

INSTRUCTOR: uma Ferramenta para Análise de Trajetórias Anômalas de Navios por Meio de Algoritmos de Agrupamento*

Claudio V. Ribeiro¹, Aline Paes¹ e Daniel de Oliveira¹

¹Instituto de Computação - Universidade Federal Fluminense
Av. Gal. Milton Tavares de Souza, s/nº, Boa Viagem – Niterói/RJ – Brasil

claudiovr@id.uff.br {alinepaes, danielcmo}@ic.uff.br

Resumo. Milhares de navios circulam diariamente e a quantidade de incidentes é expressiva, sendo tais ocorrências associadas à trajetórias consideradas anômalas. Os agentes responsáveis pela vigilância marítima precisam ser apoiados por análises visuais da situação para indicação dessas ocorrências, sobretudo de forma antecipada. Neste artigo, apresentamos a ferramenta INSTRUCTOR, que permite a análise visual de trajetórias anômalas de navios usando múltiplos algoritmos de agrupamento (ex. DBSCAN, K-Means, Birch, Spectral Clustering e Ensembles). A INSTRUCTOR pode ser obtida no repositório do GitHub¹ e vem sendo avaliada junto a especialistas da Marinha do Brasil por meio de questionários específicos com roteiros de testes.

Abstract. Thousands of vessels circulate daily and the number of incidents is significant, and such occurrences are associated with trajectories considered anomalous. The agents responsible for maritime surveillance need to be supported by visual analysis of the situation to indicate these occurrences, especially in advance. In this article, we present the INSTRUCTOR tool, which allows the visual analysis of anomalous vessel trajectories using multiple clustering algorithms (e.g. DBSCAN, K-Means, Birch, Spectral Clustering and Ensembles). The INSTRUCTOR can be obtained at GitHub repository and has been evaluated by experts from the Brazilian Navy through specific questionnaires with a test script.

1. Introdução

O transporte marítimo desempenha um papel essencial na economia global, com 90% do comércio mundial realizado por via marítima [UNCTAD 2022], o que demonstra a importância do transporte marítimo tanto em volume quanto em valor financeiro [Ribeiro et al. 2023]. Somente em 2019, o número de navios comerciais chegou a mais de 52.000 [UNCTAD 2019] e foram relatados 3.174 incidentes marítimos, sendo 95 deles graves, com 53 mortes e mais de 900 feridos [EMSA 2019]. Embora o número de incidentes tenha diminuído nos últimos anos, ainda é um desafio reduzi-los [Weng et al. 2018]. Existem diversos tipos de incidentes marítimos, sendo os tipos mais comuns: (i) colisões, (ii) naufrágios e (iii) encalhes [Ribeiro et al. 2023]. Outros tipos de eventos também podem impactar a atividade marítima, incluindo, pirataria, terrorismo e poluição [Zor and Kittler 2017]. Uma característica em comum quando tais eventos ocorrem é que o

*Vídeo da demonstração da ferramenta disponível em <https://www.youtube.com/watch?v=q6Vn7NyFAe8>

¹<https://github.com/UFFeScience/instructor>

navio apresenta uma trajetória anômala se comparada com outras rotas regulares na mesma região. Assim, a identificação de trajetórias anômalas se torna fundamental para tomar providências que evitem a ocorrência de tais incidentes.

Entretanto, se considerarmos a quantidade de navios em determinadas regiões do planeta, a detecção de trajetórias anômalas se mostra como uma tarefa não trivial e propensa a erros, se executada manualmente. Assim, abordagens automatizadas ou semi-automatizadas devem ser aplicadas nesse cenário. A detecção automática ou semi-automática de trajetórias anômalas de navios pode ajudar os operadores de sistemas de vigilância marítima em cenários críticos com restrições de tempo e um grande número de navios [Kowalska and Peel 2012]. A detecção pode ser realizada a partir de dados já existentes, como os dados do AIS (Sistema de Identificação Automática) [Sidibé and Shu 2017]. O AIS é baseado no uso de *transponders* em navios para monitorar o tráfego e representa a principal fonte de informações para a vigilância do tráfego marítimo [De Vries and Van Someren 2012]. *Transponders* são equipamentos que recebem automaticamente informações transmitidas por outras embarcações e estações fixas equipadas com AIS ao mesmo tempo em que enviam informações para outras embarcações. Assim, existe um grande volume de dados gerados por esses sistemas e que podem ser explorados de maneiras que não eram possíveis há poucos anos.

