JSST
Journal of Space Science and Technology

P-ISSN: 2008-4560
E-ISSN: 2423-4516

Journal of Space Science and Technology

Vol. 19, No. 1, pp. 95-111, 2026
https://doi.org/10.22034/jsst.2025.1558
Journal Homepage: https://www.jsstpub.com

Original Research Paper

Intelligent Detection of Spoofing Attack in Single-Frequency GNSS Receivers Using Density-Based Automatic Spatial Clustering Algorithm in Noisy Environment

Amin Mahdi Delnavaz [1], Alireza Shamsi[2], and Ebrahim Shafiee [3*]

1, 2, 3. Faculty of Electrical Engineering, University of Science and Technology of Aviation, Shahid Sattari University, Tehran, Iran Tehran, Iran

ARTICLE INFO

Article History:

Received 13 July 2025
Revised 12 October 2025
Accepted 22 October 2025
Available Online 3 November 2025

Keywords:

Machine Learning
Positioning
HDBSCAN clustering
Spoofing attack
GPS Software Receiver

ABSTRACT

With the growing application of Global Navigation Satellite Systems (GNSS), ensuring the security of these systems has become a critical priority. Among the various threats targeting GNSS, spoofing attacks pose the greatest danger, as they exploit the inability of GPS receivers to distinguish between authentic and spoofed signals. This vulnerability can lead to incorrect navigation equation solutions, ultimately resulting in erroneous position estimations. Due to the lack of a universal method for detecting spoofing attacks, various approaches have been proposed, tailored to specific requirements. One prominent category involves the use of unsupervised machine learning clustering algorithms. While density-based clustering algorithms have demonstrated promising performance, their effectiveness is often limited by a strong dependence on input parameters. In this study, spoofing signals are detected using the HDBSCAN algorithm, a hierarchical density-based clustering approach. The algorithm distinguishes spoofing signals by analyzing signal phase, cross-correlation norm, and the power differential between spoofed and genuine signals. The algorithm's performance is assessed using the Silhouette index, Dunn index, and Confusion Matrix metrics. Field tests validated the proposed algorithm, which was implemented on a software-defined receiver. The results demonstrated a spoofing signal detection accuracy of 98.43%.

* Corresponding Author's E-mail: shafiee_e@elec.iust.ac.ir

How to Cite this Article:
A.M. Delnavaz, A. Shamsi, and E. Shafiee, " Intelligent detection of spoofing attack in single-frequency GNSS receivers using density-based automatic spatial clustering algorithm in noisy environment," *Journal of Space Science and Technology*, Vol. 19, No. 1, pp. 95-111, 2026, (in Persian), https://doi.org/10.22034/jsst.2025.1558.

JSST
علوم و فناوری فضایی

شاپای چاپی: ۴۵۶۰-۲۰۰۸
شاپای الکترونیکی: ۴۵۱۶-۲۴۲۳

علوم و فناوری فضایی

سال ۱۴۰۵، دوره ۱۹، شماره ۱، صفحه ۹۵-۱۱۱
https://doi.org/10.22034/jsst.2025.1558
Journal Homepage: https://www.jsstpub.com

مقاله پژوهشی

شناسایی هوشمند حمله فریب در گیرنده‌های تک فرکانسه GNSS با استفاده از الگوریتم خوشه‌بندی خودکار فضایی مبتنی بر چگالی در محیط نویزی

امین مهدی دلنوازان[۱]، علیرضا شمسی[۲] و ابراهیم شفیعی[۳*]

۱- کارشناسی، دانشکده برق علوم و فنون هوایی، دانشگاه شهید ستاری، تهران، ایران
۲ و ۳- استادیار، دانشکده برق علوم و فنون هوایی، دانشگاه شهید ستاری، تهران، ایران

اطلاعات مقاله

تاریخچه مقاله:

دریافت ۲۲ تیر ۱۴۰۴
بازنگری ۲۰ مهر ۱۴۰۴
پذیرش ۳۰ مهر ۱۴۰۴
اولین انتشار ۱۲ آبان ۱۴۰۴

واژه‌های کلیدی:

یادگیری ماشین
موقعیت‌یابی
خوشه‌بندی HDBSCAN
حمله فریب
گیرنده نرم‌افزاری GPS

چکیده

باتوجه‌به گسترش کاربرد سامانه‌های موقعیت‌یاب جهانی GNSS، حفاظت از این سامانه به یکی از مهم‌ترین اقدامات بدل شده است. از میان حملات به این سامانه‌ها خطرناک‌ترین حملات مربوط به حمله فریب است؛ زیرا که در این حمله گیرنده سیگنال GPS نمی‌تواند بین سیگنال اصلی و سیگنال فریب تمایز قائل شود. این امر موجب حل اشتباه معادلات ناوبری و در نهایت تشخیص موقعیت اشتباه می‌شود. به علت عدم وجود یک روش یکتا برای شناسایی حمله فریب؛ روش‌های گوناگونی وجود دارند که باتوجه‌به نیاز مورداستفاده قرار می‌گیرند. یک دسته از روش‌ها استفاده از الگوریتم‌های خوشه‌بندی یادگیری ماشین بدون ناظر است. الگوریتم‌های خوشه‌بندی مبتنی بر چگالی هر چند که رویکرد خوبی از خود نشان می‌دهند بااین‌وجود به علت وابستگی شدید به پارامترهای ورودی آن چنان کارآمد نخواهند بود. در این مقاله با الگوریتم HDBSCAN که الگوریتمی سلسله‌مراتبی مبتنی بر چگالی است؛ سیگنال فریب آشکارسازی می‌گردد. این الگوریتم با تحلیل فاز، نرم هم‌بسته گیری و توان سیگنال فریب نسبت به سیگنال اصلی، جعلی بودن آن را تشخیص می‌دهد. برای بررسی عملکرد این الگوریتم از شاخص‌های Silhouette، Dunn و Confusion Matrix استفاده می‌شود. در آزمون‌های میدانی، الگوریتم پیشنهادی بر روی گیرنده نرم‌افزاری پیاده‌سازی و بادقت ۹۸/۴۳ درصد سیگنال فریب را به‌درستی تشخیص می‌دهد.

* پست الکترونیکی نویسنده مسئول: shafiee_e@elec.iust.ac.ir

How to Cite this Article:
A.M. Delnavaz, A. Shamsi, and E. Shafiee, " Intelligent detection of spoofing attack in single-frequency GNSS receivers using density-based automatic spatial clustering algorithm in noisy environment," *Journal of Space Science and Technology*, Vol. 19, No. 1, pp. 95-111, 2026, (in Persian), https://doi.org/10.22034/jsst.2025.1558.

۱ مقدمه

موقعیت‌یاب جهانی GPS[1] یکی از مهم‌ترین و کارآمدترین سامانه‌های ناوبری است که بر اساس سیگنال‌های ماهواره‌ای برای تعیین موقعیت و زمان در هر نقطه‌ای از سطح زمین استفاده می‌شود. این سامانه در ابتدا به دلیل کاربرد نظامی ابداع شد و به‌تدریج به دلیل کاربردهای عمده در زمینه‌های: ناوبری ناوگان‌ها، مخابرات سلولی، پهپادها، فعالیت ورزشی، پزشکی و مراقبت بهداشتی، موقعیت‌یابی و ردیابی اشیا و افراد وارد زندگی عموم مردم شد، به‌طوری‌که امروزه استفاده از آن به طور چشم‌گیری فزونی‌یافته است. به دلیل آن که کاربرد آن در جهان امروز انکارناپذیر است، پس حفاظت از این سامانه ناوبری کار بسیار مهمی است [۱،۶] .

یکی از مؤثرترین حملاتی که سیگنال GPS را در معرض خطر قرار می‌دهد؛ حمله فریب[2] است. حمله فریب به عملیاتی سیگنالی گفته می‌شود که دشمن یا متخاصم طی آن سعی در گمراهی گیرنده GPS دارند و سبب می‌شوند که گیرنده در حل معادلات ناوبری اشتباه نماید و درنتیجه موقعیت نادرست را به‌عنوان مکان صحیح نمایش دهد [۹، ۷].

سیگنال‌های GPS به علت این که هنگام عبور از لایه‌های جوی با توانی کم به گیرنده می‌رسند؛ بنابراین در برابر اخلال هر سیگنال دیگری آسیب‌پذیر می‌باشد. پس سیگنال فریب به علت توان بالاتر و تفاوت ناچیز با سیگنال اصلی به‌راحتی گیرنده‌هایی که توان شناسایی و حذف سیگنال فریب را ندارند؛ فریب داده و کنترل دستگاه را به‌دست می‌گیرند. گیرنده‌هایی متداول به‌شدت توسط حمله فریب آسیب‌پذیر هستند؛ بنابراین آشکارسازی سیگنال فریب و روش‌های آشکارسازی به یکی از مهم‌ترین مباحث امروزی تبدیل شده است [۱۰،۱۱].

در روش‌های مقابله با فریب دو نوع رویکرد وجود دارد؛ کاهش یا تشخیص فریب [۱۲،۱۳]. به علت این که در کاهش اثر حمله فریب نیز ابتدا باید سیگنال فریب GPS را تشخیص داد، بنابراین تشخیص سیگنال جعلی یکی از موضوعات مهم در فرآیند ضد فریب در نظر گرفته می‌شود. در آشکارسازی سیگنال فریب[3] چون روش یکتایی وجود ندارد، متناسب با نیاز روش‌های متنوعی برای رویارویی با حمله- فریب ارائه شده است. این روش‌ها می‌توانند در بخش‌های: معادلات موقعیت‌یابی، بخش داده، پردازش سیگنال و بخش موقعیت‌یابی صورت گیرد. یکی از روش‌های موثر برای شناسایی حملات فریب سیگنالGNSS ، استفاده از تحلیل همبستگی مکانی سیگنال‌ها است. این روش با بررسی تفاوت همبستگی فضایی سیگنال‌های دریافتی از ماهواره‌های واقعی و سیگنال‌های جعلی، حملات را شناسایی می‌کند [۲]. الگوریتم‌های تشخیص نفوذ IDS[4] [۳] نقش حیاتی در شناسایی الگوهای غیرعادی در ترافیک شبکه دارند و به سیستم‌های امنیتی در شناسایی فعالیت‌های مشکوک کمک می‌کنند. الگوریتم شبکه‌های مصنوعی عصبی ANN[5] [۴] برای یادگیری الگوهای پیچیده داده‌ها و شبیه‌سازی رفتارهای غیرعادی از داده‌های گذشته استفاده می‌شود. الگوریتم‌های مبتنی بر ماشین بردار پشتیبان SVM[6] [۵] با استفاده از ویژگی‌هایی چون دقت بالا و توانایی تفکیک داده‌ها در فضاهای چندبعدی، به طور خاص برای تفکیک سیگنال‌های جعلی از سیگنال‌های اصلی طراحی شده‌اند.

