreground object detection,” *IEEE Trans. Image Process.*, Vol. 13, No. 11, pp. 1459–1472, 2004.
- [29] C. Lin, Q. Zhu, S. Guo, Z. Jin, Y. R. Lin, and N. Cao, “Anomaly detection in spatiotemporal data via regularized non-negative tensor analysis,” *Data Min. Knowl. Discov.*, pp. 1–18, Mar. 2018.
- [30] “Enron Email Dataset.” [Online]. Available: <https://www.cs.cmu.edu/~enron/>.
- [31] “Trade (v4.0) — Correlates of War.” [Online]. Available: <https://correlatesofwar.org/datasets/bilateral-trade>.
- [32] J. Liao, J. Tang, W. Zeng, and X. Zhao, “Efficient and Accurate Traffic Flow Prediction via Incremental Tensor Completion,” pp. 2169–3536, 2018.
- [33] A. Vandervalk, D. Snyder, and J. Hajek, “US DOT Roadway Transportation Data Business Plan (Phase 1),” 2013.
- [20] T. G. Kolda and B. W. Bader, “Tensor Decompositions and Applications,” *SIAM Rev.*, Vol. 51, No. 3, pp. 455–500, Aug. 2009.
- [21] M. Amoozegar, B. Minaei-Bidgoli, M. Rezaghi, and H. Fanaee-T, “Extra-adaptive robust online subspace tracker for anomaly detection from streaming networks,” *Eng. Appl. Artif. Intell.*, Vol. 94, p. 103741, Sep. 2020.
- [22] J. P. van Brakel, “Smoothed z-score algorithm,” 2016. [Online]. Available: <http://stackoverflow.com/questions/22583391/peaksignal-detection-in-realtime-timeseries-data>. [Accessed: 23-June2019]. [Accessed: 01-Jan-2019].
- [23] H. Kasai, “Fast online low-rank tensor subspace tracking by CP decomposition using recursive least squares from incomplete observations,” *Neurocomputing*, Vol. 347, pp. 177–190, 2019.
- [24] A. Sobral, S. Javed, S. K. Jung, T. Bouwmans, and E. H. Zahzah, “Online Stochastic Tensor Decomposition for Background Subtraction in Multispectral Video Sequences,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, Vol. 2015-Febru, pp. 946–953.
- [25] M. Mardani, G. Mateos, and G. B. Giannakis, “Subspace learning and imputation for streaming big data matrices and tensors,” *IEEE Trans. Signal*

-
- ¹ Main subspace
² Robust
³ Sparse
⁴ Temporal Matrix Factorization
⁵ Seasonal Matrix Factorization
⁶ Rank
⁷ Blind
⁸ Drift aware
⁹ Informed
¹⁰ Order
¹¹ Robust Online Subspace Tracker
¹² Low rank
¹³ Informed Adaptive Robust Online Subspace Tracker
¹⁴ Peak
¹⁵ Proactive
¹⁶ Reactive
¹⁷ Normalized Residual Error
¹⁸ Background Subtraction
¹⁹ Precision
²⁰ Recall