Existem diversas abordagens que propõem o uso de técnicas de Aprendizado de Máquina para auxiliar na identificação de trajetórias anômalas, em especial utilizando algoritmos de agrupamento [Ribeiro et al. 2023] como o DBSCAN [Ester et al. 1996] e suas variantes [Dobrkovic et al. 2015, Zhao and Shi 2019, Wang et al. 2014]. Por exemplo, [Zhao and Shi 2019, Dobrkovic et al. 2015] utilizam o algoritmo DBSCAN tradicional com dados de posicionamento (latitude e longitude) da embarcação obtidos do AIS para identificação das trajetórias anômalas, enquanto que [Wang et al. 2014] propõem uma variante que incorpora outros dados como o COG (*Course over Ground*) e o SOG (*Speed Over Ground*), também disponíveis no AIS, para realizar o agrupamento. Apesar de tais trabalhos representarem um avanço, os operadores de vigilância marítima necessitam realizar também uma análise visual dos agrupamentos de rotas e das trajetórias anômalas para chegar a uma decisão. A comparação visual de uma trajetória sob análise com relação aos agrupamentos de rotas gerados por múltiplos algoritmos de agrupamento propostos ainda não é fornecida por nenhuma ferramenta atualmente.

De forma a suprir essa lacuna, o presente artigo apresenta a ferramenta INSTRUCTOR (do inglês *IdeNtification of anomalous veSsel TRajectory Using ClusTers Of tRajectories*). A ideia da INSTRUCTOR é permitir ao operador de tráfego marítimo avaliar visualmente os agrupamentos de rotas gerados a partir de dados AIS históricos e comparar trajetórias atuais com tais agrupamentos gerados por múltiplas abordagens. A partir dessa comparação visual e de informações complementares apresentadas pela ferramenta (*e.g.*, velocidade da embarcação), a ideia é que o operador seja capaz de identificar e classificar a trajetória atual como anômala (ou não). A ferramenta INSTRUCTOR vem sendo avaliada em conjunto com especialistas da Marinha do Brasil. Esse artigo de demonstração está organizado em três seções além da Introdução. A Seção 2 apresenta a ferramenta INSTRUCTOR. A Seção 3 discute a demonstração, e, finalmente, a Seção 4 conclui o artigo.

2. A Ferramenta INSTRUCTOR

A INSTRUCTOR é uma ferramenta *Web* que oferece uma interface onde usuários operadores de vigilância marítima carregam os dados de AIS de diferentes fontes. A partir de tal carga, os dados são processados de forma automática para a geração de agrupamentos de rotas utilizando múltiplos algoritmos de agrupamento. Além dos dados do AIS, a INSTRUCTOR importa dados de outras fontes, para incluir dados adicionais, como a velocidade do vento e dados ambientais. Tais informações complementares são fundamentais para a identificação de trajetórias anômalas. A INSTRUCTOR permite a visualização e interação com o mapa da região em questão em diversos níveis de detalhes. A arquitetura da INSTRUCTOR é apresentada na Figura 1