همچنین، روش‌هایی با گروه‌بندی داده‌های مشابه و شناسایی الگوهای غیرعادی، به‌ویژه در شناسایی حملات جدید، مؤثر هستند. شبکه عصبی چندلایه (MLP) [۶] توانایی تحلیل ویژگی‌های پیچیده و غیرخطی را دارند و برای تشخیص نفوذهای پیچیده و حملات نوظهور مفید هستند. الگوریتمK- نزدیک‌ترین همسایه‌ها (KNN) [۷] سیگنال‌های جدید را بر اساس نزدیکی به نمونه‌های آموزشی شبیه‌سازی و دسته‌بندی می‌کند. این الگوریتم‌ها به طور ترکیبی و مستقل برای شناسایی و تفکیک سیگنال‌های جعلی در ترافیک شبکه استفاده می‌شوند.

خوشه‌بندی یک روش در زمینه یادگیری ماشین و داده‌کاوی است که داده‌ها را بر اساس شباهت آن‌ها به یکدیگر خوشه‌بندی می‌کند. چند الگوریتم خوشه‌بندی عبارتند از k-means ، DBSCAN[7]، OPTICS[8]، GMM[9] و غیره که در خوشه‌بندی باتوجه‌به کاربرد استفاده می‌شوند [۱۹ ،۱۸] . در [۸] نیز توسط الگوریتم‌های خوشه- بندی DBSCAN و FCM و Subtractive شناسایی سیگنال فریب مورد آزمایش قرارگرفته است.

باتوجه به اینکه k-means یکی از معروف‌ترین و پرکاربردترین الگوریتم خوشه‌بندی می‌باشد بااین‌حال نمی‌تواند داده‌هایی را که به‌صورت توزیع غیرنرمال و با چگالی‌هایی متنوع در فضا پخش شده‌اند را خوشه‌بندی کند. الگوریتم DBSCAN یکی از شناخته‌شده‌ترین الگوریتم خوشه‌بندی است که بر اساس چگالی به عمل خوشه‌بندی می‌پردازد. یکی از مشکلات این الگوریتم وابستگی شدید آن به پارامترهای ورودی آن است و اگر این پارامترها

1. Global Positioning System
2. Spofing Attack
3. Spofing Signal Detection
4. Intrusion Detection System
5. Artificial Neural Network
6. Support Vector Machine
7. Density-Based Spatial Clustering of Applications with Noise
8. Ordering Points To Identify the Clustering Structure
9. Gaussian Mixture Modelling

به‌درستی انتخاب نکردند ممکن است حتی به خوشه‌بندی اشتباه دچار شود [۲۳، ۲۰].

الگوریتم پیشنهادی در این مقاله الگوریتم HDBSCAN[1] است که یک الگوریتم سلسله‌مراتبی مبتنی بر چگالی می‌باشد. این الگوریتم که بر مبنای دو الگوریتم DBSCAN و OPTICS ایجاد شده است به‌صورت خودکار خوشه‌بندی داده‌ها را انجام می‌دهد. در این روش نیاز به تعیین دقیق پارامترهای ورودی نمی‌باشد و الگوریتم با گرفتن داده‌ها شروع به خوشه‌بندی آن‌ها در خوشه‌های متفاوت می‌کند و این امر باعث قدرت زیاد در تشخیص حمله فریب می‌شود؛ چون در حملات فریب تعیین دقیق پارامترها برای عمل خوشه‌بندی عملی پیچیده‌ای است.

۱-۱ فریب و آشکارسازی فریب

فریب سیگنال ماهواره GPS از خطرناک‌ترین نوع حمله محسوب می‌شود؛ زیرا گیرنده سیگنال GPS نمی‌تواند بین سیگنال معتبر و فریب تمایز قائل شود. توان سیگنال فریب به علت آن که مرجع ارسال آن از سطح زمین است، تلفات مسیر ندارد. به همین علت در محل گیرنده در هنگامی که هر دو سیگنال فریب و اصلی به طور هم‌زمان در گیرنده حضور داشته باشند؛ گیرنده به‌اشتباه سیگنال فریب را دریافت کرده در حل معادلات ناوبری و در نهایت شناسایی موقعیت دقیق دچار اشتباه می‌شود [۹]. ازاین‌رو شناسایی سیگنال فریب به طور گسترده موردتوجه محققان این حوزه قرار گرفته است. باتوجه به این که حمله یکتایی برای فریب سیگنال وجود ندارد؛ ازاین‌رو روشی یکتایی نیز برای شناسایی سیگنال فریب وجود ندارد. به علت این که در چنین حملاتی آگاهی از ویژگی سیگنال جعلی در دست نیست، استفاده از الگوریتم‌های خوشه‌بندی یادگیری ماشین بدون ناظر کمک زیادی به شناسایی می‌کند [۹].

۱-۲ خوشه‌بندی

خوشه‌بندی روشی تکرارشونده در داده‌کاوی است که در آن باتوجه‌به ویژگی داده‌ها آن‌ها را به خوشه‌های مختلف تفکیک می‌کند. بدین‌گونه که داده‌هایی که از نظر یک ویژگی شبیه به هم هستند در یک خوشه قرار می‌گیرند و از داده‌های دیگر تمیز داده می‌شوند. الگوریتم OPTICS نیز که بر اساس الگوریتم DBSCAN نوشته شده هست؛ در آن حساسیت به پارامترهای ورودی کم شده است به‌گونه‌ای که دیگر نیاز به انجام مقداردهی به پارامتر ε (Epsilon) نمی‌باشد و خود الگوریتم شعاع همسایگی (ε) را بر اساس تعداد نقاط همسایگی ($MinPts$) داده شده، به دست می‌آورد [۱].

۲ روش پیشنهادی تشخیص فریب مبتنی بر الگوریتم HDBSCAN

خوشه‌بندی بر پایه چگالی یک الگوی مؤثر برای خوشه‌بندی است. روش‌های موجود مانند DBSCAN دارای محدودیت‌هایی مانند ناتوانی در خوشه‌بندی داده‌ها با تراکم بالا و حساسیت زیاد به پارامترهای ورودی می‌باشند و می‌توانند تنها یک برچسب «صاف» (بدون سلسله‌مراتب) برای داده‌ها بر اساس یک آستانه چگالی ($MinPts$) ارائه دهند. استفاده از یک آستانه چگالی نمی‌تواند مجموعه‌داده‌های دارای خوشه‌هایی با چگالی‌های بسیار متفاوت و یا تودرتو را به‌درستی مشخص کند. همچنین الگوریتم‌هایی که یک سلسله‌مراتب از خوشه‌بندی ارائه می‌دهند، قادر به ساده‌سازی خودکار سلسله‌مراتب به یک نمایش قابل‌تفسیر که تنها شامل خوشه‌های مهم‌تر باشد، نیستند. علاوه بر این سرعت این الگوریتم نسبت به الگوریتم OPTICS بسیار بیشتر است.

الگوریتم HDBSCAN در [۲] ارائه شد و هیچ‌کدام از محدودیت‌های گفته شده را ندارد و این برای تشخیص سیگنال فریب از سیگنال اصلی بسیار حائز اهمیت است. برای دست یابی به این الگوریتم، نویسندگان [۲] الگوریتم DBSCAN را مورد بازبینی قرار داده و DBSCAN* را ارائه داده و با کمک آن توانسته‌اند الگوریتم HDBSCAN را پیشنهاد کنند که در آن یک مقدار جدید برای ε به نام ε' به دست می‌آید. مراحل زیر نشان‌دهنده فرآیند دستیابی به الگوریتم مورد نظر می‌باشد [۲۵،۲۶].

۲-۱ الگوریتم DBSCAN*

فرض کنید که $X= \{x_1, \ldots, x_n\}$ یک مجموعه داده از n داده است و همچنین D یک ماتریکس $n \times n$ می‌باشد که شامل فواصل زوجی (جفتی) $d(x_p,x_q)$ هست. x_p و x_q عضوهایی از X برای یک فاصله متریک $d(\cdot,\cdot)$ هستند. در تعریف خوشه‌های مبتنی بر چگالی بر اساس داده‌های مرکزی، هدف شناسایی خوشه‌هایی می‌باشد که داده‌های مرکزی به عنوان داده اصلی و مهم آن‌ها هستند. این بدین معناست که خوشه‌ها بر اساس داده مرکزی که دارای چگالی بالا هستند، تعریف و شناسایی می‌شوند. این روش تحلیل داده‌ها بر اساس داده‌های کلیدی و مهم، به جای استفاده از تمام داده‌ها، صورت می‌گیرد و در تحلیل داده‌های پر نویز و با ابعاد بالا مفید می‌باشد [۱۰] و [۱۱]. با استفاده از تعاریف زیر می‌توان الگوریتم مورد نظر را بدست آورد [۱۲].

1. Hierarchical Density-Based Spatial Clustering of Applications with Noise

تعریف ۱. داده مرکزی:

به یک داده x_p داده مرکزی[1] گفته می‌شود اگر به شعاع $\boldsymbol{\varepsilon}$ از آن داده، حداقل شامل $MinPts$ داده در محدوده آن شعاع باشد. به داده‌ای که مرکزی نباشد داده غیرمرکزی یا پرت گفته می‌شود. باتوجه‌به رابطه (۱) داریم [۲]:

$$N_\varepsilon(x_p) = \{x \in X \mid d(x, x_p) \le \varepsilon\}$$

اگر

$$\text{آنگاه} \begin{cases} N\varepsilon(x_p) \ge MinPts & \text{داده مرکزی} \\ \text{درغیراین صورت} & \text{داده پرت} \end{cases} \quad (۱)$$

تعریف ۲. (ε – قابل‌دسترس):

دو داده مرکزی x_p و x_q نسبت به ε قابل‌دسترس هستند اگر طبق رابطه (۲) و (۳) برقرار باشند [۱].

$$x_p \in N_\varepsilon(x_q) \quad (۲)$$

$$x_q \in N_\varepsilon(x_p) \quad (۳)$$

تعریف ۳. (چگالی – متصل):

اگر دو داده مرکزی x_p و x_q به فاصله‌ای برابر یا کمتر از ε از یکدیگر قرار بگیرند و در فاصله بین آن دو حداقل $MinPts$ داده وجود داشته باشد؛ در این صورت می‌توان گفت آن دو داده از نظر چگالی متصل می‌باشند.

تعریف ٤. (خوشه):

یک خوشه C از مجموعه‌داده X، بزرگ‌ترین زیرمجموعه غیر خالی است که تمام داده‌های آن از نظر چگالی از یکدیگر قابل‌دسترس باشند.

بنابراین، الگوریتم DBSCAN* به شرح زیر است:

۱. برای هر داده x در مجموعه X اگر x هنوز برچسب نخورده باشد:

- پیداکردن همسایگان x (شامل خود x) که فاصله آن‌ها از x کمتر یا مساوی ε است.
- اگر تعداد همسایگان بیشتر یا مساوی $MinPts$ باشد:
- داده‌های همسایه را به‌عنوان داده مرکزی برچسب زده و آن‌ها را به یک خوشه جدید اضافه می‌کند.
  - خوشه را با اضافه‌کردن همسایگان آن به فهرست داده‌های موجود در خوشه، گسترش می‌دهد.
- در غیر این صورت، x را به‌عنوان یک داده نویزی برچسب زده.

۲. خروجی الگوریتم شامل فهرست خوشه‌ها و داده‌های با برچسب نویز است.

این الگوریتم به طور مفهومی خوشه‌ها را به‌عنوان اجزای متصل یک گراف تعریف می‌کند که داده‌های X به‌عنوان رئوس آن و درصورتی‌که داده‌های متناظر از یکدیگر قابل‌دسترس باشند هر زوج رأس، مجاور یکدیگر هستند. داده‌های غیرمرکزی به‌عنوان داده نویزی معرفی می‌شوند [۲۸،۱].

۲-۲ HDBSCAN؛ الگوریتم DBSCAN* براساس سلسله مراتب

پارامتر ورودی $MinPts$ یک عامل پردازنده کلاسیک در تخمین چگالی است. برای فرموله‌کردن صحیح سلسله‌مراتبی مبتنی بر چگالی نسبت به یک مقدار $MinPts$؛ از تعاریف زیر کمک گرفته می‌شود.

تعریف ۵. (فاصله هسته'ε):

فاصله هسته‌ای یک داده x_p؛ برابر فاصله x_p تا نزدیک‌ترین داده همسایه خود است که حداقل شامل $MinPts$ داده باشد. طبق رابطه (۴) داریم [۱]:

$$N_\varepsilon(x_p) = \{x \in X | d(x, x_p) \le \varepsilon\}$$

اگر

$$\text{آنگاه} \begin{cases} \text{Nε(xp)} <= \varepsilon \\ \text{Nε(xp)} = \text{dcore(xp)} \end{cases} \quad (۴)$$

تعریف ٦. (داده مرکزی نسبت به 'ε):

داده x_p به شرطی داده مرکزی ε' محسوب می‌شود که برای هر مقدار ε کوچک‌تر یا مساوی فاصله هسته x_p باشد. یعنی طبق رابطه (۵) [۱]:

$$d_{core}(x_p) \le \varepsilon \quad (۵)$$

تعریف ۷. (فاصله قابل‌دسترسی متقابل):

فاصله قابل‌دسترسی متقابل بین دو داده x_p و x_q باتوجه‌به رابطه (۶) برابر است با [۲]:

$$d_{mreach}(x_p, x_q) = max\{d_{core}(x_p), d_{core}(x_q), d(x_p, x_q)\} \quad (۶)$$

در این تعریف d_{mreach} بزرگ‌ترین فاصله در میان دو فاصله مرکزی x_p و x_q و همچنین فاصله بین x_p و x_q است.

تعریف ۸. (گراف قابل دسترسی متقابل):

این تعریف یک گراف کامل به نام G_{mpts} است که در آن هر یک از داده‌های مجموعه X به‌عنوان یک رأس در گراف نمایش داده می‌شود و وزن هر یال برابر بافاصله قابل‌دسترسی متقابل بین هر زوج داده می‌باشد.

1. Core Objects

با استفاده از گراف G_{mpts} می‌توان به گراف $G_{mpts,\varepsilon'}$ دست‌یافت به این‌گونه که تمام یال‌هایی‌که وزن بزرگ‌تری نسبت به ε' دارند، حذف می‌شوند. باتوجه‌به تعاریف گفته شده در این صورت می‌توان گفت که خوشه‌ها بر اساس الگوریتم DBSCAN* (نسبت به $MinPts$ و ε')؛ داده‌های قابل‌دسترس از یکدیگر در $G_{mpts,\varepsilon'}$ هستند و داده‌های باقی‌مانده به‌عنوان نویز شناخته می‌شوند؛ بنابراین، تمام تقسیم‌بندی‌های DBSCAN* برای مقادیر مختلف ε در بازه $[0,\infty)$ می‌توانند به‌صورت پی‌درپی با حذف یال‌ها به ترتیب کاهشی از G_{mpts} تولید شوند؛ پس می‌توان یک سلسله‌مراتب از داده‌ها بر اساس چگالی به دست آورد [۱، ۲۹، ۳۰].

۲-۳ خوشه‌بندی غیر سلسله‌مراتبی بهینه

برای شناسایی حمله فریب نیاز به یک آستانه چگالی است که عمل خوشه‌بندی بر اساس آن به‌درستی انجام گیرد. در وقوع حمله فریب به علت نامفهوم بودن نحوه توزیع سیگنال و چگونگی توزیع چگالی، پیداکردن یک مقدار بهینه بسیار دشوار و با آزمون‌وخطا همراه است. درضمن در هنگام وقوع این حمله زمان زیادی برای این امر در دست مقابله‌کننده با حمله فریب نمی‌باشد و چه‌بسا اگر این مقدار به‌درستی انتخاب نشود، سبب عدم تشخیص فریب می‌شود. در الگوریتم HDBSCAN با استفاده از پایداری خوشه‌ها و بیشینه بهینه‌سازی می‌توان به یک مقدار خودکار برای هر داده‌ای با هر توزیعی دست‌یافت [۱۳].

۲-۳-۱ پایداری خوشه

خوشه‌ها به‌عنوان زیرمجموعه‌ای از بزرگ‌ترین ارتباط بین چگالی‌های مشابه تشکیل شده‌اند که این بیان تعریف الگوریتم‌های خوشه‌بندی مبتنی بر چگالی می‌باشد. اگر x ویژگی داده باشد، تفاوت بین این الگوریتم‌ها بر اساس چگونگی تخمین چگالی $f(x)$ در سطح λ است. برای الگوریتم BDSCAN* با فرض اینکه $\lambda = 1/\varepsilon$، از استفاده با باشد چگالی آستانه یک $1/dcore(x)$ خوشه‌ها و یک تخمینی از چگالی مانند $f(x)$؛ خطوط چگالی را تخمین می‌زنند. HDBSCAN تمام حالت‌های قابل‌انجام DBSCAN* را نسبت به یک آستانه داده شده برابر $MinPts$ و همه آستانه‌های برابر بازه در $[0,\infty)$ $\lambda = 1/\varepsilon$ را تولید می‌کند. با افزایش سطح چگالی λ، خوشه‌ها کوچک‌تر می‌شوند، تا زمانی که از بین می‌روند یا به زیر خوشه‌ها شکسته می‌شوند. خوشه‌های برجسته زمان بیشتری بعد از ظاهر شدنشان "زنده" (باقی) می‌مانند. باتوجه‌به مفهوم فزونی جرم؛ فرض کنید که سطح چگالی λ را افزایش داده و با این کار یک خوشه دور-توانی (چگالی داده‌ها با دور شدن از داده مرکزی به صورت توانی کاهش می‌یابد) با چگالی C_i در سطح $\lambda_{min}(C_i)$ ظاهر می‌شود [۱] و [۳۲]. فزونی جرم C_i توسط رابطه (۷) بیان شده [۱] و در شکل (۱) نشان داده می‌شود، جایی که نواحی تیره‌تر فزونی‌جرم سه خوشه، C_3، C_4، و C_5 را نشان می‌دهد. فزونی جرم C_2 (که در تصویر معرفی نشده) شامل فزونی‌های نماینده‌های خود یعنی C_4 و C_5 است.

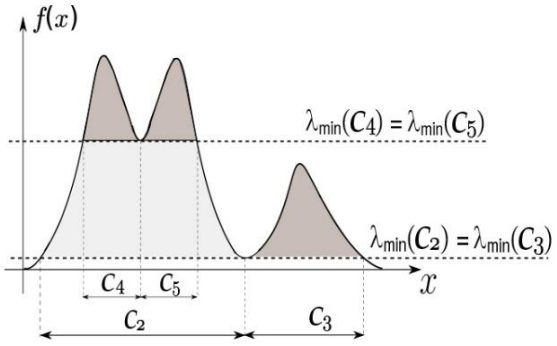
$$E(C_i) = \int_{x\in C_i} (f(x) - \lambda_{min}(C_i))\, dx \qquad (۷)$$

توده نسبی[1] (کمیتی که یک خوشه دارا می‌باشد و کمیتی که از همان خوشه مورد انتظار است)؛ رفتار یکنواختی را در امتداد هر شاخه از درخت خوشه‌ای سلسله‌مراتبی حاصل از الگوریتم HDBSCAN ، از خود نشان می‌دهد. این معیار نمی‌تواند برای مقایسه پایداری خوشه‌های توده‌ای درون یافته مانند C_2 در مقابل C_4 و C_5 استفاده شود. برای این منظور، مفهوم اضافه توده نسبی یک خوشه C_i طبق رابطه (۹) در سطح $\lambda_{min}(C_i)$ به دست می‌آید [۱]:

$$E_R(C_i) = \int_{x\in C_i} (\lambda_{max}(x, C_i) - \lambda_{min}(C_i))\, dx \qquad (۹)$$

که در آن طبق رابطه (۱۰) به دست می‌آید:

$$\lambda_{min}(x, C_i) = min\{f(x), \lambda_{max}(C_i)\} \qquad (۱۰)$$



شکل ۱- تصویر تابع چگالی، خوشه‌ها و مازاد جرم [۱]

Fig. 1. The density function, clusters, and mass excess

$\lambda_{max}(C_i)$ سطح چگالی است که خوشه C_i در آن تقسیم یا ناپدید می‌شود. به‌عنوان‌مثال، برای خوشه C_2 در شکل (۱)، داریم: $\lambda_{max}(C_2) = \lambda_{min}(C_4) = \lambda_{min}(C_5)$. اضافه توده نسبی متناظر با آن، توسط ناحیه پررنگ‌تر در شکل (۱) نمایش‌داده‌شده است.

1. Excess of Mass

برای الگوریتم سلسله‌مراتبی HDBSCAN که دارای یک مجموعه‌داده محدود X می‌باشد، برچسب‌های خوشه و آستانه‌های چگالی، مرتبط با هر سطح سلسله‌مراتب هستند. با تطبیق معادله (۹) پایداری یک خوشه Ci را می‌توان طبق رابطه (۱۱) تعریف کرد [۱].

$$S(C_i) = \sum_{x_i \in C_i} (\lambda_{max}(x_i, C_i) - \lambda_{min}(C_i)) = \sum_{x_i \in C_i} \left(\frac{1}{\varepsilon_{min}(x_i, C_i)} - \frac{1}{\varepsilon_{max}(C_i)} \right) \quad (۱۱)$$

که در آن $\lambda_{min}(C_i)$ نشان‌دهنده حداقل سطح چگالی می‌باشد که خوشه Ci در آن وجود دارد. $\lambda_{max}(x_j, C_i)$ نشان‌دهنده سطح چگالی است که فراتر از آن داده x_j دیگر به خوشه C_i تعلق ندارد و $\varepsilon_{max}(C_i)$ و $\varepsilon_{min}(x_j, C_i)$ نیز مقادیر متناظر ε هستند[۱].