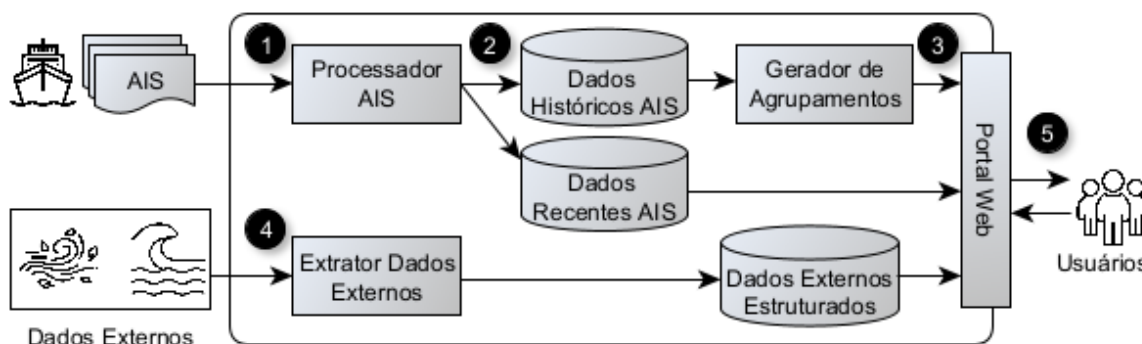


Figura 1. Arquitetura da INSTRUCTOR

A arquitetura da INSTRUCTOR é composta de seis componentes: (i) Processador AIS, (ii) Base de Dados AIS, (iii) Gerador de Agrupamentos, (iv) Extrator de Dados Externos, (v) Base de Dados Externos Estruturados e (vi) Portal *Web*. A INSTRUCTOR recebe como entrada (passo ①) um conjunto de arquivos contendo dados do AIS (*i.e.*, múltiplas posições do navio ao longo do tempo) de diversos navios. Esses dados são processados pelo Processador AIS, que os insere em um banco de dados com as rotas históricas importadas (passo ②). Esse banco de dados contém o *timestamp* da leitura do AIS, latitude, longitude, velocidade e rumo, *etc.* Esse banco de dados histórico é acessado pelo *Gerador de Agrupamentos* que realiza o agrupamento (passo ③) de acordo com um catálogo de algoritmos de agrupamento disponíveis e usualmente empregados (*e.g.*, DBSCAN, *K-Means*, *Spectral Clustering*, *Birch* e *Agglomerative Clustering*), bem como em variações de uso conjunto na forma de *Ensembles*. Além dos dados do AIS, a INSTRUCTOR importa dados de fontes complementares, como dados de vento e ondas. O *Extrator de Dados Externos* possui um catálogo de fontes externas e periodicamente importa esses dados para a *Base de Dados Externos Estruturados* (passo ④). Finalmente, os usuários acessam a INSTRUCTOR via *portal Web* (passo ⑤) que apresenta os agrupamentos de rotas e dados complementares de forma visual para avaliação do usuário.

3. Demonstração

A demonstração da INSTRUCTOR seguirá um estudo de caso com dados de rotas de navios na Baía de Delaware², na costa nordeste dos Estados Unidos (EUA), que é um estuário na foz do rio Delaware. Por conta da geografia e por ser um local onde a água doce e a água

²<https://goo.gl/maps/iGqbpBEUthpZLUPB9>

salgada do Atlântico se misturam numa vasta extensão, a baía é um local de criação de ostras. Além disso, serve como rota de escoamento da produção da Filadélfia. Por conta dessas características, o local possui um fluxo intenso de navios todos os dias. Os dados do AIS e os dados ambientais (*e.g.*, vento, ondas, *etc*) foram obtidos do NOAA (*National Oceanic and Atmospheric Administration*)³.

A demonstração se inicia com a definição da área de interesse no mapa da baía de Delaware. O usuário define os pontos de interesse para formação de um polígono (Figura 2(a)) e uma grade é criada sobre essa área (Figura 2(b)). O usuário define a granularidade desejada nesse momento, podendo definir quadrículas de 100, 200, 500, 1000, 2000 e 5000 metros de lado para compor a grade. A grade pode ser salva para ser reutilizada posteriormente. A partir da grade definida, os algoritmos de agrupamento podem ser executados. A Figura 3 apresenta a interface do portal *Web* da INSTRUCTOR e possui quatro regiões principais: Mapa (área ③), Dados Históricos (áreas ① e ②), Análise de Nova Trajetória (áreas ④ e ⑤) e Dados Complementares (área ⑥).