شکل‌ها و نمودارها

ئد ریشه C_1 و $S(C_i)$ به معنای ارزش پایداری هر خوشه است. فرض کنید خوشه‌های به‌دست‌آمده از درخت سلسله‌مراتبی ساده شده توسط الگوریتم HDBSCAN شامل $\{C_2,\ldots,C_\kappa\}$ به جز C_1 و $S(C_i$ باشد. هدف آن است که بتوان خوشه‌های برجسته‌تر که احتمالاً همراه با نویز هستند را به‌صورت «صاف» و غیرتداخلی استخراج کرد. می‌توان یک مسئله بهینه‌سازی را باهدف بیشینه‌کردن جمع پایداری خوشه‌های استخراج‌شده طبق رابطه (۱۲) فرموله کرد [۱]:

$$\max_{\delta_2,\ldots,\delta_k} \quad J = \sum_{i=2}^{k} \delta_i S(C_i) \qquad subject\ to \begin{cases} \delta \in \{0,1\}, \ i = 2,\ldots,k \\ \sum_{j=I_h} \delta_j = 1, \ \forall h \in L \end{cases} \quad (۱۲)$$

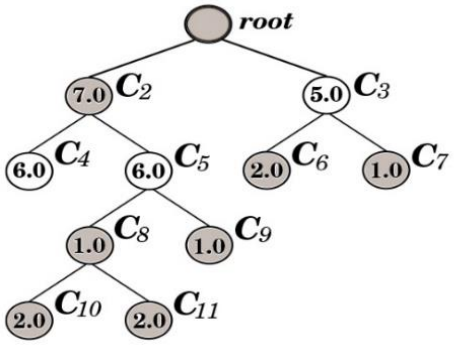
که در آن $\delta_i (i = ۲, \ldots, \kappa)$ نشان‌دهنده این است که آیا خوشه C_i در صفحه یک‌پارچه خوشه[1] وارد شده است:

$\{C_h$، بزرگترین خوشه باشد $| h \} L =$ همچنین. $(\delta_i = ۰)$ خیر یا $(\delta_i = ۱)$

مجموعه‌ای از نمایه‌های خوشه‌های با جرم مازاد[2] را نشان می‌دهد، جایی که C_h یک خوشه بزرگ است. I_h نیز مجموعه‌ای از نمایه‌های تمام خوشه‌هایی می‌باشد که بر روی مسیر از خوشه C_h تا ریشه (به‌جز خود ریشه) به همراه خوشه‌های زیرمجموعه C_h قرار دارند. در این تعریف خوشه C_h نیز در محدوده قرار می‌گیرد. این محدودیت‌ها از انتخاب خوشه‌های تودرتو بر روی مسیر یک‌پارچه خوشه جلوگیری می‌کنند، به این معنا که اگر یک خوشه وارد صفحه رشد خوشه شود؛ خوشه‌های بزرگ‌تر نتوانند بر روی مسیر رشد خوشه تا ریشه، وارد شوند.

برای حل مسئله (۱۲)، می‌توان هر گره (بخشی از یک ساختار درختی یا گراف می‌باشد که دارای اطلاعات، پیوندها و ویژگی‌های خاصی است) را به جز خود ریشه پردازش کرد. از چگالی‌های کم به زیاد حرکت کرده و در هر گره Ci تصمیم گرفته شود که آیا باید خود گره Ci؛ یا بهترین انتخابی که تا آن لحظه در زیر درخت‌های گره Ci انجام شده است، انتخاب شود. برای اینکه بتوان این تصمیم را به‌صورت محلی در گره Ci گرفت، می‌توان مجموع ثبات‌های $\hat{s}$ (Ci) تمام خوشه‌های انتخاب شده در زیر درختی که ریشه آن گره Ci است را با استفاده از یک روش بازگشتی، منتقل و به‌روزرسانی کرد [۱]. طبق رابطه (۱۳):

$$\hat{S}(C_i) = \begin{cases} S(C_i), & \text{اگر } Ci \text{ یک گره اصلی باشد} \\ max\{S(C_i), \hat{S}(C_{il}) + \hat{S}(C_{ir})\} & \text{اگر } Ci \text{ یک گره داخلی باشد} \end{cases} \quad (۱۳)$$



شکل ۲- تصویری از انتخاب بهینه از خوشه‌های یک درخت خوشه‌ای معین [1]

Fig. 2. Optimal selection of clusters from a given cluster tree [1]

شکل (۲) خوشه‌بندی سلسله‌مراتبی را نشان می‌دهد. این الگوریتم با خوشه‌های فرد از C_1 تا C_{11} شروع می‌شود و آن‌ها را بر اساس یک معیار شباهت به‌صورت تکراری ادغام می‌کند. اعداد داخل دایره‌ها هزینه (هزینه نشان‌دهنده فاصله یا تفاوت بین ویژگی‌های داده‌ها در هر خوشه است) ادغام این خوشه‌ها را نشان می‌دهد. هدف یافتن مجموعه بهینه از خوشه‌ها است که کمترین هزینه را ایجاد کنند. این الگوریتم، مجموعه‌های مختلف از خوشه‌ها را مقایسه می‌کند و مجموعه‌هایی با هزینه‌های بالاتر را نادیده می‌گیرد [۱]. باتوجه‌به شکل (۲) طبق فرآیند آورده شده:

1. flat
2. leaf clusters

(C_{10}، C_{11}) بهتر از C_8 است؛ زیرا که ادغام C_{10} و C_{11} نسبت به نگه‌داشتن C_8 به‌تنهایی باعث کاهش هزینه می‌شود؛ بنابراین، C_8 رد شده است. (C_{10}، C_{11}، C_9) از C_5 بدتر است؛ زیرا ادغام C_{10}، C_{11} و C_9 نسبت به نگه‌داشتن C_5 به‌تنهایی باعث افزایش هزینه می‌شود؛ بنابراین، این گروه رد می‌شود. (C_4) و (C_5) از (C_2) بهتر هستند؛ زیرا که ادغام C_4 و C_5 نسبت به نگه‌داشتن C_2 به‌تنهایی باعث کاهش هزینه می‌شود، پس C_2 رد می‌شود.

C_3 از (C_6، C_7) بهتر است، چون C_3 یک انتخاب بهتر نسبت به ادغام C_6 و C_7 است، بنابراین C_6 و C_7 رد می‌شوند. راه‌حل بهینه نهایی شامل خوشه‌های C_3، C_4 و C_5 با هزینه کل ۱۷ است. این راه‌حل نشان‌دهنده بهترین تعادل بین کمینه‌کردن هزینه و حفظ تعداد منطقی از خوشه‌ها است. در نتیجه، الگوریتم به طور مکرر خوشه‌ها را بر اساس هزینه‌شان ادغام می‌کند و به‌تدریج خوشه‌هایی با هزینه‌های بالاتر را کنار می‌گذارد، و در نهایت به راه‌حل بهینه با کمترین هزینه کل می‌رسد.

۳ گام‌های آزمودن رویکرد پیشنهادی

۳- ۱ گام اول - داده‌یابی

فریب سیگنال یکی از زیرشاخه‌های جنگ الکترونیک است که با هدف اختلال در دریافت، تحلیل و استفاده از سیگنال‌های دشمن انجام می‌شود. این عملیات به سه سطح ساده، متوسط و پیشرفته تقسیم می‌شود. فریب ساده شامل اقدامات اولیه‌ای مانند ارسال سیگنال‌های جعلی یا گمراه‌کننده است که طراحی پیچیده‌ای ندارند و برای انحراف سامانه‌های اولیه دشمن، مانند رادارهای ساده یا حسگرهای قدیمی، به کار می‌روند. فریب متوسط شامل تکنیک‌های پیشرفته‌تر، از جمله تغییر پارامترهای سیگنال مانند فرکانس، دامنه یا فاز برای تقلید دقیق‌تر از سیگنال‌های واقعی است. این روش برای اختلال در سامانه‌های پیچیده‌تر و کاهش احتمال شناسایی فریب به کار می‌رود. فریب پیشرفته به معنای استفاده از سامانه‌های هوشمند و تکنیک‌های پیچیده مانند پردازش سیگنال تطبیقی، تحلیل داده‌های بلادرنگ و هماهنگی چندسکویی است تا نه‌تنها سیگنال‌های دشمن را شبیه‌سازی، بلکه رفتار آنها را پیش‌بینی و پاسخ مناسب را تولید کند. این سطح از فریب به طور ویژه در مقابله با سامانه‌های راداری یا مخابراتی مدرن و مبتنی بر هوش مصنوعی اهمیت دارد [۱۴،۱۵،۱۶].

برای ارزیابی الگوریتم پیشنهادی نیاز به بررسی یک حمله فریب واقعی است. به این منظور داده‌های مربوط به یک حمله فریب متوسط مورد بررسی قرار داده می‌شود. در حمله فریب به صورت شبیه‌سازی با شدت متوسط، سیگنال اصلی GPS توسط شبیه‌ساز بازتولید شده و سپس با توانی بیشتر از سیگنال واقعی ارسال می‌شود. در نتیجه، آنتن گیرنده نرم‌افزاری به‌طور هم‌زمان سیگنال واقعی و سیگنال جعلی را دریافت می‌کند. در شکل (۳) قسمت‌های مختلف سخت‌افزار همانند فریب دهنده، سامانه فرسنده، آنتن گیرنده و گیرنده GNSS را مشاهده می‌نمایید. حال با توجه به ویژگی‌های دلتا، نرم توزیع تابع همبستگی و توان سیگنال در پهنای باند معینی از سیگنال GPS با نرخ نمونه‌برداری ۵/۷، نمونه‌برداری از سیگنال انجام می‌شود. Δ و ELP و ξ به ترتیب نرم توزیع تابع همبستگی و تغییرات فاز و توان سیگنال هستند که به‌صورت روابط (۱۴) تا (۱۶) نشان‌داده‌شده است [۶].

بنابراین، داریم:

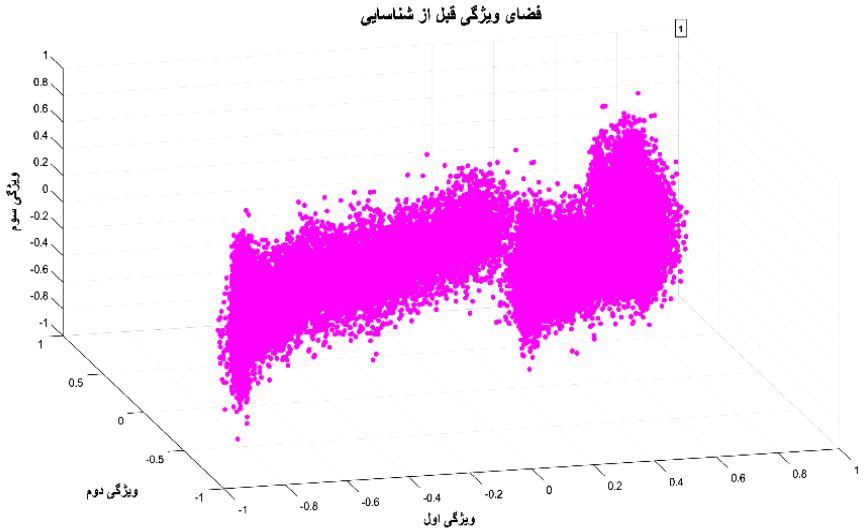
$$\Delta_\tau(t) = \left(\frac{I_{E,\tau}(t) - I_{L,\tau}(t)}{2I_p(t)}\right) \qquad (۱۴)$$

$$ELP_\tau(t) = tan^{-1}\left(\frac{Q_{L,\tau}(t)}{I_{L,\tau}(t)} - \frac{Q_{E,\tau}(t)}{I_{E,\tau}(t)}\right) \qquad (۱۵)$$

$$\xi = \frac{1}{T}\int_T |x(t,\tau)|^2 d_\tau \qquad (۱۶)$$

شکل ۳- سخت‌افزار مورد استفاده برای ذخیره‌سازی سیگنال معتبر و سیگنال فریب GPS

Fig. 3. Hardware employed for the storage of authentic and GPS spoofing signals



شکل ٤- نمایش داده سیگنال اصلی و فریب در فضای ویژگی‌های پیشنهادی قبل از آشکارسازی حمله فریب.

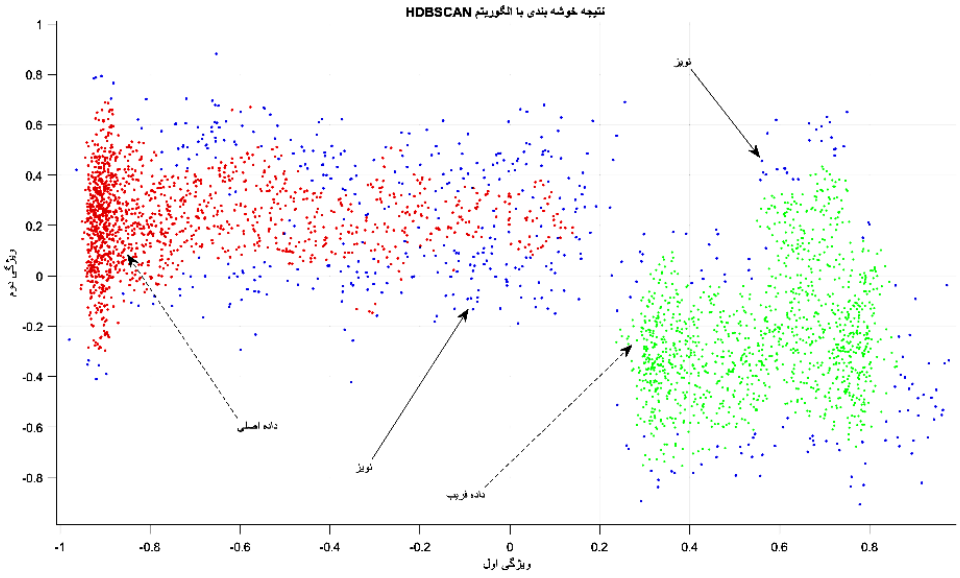
Fig. 4. Representation of authentic and spoofing signals in the proposed feature space prior to spoofing attack detection

که در آن I مؤلفه حقیقی و Q مؤلفه موهومی ردیاب گیرنده است. E و L و P به ترتیب مؤلفه زود، دیر و بی‌درنگ هم‌بسته‌ساز شاخه ردیابی و τ میزان فاصله این مؤلفه‌ها از یکدیگر است. شکل (۴) نشان‌دهنده سیگنال نمونه‌برداری شده معتبر در حضور فریب قبل از شناسایی حمله فریب می‌باشد [۶].

گام دوم - استفاده از الگوریتم برای شناسایی داده فریب

به‌منظور آشکارسازی حمله فریب، الگوریتم پیشنهادی توسط نسخه ۲۰۲۳b نرم‌افزار متلب (MATLAB) بر روی سیستمی با رم ۱۶ گیگابایت و CPU ۳.۲۰ گیگاهرتز و همچنین با سیگنال نمونه‌برداری شده ۸ بیتی با فرکانس ۵.۷؛ پردازش می‌شود. شاخص‌های موردنیاز الگوریتم، استخراج شده و بعد از پردازش اگر هر تغییری در روند سیگنال مانند تغییر الگوی سیگنال مشاهده شود به یک خوشه جدید تبدیل شده و سیگنال فریب شناسایی می‌شود.

شکل (۵) نشان‌دهنده نتیجه خوشه‌بندی توسط الگوریتم HDBSCAN است. در این تصویر داده‌های به رنگ صورتی نشان‌دهنده سیگنال اصلی و داده‌های به رنگ آبی نشان‌دهنده سیگنال فریب هستند. نقاط سبز نشان‌دهنده نویز هستند. این الگوریتم می‌تواند در محیط‌های پر نویز عمل خوشه‌بندی را به‌درستی انجام دهد.



شکل ٥- نتیجه خوشه‌بندی با الگوریتم HDBSCAN

Fig. 5. Clustering results using the HDBSCAN algorithm

گام سوم – پیاده‌سازی سخت‌افزاری

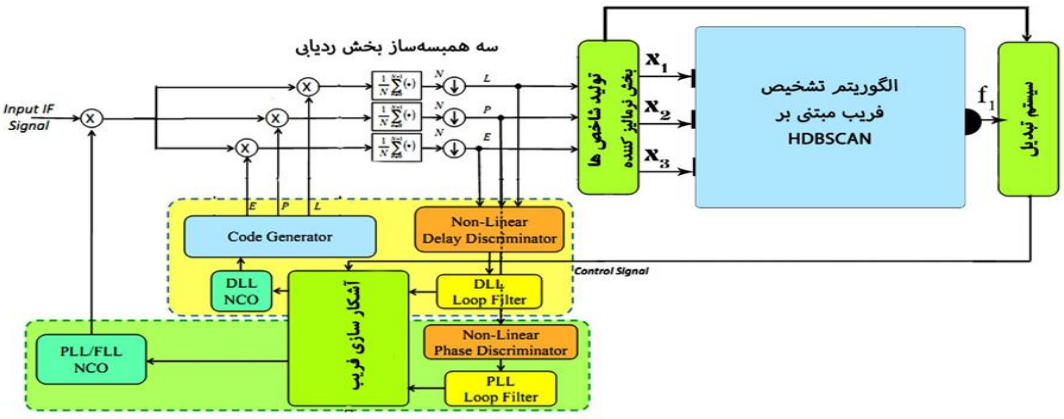
الگوریتم پیشنهادی بر روی یک گیرنده نرم‌افزاری (SDR) سیگنال GPS برنامه‌ریزی گردیده و در هنگام حمله فریب توانایی شناسایی این حمله داراست. هنگام روی‌دادن حمله فریب پردازنده مرکزی را با یک پرچم از حضور سیگنال فریب آگاه می‌نماید. از مزایای این گیرنده‌های نرم‌افزاری می‌توان به قابلیت برنامه‌ریزی سریع، حجم فیزیکی کم، قیمت مناسب و غیره اشاره کرد [۶]. بلوک دیاگرام کلی این پژوهش در شکل (۶) نشان‌داده شده است.

٤ ارزیابی الگوریتم پیشنهادی

در به‌منظور ارزیابی عملکرد آشکارساز حمله فریب GPS با استفاده از داده‌های در تست‌های میدانی واقعی و با گیرنده رادیویی ثبت شده‌اند، سناریوهای مختلف فریب از جمله فریب تأخیری، فریب باز انتشار، فریب شبیه‌ساز و فریب پیشرفته سنکرون اجرا و سیگنال RF ذخیره‌سازی شده است. در نرم‌افزار MATLAB برنامه‌ای نوشته شده است که داده‌های هر یک از سناریوهای اجرا شده را دریافت کرده و مشابه گیرنده RF، اطلاعات لحظه‌ای گیرنده را در اختیار الگوریتم آشکارساز فریب قرار می‌دهد. با استفاده از این تکنیک، می‌توان بارها و بارها از صحت عملکرد الگوریتم تشخیص فریب به‌خوبی مطمئن شد. در قسمت رفع باگ (Debug) نرم‌افزار و با قراردادن Breakpoint در خطوط مهم الگوریتم نوشته شده، می‌توان در زمان‌های حمله فریب یا خروج از فریب، پارامترهای مهم گیرنده، متغیرهای مربوط به الگوریتم‌های تشخیص فریب و پرچم فریب را بررسی کرد. در جدول (۱) انواع سناریوها واقعی سنجش شده و وضعیت گیرنده آورده شده است. شایان ذکر است که عملکرد الگوریتم خروج در لحظه می‌باشی.

سناریوی فریب به‌صورت مختصر تشریح می‌شود:

سناریوی شماره (۱) فریب با شبیه‌ساز است. این داده در محیط دانشگاه علم و صنعت ایران ضبط شده است. در شکل (۶)، تصویر نقشه خیابانی سناریو حمله فریب GPS و فاصله بین مکان اصلی گیرنده و مکان فریب‌خورده گیرنده نشان‌داده‌شده است. مکان فریب همان‌طور که در شکل (۷) مشخص شده است، مسیر فریب بیضی‌شکل اطراف فرودگاه مهرآباد تهران است. شکل (۸)، نقشه ماهواره‌ای مکان فریب‌خورده سناریو حمله فریب تعریف شده را در اطراف فرودگاه بین‌المللی مهرآباد تهران نشان می‌دهد. شکل (۷–۹)، تصویر ماهواره‌ای مکان پیاده‌سازی آشکارساز حمله فریب GPS را در ورزشگاه دانشگاه علم و صنعت ایران نشان می‌دهد. در لحظه ۱۳:۲۵:۲۶، سیگنال جمینگی برای NO FIX شدن گیرنده، فرستاده می‌شود. سرانجام در لحظه ۱۳:۳۲:۴۰ ساعت گیرنده به ۰۰:۰۰:۲۸ و تاریخ از ۰۶/۲۷/۲۰۲۳ به ۱۲/۲۱/۲۰۲۱ تغییر می‌کند. همچنین عرض جغرافیایی به مقداری اندک تغییر می‌کند؛ بنابراین، بلافاصله در حدود چند میلی‌ثانیه پس از شروع حمله فریب، تابع تصمیم‌گیر اعلام فریب می‌کند. گیرنده همچنان در حالت NO FIX قرار دارد و موقعیت‌یابی صحیح صورت نگرفته است. در لحظه ۰۰:۰۴:۵۲، گیرنده به سیگنال‌های معتبر سوئیچ کرده و پرچم اعلام فریب همچنان فعال است. اما به دلیل اینکه هنوز عرض جغرافیایی FIX نشده، به چند ثانیه تأخیر، پرچم فریب مربوطه غیرفعال می‌شود. تابع تصمیم‌گیر به‌صورت کلی، صفر کردن پرچم فریب را تصمیم‌گیری می‌کند و خروج از فریب در این آزمایش به‌خوبی توسط آشکارساز اعلام می‌شود. گیرنده‌های متداول هر ثانیه یک موقعیت خروجی می‌دهد. در گیرنده‌های هوانوردی خروجی ۱۰ هرتز می‌باشد. از لحاظ میزان بلادرنگ بودن الگوریتم تشخیص فریب می‌تواند در زمان قابل‌قبولی تشخیص انجام می‌شود. در جدول (۲) بلادرنگ بودن الگوریتم پیشنهادی با الگوریتم‌های مورد بحث مقایسه شده است.