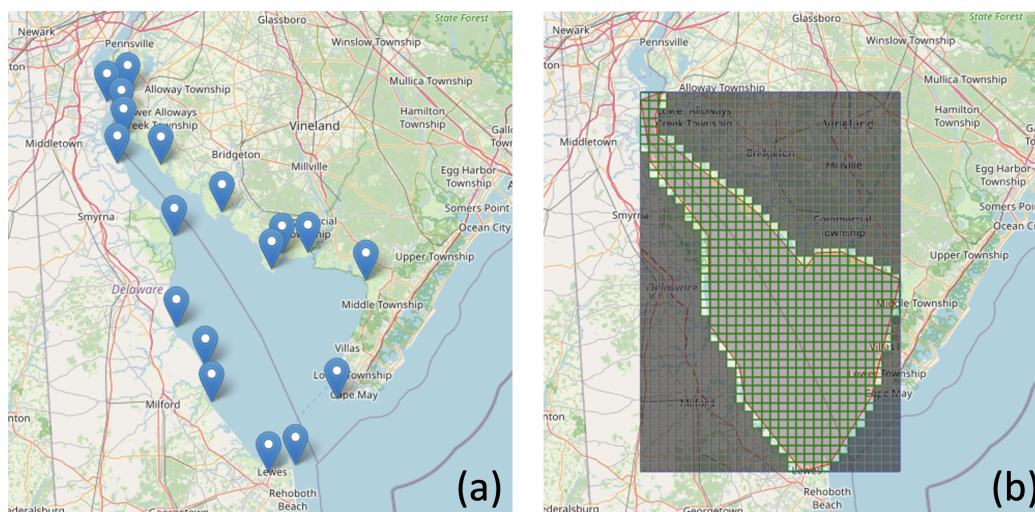


Figura 2. Definição da área de interesse e da grade sobre o mapa.

Após definir a grade, o usuário inicia o carregamento dos dados históricos do AIS. O usuário seleciona o arquivo e o carrega na INSTRUCTOR (área ①). Com os dados já carregados, o usuário seleciona qual algoritmo de agrupamento deseja aplicar sobre os dados (área ②). O usuário pode ainda definir outros parâmetros como o número de agrupamentos (necessário dependendo do algoritmo escolhido). Uma vez que os agrupamentos de rotas tenham sido gerados, eles são apresentados para o usuário por meio de cores distintas (área ③), sendo a cor preta reservada para indicar pontos que não foram vinculados a qualquer agrupamento. O usuário então pode selecionar um arquivo com dados de AIS de embarcações em trânsito (área ④) e comparar a trajetória das mesmas com os agrupamentos de rotas gerados, sendo apresentado o percentual de similaridade entre a trajetória em análise e cada agrupamento (área ③). Se um mesmo arquivo possuir dados do AIS de múltiplas embarcações, o usuário pode escolher qual delas deseja comparar com os agrupamentos de rotas. Finalmente, o usuário pode verificar os dados de rumo e velocidade da embarcação ponto a ponto (área ⑤) da trajetória e comparar com dados históricos exibidos

³<https://www.noaa.gov/>

em uma rosa dos ventos (área 6) para complementar as análises.