شکل ٦– بلوک دیاگرام روش پیشنهادی.

Fig. 6. Block diagram of the proposed method

جدول ۱– مشخصات سناریو و عملکرد آشکارساز در لحظه ورود و خروج سیگنال فریب از گیرنده

Table 1. Scenario specifications and detector performance at the onset and offset of the spoofing signal

Parameters			
Senario No.	Receiver Motion State	Simulated GPS Spoofing Frequencies	Performance: Detection Time (ms) after Spoofing Entry
1	Stationary	L1=1/575 GHz	335
2	Mobile	L1=1/575 GHz	224
3	Stationary	L1=1/575 GHz L2=1/227 GHz	262
4	Mobile	L1=1/575 GHz L2=1/227 GHz	281
5	Mobile	L1=1/575 GHz L2=1/227 GHz	249
6	Stationary	L1=1/575 GHz L2=1/227 GHz	245
7	Stationary	L1=1/575 GHz	361

شکل ۷– تصویر نقشه خیابانی سناریو حمله فریب GPS تعریف شده در دانشگاه علم و صنعت ایران

Fig. 7. Street map of the GPS spoofing attack scenario defined at Iran University of Science and Technology (IUST)

شکل ۸- تصویر نقشه ماهواره‌ای مکان فریب خورده سناریو حمله فریب GPS اطراف فرودگاه بین‌المللی مهرآباد تهران.

Fig. 8. Satellite imagery showing the spoofed location within the GPS spoofing attack scenario near Tehran Mehrabad International Airport

شکل ۹- تصویر نقشه ماهواره‌ای مکان اجرای سناریو حمله فریب GPS تعریف شده در ورزشگاه دانشگاه علم و صنعت ایران.

Fig. 9. Satellite imagery of the execution site for the GPS spoofing attack scenario defined at the Iran University of Science and Technology (IUST) stadium

۴-۱ صحت‌سنجی روش‌های پیشنهادی بر اساس معیار Silhouette و Dunn و Confusion Matrix[۱]

نویسندگان ماتریس درهم‌ریختگی در مواقعی مورداستفاده قرار می‌گیرد که تعلق به یک خوشه یا دسته مورد اهمیت واقع شود. باتوجه‌به این معیار چهار حالت زیر رخ می‌دهد که در حالات زیر خوشه منفی همان عدم وجود فریب و خوشه مثبت وجود فریب می‌باشد:

- نمونه عضو خوشه مثبت باشد و عضو خوشه مثبت تشخیص داده شود (مثبت صحیح یا True Positive(TP)
- نمونه عضو خوشه مثبت باشد و عضو خوشه منفی تشخیص داده شود (منفی کاذب یا False Negative(FN)
- نمونه عضو خوشه منفی باشد و عضو خوشه منفی تشخیص داده شود (منفی صحیح یا True Negative (TN)

۱ ماتریس درهم‌ریختگی

- نمونه عضو خوشه منفی باشد و عضو خوشه مثبت تشخیص داده شود (مثبت کاذب یا False Positive (FP)

از پارامترهای گوناگونی می‌توان برای تحلیل ماتریس درهم‌ریختگی استفاده کرد که یکی از آن‌ها، پارامتر صحت[1] است. پارامتر صحت متداول‌ترین و ساده‌ترین معیار برای ارزیابی کیفیت یک خوشه‌بندی می‌باشد که نشان‌دهنده میزان الگوهایی است که درست تشخیص داده شدند و طبق رابطه (۱۷) فرموله می‌شود:

$$Accuracy = \frac{(TP + TN)}{(TP + FN + FP + TN)} \qquad (۱۷)$$

شاخص نیم‌رخ[2]، هم به میزان پیوستگی داده‌ها در یک خوشه و هم به میزان تفکیک داده‌ها از خوشه مجاور بستگی دارد. این معیار برای هر داده میزان وابستگی آن داده را به خوشه‌اش در مقایسه با خوشه مجاور بررسی می‌کند. مقدار این شاخص بین ۱− تا ۱+ متغیر است. مقدار نزدیک به ۱ بیانگر انطباق خوب بین داده و خوشه‌اش نسبت به خوشه مجاور است. اگر معیار نیم‌رخ برای همه نقاط درون خوشه‌ها نزدیک به ۱ باشد، عمل خوشه‌بندی به‌درستی انجام شده است. درحالی‌که کوچک بودن مقدار نیم‌رخ برای خوشه‌ها، بیانگر ضعیف‌بودن الگوریتم خوشه‌بندی در انجام خوشه‌بندی است.

شاخص دان[3] یک معیار آماری است که برای ارزیابی کیفیت خوشه‌بندی در داده‌ها مورداستفاده قرار می‌گیرد. این شاخص به بررسی میزان تفکیک‌پذیری بین خوشه‌ها و فشردگی درون هر خوشه می‌پردازد. مقدار بالاتر شاخص دان نشان‌دهنده خوشه‌های بهتری است، چرا که فاصله بین خوشه‌ها بیشتر و توزیع داده‌ها درون هر خوشه فشرده‌تر است. این شاخص با استفاده از نسبت کوچک‌ترین فاصله بین خوشه‌ها به بزرگ‌ترین قطر درونی خوشه‌ها محاسبه می‌شود و برای ارزیابی الگوریتم‌های خوشه‌بندی در حوزه‌هایی نظیر داده‌کاوی و یادگیری ماشین کاربرد فراوان دارد [۱۴].

شکل (۱۱) نشان‌دهنده شاخص نیم‌رخ برای الگوریتم HDBSCAN است. جدول (۲و۳) نتیجه شاخص دان و صحت را برای چند الگوریتم نشان می‌دهد. همچنین مزایا و معایب رویکردهای مختلف تشخیص فریب را مبتنی بر خوشه‌بندی مراجع مختلف در جدول (۴) با یکدیگر مقایسه شده‌اند. بر اساس نتایج حاصل از مقایسه‌ی روش‌های مختلف تشخیص فریب در جدول (۴)، مشخص می‌شود که الگوریتم پیشنهادی HDBSCAN در مقایسه با سایر الگوریتم‌ها از نظر دقت، پایداری و کارایی در محیط‌های نویزی برتری قابل توجهی دارد. این الگوریتم با بهره‌گیری از ساختار سلسله‌مراتبی مبتنی بر چگالی و بدون نیاز به تعیین پارامترهای اولیه و تعداد خوشه‌ها، توانسته است داده‌هایی با چگالی‌های متفاوت را به‌صورت خودکار و دقیق خوشه‌بندی و تفکیک کند. مزیت اصلی روش پیشنهادی در مقایسه با الگوریتم‌های سنتی مانند DBSCAN و K-means، در توانایی آن برای شناسایی خودکار داده‌های پرت و نویزی نهفته است که این ویژگی موجب افزایش دقت و پایداری نتایج در شرایط متغیر و پیچیده محیطی شده است. افزون بر این، یافته‌ها نشان می‌دهد که الگوریتم HDBSCAN ضمن دستیابی به دقت بالا، از سرعت پردازش مناسبی نیز برخوردار است و قادر است در بازه‌ی زمانی حدود ۲۶۰ تا ۳۵۰ میلی‌ثانیه حمله‌ی فریب را شناسایی نماید؛ موضوعی که امکان استفاده از آن را در سامانه‌های نزدیک به بلادرنگ (Near Real-Time) فراهم می‌سازد.

مقایسهٔ کمی میان الگوریتم‌های موردبررسی نشان می‌دهد که روش پیشنهادی HDBSCAN بادقت تشخیص ۹۸/۴۳ درصد و نرخ هشدار غلط ۵۳/۱ درصد، بهترین عملکرد را در میان روش‌های آزموده‌شده ارائه کرده است. در مقابل، الگوریتم‌های OPTICS و GMM هرچند در برخی سناریوها نتایج قابل‌قبولی به‌دست آورده‌اند، اما به دلیل وابستگی بالا به پارامترهای اولیه و زمان محاسباتی طولانی‌تر، از کارایی کمتری نسبت به الگوریتم پیشنهادی برخوردار بوده‌اند. همچنین، الگوریتم‌های مبتنی بر یادگیری ماشین نظیر شبکه عصبی بازگشتی (RNN) و فیلتر کالمن، اگرچه از نظر سادگی پیاده‌سازی و هزینه‌ی سخت‌افزاری مزایایی دارند، اما دقت شناسایی آن‌ها به ترتیب کمتر از ۷۰ درصد و حدود ۳۳ درصد گزارش شده است که نشان‌دهنده‌ی ضعف نسبی در عملکرد آن‌ها در محیط‌های نویزی و دینامیک است.

در مجموع، نتایج تجربی این پژوهش نشان می‌دهد که الگوریتم پیشنهادی HDBSCAN به دلیل دقت بالا، پایداری در شرایط نویزی، عدم وابستگی به تنظیمات دستی و قابلیت تشخیص خودکار داده‌های پرت، می‌تواند به‌عنوان یک راهکار بهینه و هوشمند برای شناسایی حملات فریب در گیرنده‌های GNSS مورد استفاده قرار گیرد. این الگوریتم نه‌تنها در محیط‌های آزمایشگاهی بلکه در سناریوهای واقعی نیز عملکرد قابل اطمینانی از خود نشان داده است و می‌تواند به‌عنوان پایه‌ای برای توسعهٔ سامانه‌های ضد فریب هوشمند و بلادرنگ در کاربردهای حیاتی نظیر پهپادها، هواپیماها و سامانه‌های خودران به کار گرفته شود.

رابطه (۱۸) نشان‌دهندهٔ ماتریس درهم‌ریختگی برای الگوریتم‌های HDBSCAN است که در آن درایه اول، دوم، سوم و چهارم به ترتیب بیانگر TP، FN، FP و TN هستند.

$$CM_{HDBSCAN} = \begin{pmatrix} ۱۴۵۸۶ & ۲۴۵ \\ ۲۳۲ & ۱۴۹۲۶ \end{pmatrix} \qquad (۱۸)$$

1. Accuracy
2. Silhouette
3. Dunn Index

برای برسی درصد میزان درستی و درصد میزان خطا مبتنی بر الگوریتم HDBSCAN از روابط(۱۹) و (۲۰) زیر استفاده می‌کنیم:

$$Precision_{HDBSCAN} = \frac{TP}{TP+FP} \times ۱۰۰ = ٪ ۹۸/۴۳ \quad (۱۹)$$

$$False\ alarm_{HDBSCAN} = \frac{FP}{TN+FP} \times ۱۰۰ = ٪ ۱/۵۳ \quad (۲۰)$$

که برای الگوریتم HDBSCAN درصد دقت درستی و درصد اعلام هشدار غلط به ترتیب برابر ۹۸/۴۳ و ۱/۵۳ درصد هستند.

جدول ۲- مقایسه کمی الگوریتم پیشنهادی با الگوریتم‌های پیشین تشخیص فریب

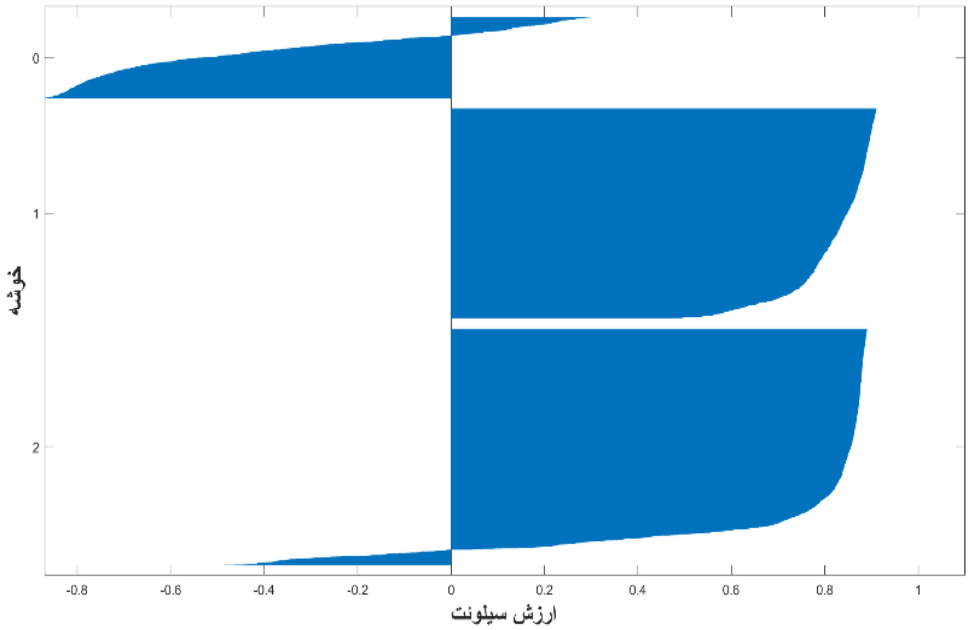
Table 2. Quantitative comparison of the proposed algorithm with existing spoofing detection algorithms

Algoritm	Parameter	
	Accuracy	Dunn
HDBSCAN	0/9841	0/2661
OPTICS	0/9398	0/6009
GMM	0/9903	0/4571
DBSCAN [5]	0/9947	0/8592

جدول ۳- مقایسه الگوریتم پیشنهادی با الگوریتم های مورد بحث از نظر بلادرنگ بودن

Table 3. Comparison of the proposed algorithm with the discussed algorithms in terms of real-time capability

الگوریتم	بلادرنگ بودن (زمان تشخیص)	ویژگی‌ها	مراجع
HDBSCAN	حدود ۲۶۲ میلی‌ثانیه	عملکرد خوب در محیط نویزی، سرعت خوشه‌بندی بالا، مناسب برای داده‌های پیچیده با چگالی متفاوت.	[۱]
K-means	وابسته به تعداد خوشه	سرعت بالا اما نیازمند تعیین تعداد خوشه‌ها، دقت کمتر برای توزیع‌های غیرنرمال.	[۱۵]
DBSCAN	معمولاً بلادرنگ نیست	تشخیص نویز خوب، حساسیت بالا به پارامترهای ورودی.	[۱۲]
GMM	حدود ۳۵۰ میلی‌ثانیه	انعطاف‌پذیری بالا، حساس به تعداد خوشه‌ها و پارامترهای اولیه.	[۲]
OPTICS	کندتر (وابسته به داده‌ها)	عدم نیاز به تعیین تعداد خوشه‌ها، مناسب برای داده‌های ناهمگن و نامتوازن اما با زمان پردازش طولانی‌تر.	[۱۶]
فیلتر کالمن	حدود ۳۳ درصد دقت	بلادرنگ، قابل پیاده‌سازی روی پردازنده‌های ارزان‌قیمت اما دقت کم در شناسایی حمله.	[۱۷]
شبکه عصبی بازگشتی	کمتر از ۷۰ درصد دقت	دقت پایین‌تر، مناسب برای سیستم‌های ساده‌تر.	[۱۸]
FCM	وابسته به داده‌ها	حساسیت بالا به انتخاب پارامترها، دقت بالا در تشخیص الگوهای پیچیده اما کندتر.	[۱۷]



شکل ۱۰- نتیجه شاخص نیم‌رخ برای الگوریتم HDBSCAN، (۰) نشان‌دهنده نویز می‌باشد.

Fig. 10. HDBSCAN algorithm profile index (0: environmental noise)

جدول ٤- مقایسه مهمترین ویژگی‌های روش‌های پیشین و پیشنهادی شناسایی فریب GPS

Table 4. Comparison of the key features between the proposed method and existing GPS spoofing detection methods

روش‌های تشخیص فریب	مزایا	معایب	مقالات با نتایج مشابه
HDBSCAN	سرعت بالا در خوشه‌بندی، عدم نیاز به تعیین پارامترهای اولیه، توانایی بالا در محیط‌های پر نویز، توانایی بالا در داده‌های با چگالی زیاد، پایداری بیشتر نسبت به DBSCAN	پیچیدگی محاسباتی بالا، تفسیر پیچیده‌تر نتایج، عملکرد وابسته به توزیع داده‌ها و تراکم آن‌ها، نتایج ممکن است به دلیل شناسایی نویز برخی نقاط کامل خوشه‌بندی نشوند	[۲۱]
K-means	سادگی، سرعت بالا، عملکرد خوب برای داده‌های خطی.	نیاز به تعیین تعداد خوشه، ناتوان در تشخیص داده‌ها با توزیع غیرنرمال، دقت پایین در داده‌های با شکل غیرمحدب.	[۱۵]
GMM	انعطاف‌پذیری در تعیین خوشه‌ها، قابلیت مدل‌سازی احتمالاتی، قابلیت تطبیق با توزیع غیریکنواخت.	نیاز به تعیین تعداد خوشه، حساسیت به پارامترها.	[۲]
OPTICS	عدم نیاز به تعیین تعداد خوشه، عدم نیاز به تعیین دقیق پارامتر ε، قابلیت کار با داده‌های نامتوازن و ناهمگن.	افزایش زمان محاسبه، افزایش حافظه نگهداری، افزایش هزینه.	[۱۶]
DBSCAN	توانایی تشخیص نویز، عدم نیاز به تعیین تعداد خوشه.	نیاز به تعیین دقیق پارامتر ε و تغییر آن در صورت نیاز، عدم توانایی در تشخیص نقاط مرزی.	[۲۱]
فیلتر کالمن	نیاز به پارامترهای ورودی، پیاده‌سازی روی پردازنده‌های ارزان‌قیمت.	وابستگی به پارامترهای ورودی، دقت شناسایی پایین حدود ۳۳ درصد.	[۱۷]
شبکه عصبی بازگشتی	پیاده‌سازی روی پردازنده‌های ارزان‌قیمت	وابستگی به پارامترهای ورودی، دقت شناسایی کمتر از ۷۰ درصد.	[۱۷]
FCM	انعطاف‌پذیری، قابلیت استفاده در محیط پر نویز، تشخیص الگوهای پیچیده.	حساسیت به انتخاب پارامترها، پایین‌بودن سرعت اجرا، وابستگی به شروط اولیه.	[۲۰]
Subtractive Clustering	سرعت اجرای بالا، قابلیت تطبیق با داده‌های پر نویز.	حساسیت به انتخاب پارامترها، عدم قابلیت تطبیق با خوشه‌بندی پویا.	[۲۱]

۵ پیشنهادات و محدودیت‌ها

باوجود عملکرد مطلوب الگوریتم پیشنهادی در تشخیص سیگنال فریب و دقت بالای به‌دست‌آمده، این پژوهش نیز همانند سایر مطالعات دارای محدودیت‌هایی است. از جمله می‌توان به وابستگی نسبی عملکرد الگوریتم HDBSCAN به نوع و توزیع داده‌ها اشاره کرد. هرچند این روش در مقایسه با الگوریتم‌های مشابه مانند DBSCAN حساسیت کمتری نسبت به پارامترهای اولیه دارد، اما در داده‌هایی با چگالی‌های نامنظم یا ساختارهای پیچیده ممکن است فرایند خوشه‌بندی به طور کامل انجام نشود. علاوه بر این، پیچیدگی محاسباتی بالای الگوریتم به دلیل ماهیت سلسله‌مراتبی و نیاز به محاسبهٔ شاخص‌های چگالی، موجب افزایش زمان پردازش می‌شود

همچنین، ارزیابی‌های انجام‌شده تنها بر روی داده‌های تک‌فرکانسه (L1)و در محیط‌های کنترل‌شده آزمایشگاهی صورت‌گرفته است و بنابراین تعمیم نتایج به شرایط واقعی‌تر از جمله محیط‌های شهری، چندمسیره و دارای تغییرات جوی نیازمند بررسی‌های بیشتر است. از سوی دیگر، تمرکز این پژوهش بر روی فریب‌های شبیه‌سازی‌شده با توان متوسط بوده و انواع فریب‌های پیشرفته‌تر نظیر فریب هماهنگ چند ماهواره‌ای یا حملات دینامیک چند منبعی مورد تحلیل قرار نگرفته‌اند.