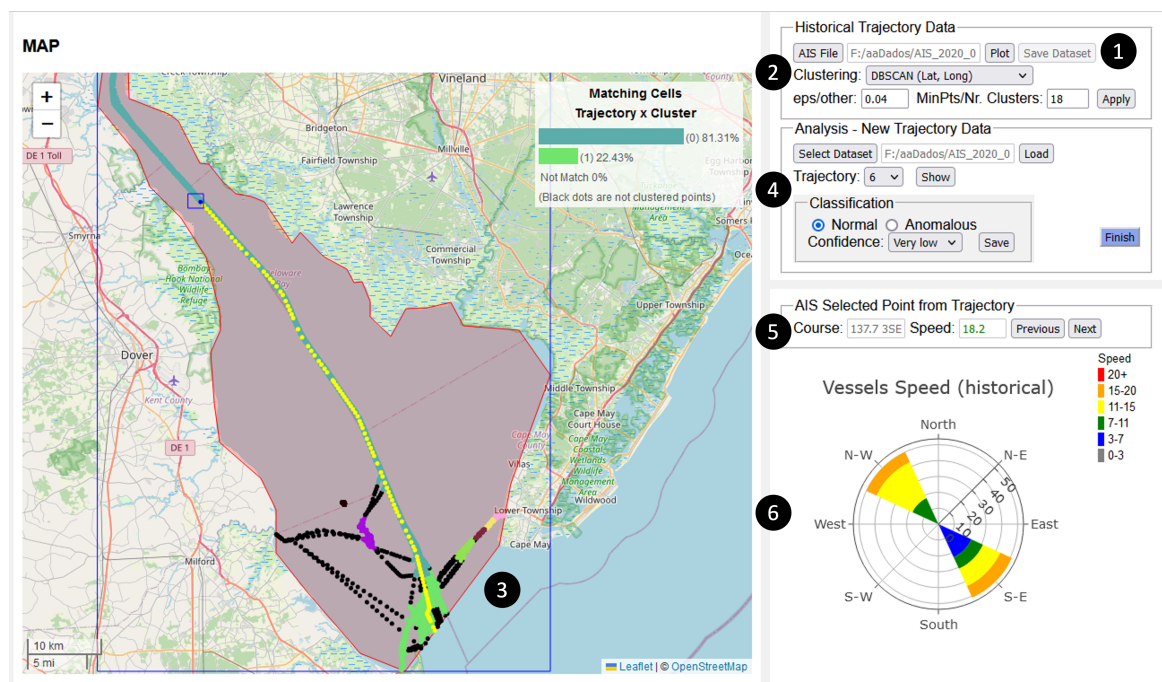


Figura 3. Interface do Portal Web da INSTRUCTOR

4. Conclusões

A análise e identificação de trajetórias anômalas de navios desempenham um papel crucial para os usuários envolvidos no monitoramento marítimo. Devido ao grande volume de navios, realizar essa análise manualmente se torna tedioso e propenso a erros. Existem soluções que já propõem o uso de algoritmos de agrupamento para identificar trajetórias anômalas por meio da comparação com agrupamentos de rotas regulares de navios. Cada solução possui pequenas diferenças (e.g., algumas consideram apenas dados de posicionamento, enquanto outras incluem também a velocidade da embarcação). Comumente, os usuários precisam comparar agrupamentos de rotas gerados por diferentes soluções para chegar a uma conclusão. Neste artigo de demonstração, apresentamos a INSTRUCTOR, que permite analisar agrupamentos de rotas gerados por múltiplos algoritmos e soluções, a fim de apoiar o processo de decisão dos especialistas. Além disso, o INSTRUCTOR fornece dados complementares, como a velocidade do vento e da embarcação, para auxiliar na tomada de decisão. No momento, a ferramenta está sendo avaliada por especialistas da Marinha do Brasil. Estes avaliam se os agrupamentos realizados pelos algoritmos estão condizentes com os padrões de navegação em uma determinada área, fazendo a análise com trajetórias que não compuseram tais agrupamentos para a geração de um *ground truth* de trajetórias anômalas, viabilizando, assim, a análise das soluções dos algoritmos pelo pesquisador. Trabalhos futuros incluem a importação de diferentes tipos de dados, incluindo imagens de satélite com previsões meteorológicas em tempo real e informações textuais análogas extraídas de Tweets.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Os autores gostariam

ainda de agradecer ao CNPq (*grant* 311898/2021-1) e FAPERJ (*grant* E-26/202.806/2019) pelo apoio financeiro.

Referências

- De Vries, G. K. D. and Van Someren, M. (2012). Machine learning for vessel trajectories using compression, alignments and domain knowledge. *Expert Systems with Applications*, 39(18):13426–13439.
- Dobrkovic, A., Iacob, M.-E., and van Hillegersberg, J. (2015). Using machine learning for unsupervised maritime waypoint discovery from streaming ais data. *Proc. of the i-KNOW '15*, pages 1–8.
- EMSA (2019). Annual overview of marine casualties and incidents. Technical report.
- Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *Proceedings of the KDD'96*, page 226–231. AAAI Press.
- Kowalska, K. and Peel, L. (2012). Maritime anomaly detection using gaussian process active learning. *2012 15th International Conference on Information Fusion*, pages 1164–1171. IEEE.
- Ribeiro, C. V., Paes, A., and de Oliveira, D. (2023). Ais-based maritime anomaly traffic detection: A review. *Expert Systems with Applications*, page 120561.
- Sidibé, A. and Shu, G. (2017). Study of automatic anomalous behaviour detection techniques for maritime vessels. *The journal of Navigation*, 70(4):847–858.
- UNCTAD (2019). *Handbook of Statistics*. United Nations Conference on Trade and Development. Available at https://unctad.org/en/PublicationsLibrary/tdstat44_en.pdf.
- UNCTAD (2022). *Handbook of Statistics*. United Nations Conference on Trade and Development. Available at https://unctad.org/system/files/official-document/tdstat47_en.pdf.
- Wang, X., Liu, X., Liu, B., de Souza, E. N., and Matwin, S. (2014). Vessel route anomaly detection with hadoop mapreduce. *2014 IEEE international conference on big data (big data)*, pages 25–30. IEEE.
- Weng, J., Yang, D., Qian, T., and Huang, Z. (2018). Combining zero-inflated negative binomial regression with mlrt techniques: an approach to evaluating shipping accident casualties. *Ocean Engineering*, 166:135–144.
- Zhao, L. and Shi, G. (2019). Maritime anomaly detection using density-based clustering and recurrent neural network. *The Journal of Navigation*, 72(4):894–916.
- Zor, C. and Kittler, J. (2017). Maritime anomaly detection in ferry tracks. *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2647–2651. IEEE.