در راستای توسعهٔ تحقیقات آینده، پیشنهاد می‌شود الگوریتم پیشنهادی برای گیرنده‌های چندفرکانسه و چند آنتنی توسعه یابد تا امکان تشخیص فریب‌های پیچیده‌تر فراهم شود. همچنین، انجام آنالیز حساسیت سیستماتیک بر اساس پارامترهای مؤثر بر فریب و بررسی پایداری روش در سطوح مختلف نویز و توان سیگنال می‌تواند گامی مؤثر در بهبود دقت روش باشد. ترکیب الگوریتم HDBSCAN با روش‌های یادگیری عمیق یا شبکه‌های عصبی خود رمزگذار (Autoencoder) نیز می‌تواند به بهبود دقت و افزایش توانایی تفکیک الگوهای غیرخطی در داده‌های واقعی کمک نماید.

از سوی دیگر، توسعهٔ نسخه‌ای بهینه‌سازی‌شده برای اجرای بلادرنگ (Real-Time Optimized) به‌منظور پیاده‌سازی در

سامانه‌های قابل‌حمل یا پهپادها می‌تواند مسیر جدیدی از کاربردهای عملی این پژوهش را فراهم کند. در نهایت، پیشنهاد می‌شود در ادامهٔ این مسیر، یک‌لایه تصمیم‌گیری تطبیقی طراحی شود تا بتواند بر اساس شدت حمله یا تغییرات محیطی، پارامترهای الگوریتم را به‌صورت پویا تنظیم کند و پایداری سیستم را در مواجهه با حملات متنوع حفظ نماید.

٦ نتیجه‌گیری

در این پژوهش، الگوریتم خوشه‌بندی خودکار فضایی مبتنی بر چگالی HDBSCAN برای شناسایی هوشمند حملات فریب در گیرنده‌های تک فرکانسه GNSS ارائه و ارزیابی شد. این الگوریتم با تحلیل ویژگی‌های مؤثر سیگنال شامل فاز، نرم هم‌بستگی شاخه‌های ردیابی و توان نسبی سیگنال دریافتی توانست داده‌های مربوط به سیگنال‌های اصلی و جعلی را به‌صورت خودکار خوشه‌بندی و جداسازی کند. استفاده از ساختار سلسله‌مراتبی مبتنی بر چگالی باعث شد تا الگوریتم پیشنهادی نسبت به روش‌های پیشین مانند DBSCAN و OPTICS عملکرد پایدارتری داشته باشد و بدون نیاز به تعیین دقیق پارامترهای اولیه، خوشه‌بندی بهینه‌ای از داده‌ها را در شرایط نویزی و پیچیده انجام دهد. نتایج حاصل از آزمایش‌های میدانی و شبیه‌سازی‌های انجام‌شده بر روی داده‌های واقعی نشان داد که الگوریتم پیشنهادی قادر است بادقت شناسایی ۹۸/۴۳ درصد و نرخ هشدار غلط ۱/۵۳ درصد، حملات فریب را از سیگنال‌های معتبر با موفقیت تشخیص دهد. ارزیابی‌ها بر اساس شاخص‌های Silhouette، Dunn و Confusion Matrix صحت و کارایی بالای خوشه‌بندی را تأیید کردند. همچنین، مقایسه‌ی الگوریتم پیشنهادی با روش‌های دیگر نشان داد که HDBSCAN از نظر سرعت خوشه‌بندی، دقت در تفکیک داده‌ها و مقاومت در برابر نویز عملکرد برتری دارد و می‌تواند در شرایط واقعی با تغییر چگالی داده‌ها، پایداری قابل توجهی از خود نشان دهد. آزمون‌های سخت‌افزاری نیز بیانگر آن بود که الگوریتم در سناریوهای واقعی، در زمانی حدود ۲۶۰ تا ۳۵۰ میلی‌ثانیه حمله را شناسایی کرده و پرچم هشدار را به پردازنده ناوبری اعلام می‌کند؛ امری که قابلیت استفاده از روش پیشنهادی را در سامانه‌های نزدیک به بلادرنگ امکان‌پذیر می‌سازد.

در مجموع، نتایج این تحقیق نشان می‌دهد که الگوریتم HDBSCAN به دلیل پایداری بالا، دقت چشمگیر، و توانایی در شناسایی خودکار داده‌های پرت و نویزی، ابزاری کارآمد برای مقابله با حملات فریب در گیرنده‌های GNSS محسوب می‌شود و می‌تواند به‌عنوان پایه‌ای برای توسعهٔ سامانه‌های ضد فریب هوشمند در کاربردهای حساس نظیر پهپادها، هواپیماها و سامانه‌های خودران مورداستفاده قرار گیرد. بااین‌وجود، برای گسترش قابلیت‌های روش پیشنهادی، پیشنهاد می‌شود در تحقیقات آتی عملکرد آن در شرایط چند فرکانسه و چند ماهواره‌ای موردبررسی قرار گیرد و همچنین با انجام تحلیل حساسیت پارامترها، بهینه‌سازی زمان اجرا و ترکیب الگوریتم HDBSCAN با روش‌های یادگیری عمیق، کارایی آن برای محیط‌های عملیاتی پیچیده‌تر و بلادرنگ ارتقا یابد.

تعارض منافع

هیچگونه تعارض منافع توسط نویسندگان بیان نشده است

مراجع

[1] T. R. Das, S. Hasan, S. Sarwar, J. K. Das, and M. A. Rahman, "Facial spoof detection using support vector machine," in *Proceedings of International Conference on Trends in Computational and Cognitive Engineering: Proceedings of TCCE 2020*, 2021: Springer, pp. 615-625, https://doi.org/10.1007/978-981-33-4673-4_50.

[2] R. A. Sowah, K. B. Ofori-Amanfo, G. A. Mills, and K. M. Koumadi, "Detection and prevention of man-in-the-middle spoofing attacks in MANETs using predictive techniques in Artificial Neural Networks (ANN)," *Journal of Computer Networks and Communications,* vol. 2019, 2019, Art. no. 4683982, https://doi.org/10.1155/2019/4683982.

[3] E. Shafiee, M. R. Mosavi, and M. Moazedi, "Detection of spoofing attack using machine learning based on multi-layer neural network in single-frequency GPS receivers," *The Journal of Navigation,* vol. 71, no. 1, pp. 169-188, 2018, https://doi.org/10.1017/S0373463317000558.

[4] K. Radoš, M. Brkić, and D. Begušić, "Recent advances on jamming and spoofing detection in GNSS," *Sensors,* vol. 24, no. 13, 2024, Art. no. 4210, https://doi.org/10.3390/s24134210.

[5] Z. Sarpanah, M. R. Mosavi, and E. Shafiee, "GPS receivers spoofing detection based on subtractive, FCM and DBScan clustering algorithms," *Journal of Circuits, Systems and Computers,* vol. 32, no. 9, 2023, Art. no. 2350152, https://doi.org/10.1142/S0218126623501529.

[6] T. Talaei Khoei, S. Ismail, and N. Kaabouch, "Dynamic selection techniques for detecting GPS spoofing attacks on UAVs," *Sensors,* vol. 22, no. 2, 2022, Art. no. 662, https://doi.org/10.3390/s22020662.

[7] L. Wang, P. Chen, L. Chen, and J. Mou, "Ship AIS trajectory clustering: An HDBSCAN-based approach," *Journal of Marine Science and Engineering,* vol. 9, no. 6, 2021, Art. no. 566, https://doi.org/10.3390/jmse9060566.

[8] M. de Berg, A. Gunawan, and M. Roeloffzen, "Faster DB-scan and HDB-scan in low-dimensional Euclidean spaces," *arXiv preprint arXiv:1702.08607,* 2017, https://doi.org/10.48550/arXiv.1702.08607.

[9] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu, "DBSCAN revisited, revisited: why and how you should (still) use DBSCAN," *ACM Transactions on Database Systems (TODS),* vol. 42, no. 3, pp. 1-21, 2017, v Art. no.19, https://doi.org/10.1145/3068335.

[10] T. V. Sai Krishna, A. Yesu Babu, and R. Kiran Kumar, "Determination of optimal clusters for a Non-hierarchical clustering paradigm K-Means algorithm," in *Proceedings of International Conference on Computational Intelligence and Data Engineering,* Singapore: Springer, 2018, pp. 301-316, https://doi.org/10.1007/978-981-10-6319-0_26.

[11] J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," *Journal of Cybernetics*, vol. 3, no. 3, pp. 32-57, 1973, https://doi.org/10.1080/01969727308546046.

[12] Z. Huang, "Extensions to the k-means algorithm for clustering large data sets with categorical values," *Data Mining and Knowledge Discovery,* vol. 2, no. 3, pp. 283-304, 1998, https://doi.org/10.1023/A:1009769707641.

[13] H. Qiu, Y. Liu, N. A. Subrahmanya, and W. Li, "Granger causality for time-series anomaly detection," in *2012 IEEE 12th International Conference on Data Mining*, 2012: IEEE, pp. 1074-1079, https://doi.org/10.1109/ICDM.2012.73.

[14] M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J. Sander, "OPTICS: Ordering points to identify the clustering structure," *ACM Sigmod Record,* vol. 28, no. 2, pp. 49-60, 1999, https://doi.org/10.1145/304181.304187.

[15] M. R. Mosavi, M. J. Rezaei, N. Hosseinzadeh, R. A. Kiaamiri, "Two New Intelligent Methods for Detecting and Cancelling Spoofing Effects on GPS Receivers", *Journal of Electronical & Cyber Defence*, vol. 2, no. 1, pp.71-81, 2014, (In persian).

[16] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Computers & Geosciences,* vol. 10, no. 2-3, pp. 191-203, 1984, https://doi.org/10.1016/0098-3004(84)90020-7.

[17] G. Stewart and M. Al-Khassaweneh, "An implementation of the HDBSCAN* clustering algorithm," *Applied Sciences,* vol. 12, no. 5, 2022, Art. no. 2405, https://doi.org/10.3390/app12052405.

[18] K. Khan, S. U. Rehman, K. Aziz, S. Fong, and S. Sarasvady, "DBSCAN: Past, present and future," in *The fifth international conference on the applications of digital information and web technologies (ICADIWT 2014)*, 2014: IEEE, pp. 232-238, https://doi.org/10.1109/ICADIWT.2014.6814687.

[19] N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image segmentation using K-means clustering algorithm and subtractive clustering algorithm," *Procedia Computer Science,* vol. 54, pp. 764-771, 2015, https://doi.org/10.1016/j.procs.2015.06.090.

[20] E. Shafiee, S. M. R. Mosavi, and M. Moazedi, "Detection of Spoofing Attack Based on Multi-Layer Neural Network in Single-Frequency GPS Receivers," *Journal of Electronic and Cyber Defence,* vol. 3, no. 1, pp. 69-80, 2015.

[21] E. Shafiee, M. R. Mosavi, and M. Moazedi, "A modified imperialist competitive algorithm for spoofing attack detection in single-frequency GPS receivers." *Wireless Personal Communications,* vol. 119, pp. 919-940, 2021, https://doi.org/10.1007/s11277-021-08244-2.