

Algoritmos de detección de anomalías y sus aplicaciones en el ámbito marítimo

Autor: Lasso Mula, Alberto

Director: Fernández García, Norberto.

Contacto: alberto_lasso@hotmail.com

Resumen: En los últimos años una de las principales tecnologías que se ha desarrollado exponencialmente posibilitando la transformación digital del mundo es la Inteligencia Artificial, la cual tiene una de sus principales aplicaciones en la detección de anomalías.

La detección de anomalías presenta la utilidad de alertar de comportamientos que se salen de lo normal y que pueden presentar un problema, en el caso concreto del ámbito marítimo se pretende con ello detectar los comportamientos ilícitos y las situaciones de riesgo de las embarcaciones.

Este trabajo hace un extenso análisis del estado del arte de la detección de anomalías en el tráfico marítimo, comparando y clasificando un importante número de estudios relevantes.

Se valoran todos los aspectos concernientes a un sistema de este tipo:

- Las fuentes de información que se pueden utilizar como entrada al algoritmo, las que se analizan detalladamente y de las que el sistema de comunicaciones AIS es la más relevante.
- Se estudian muchas de las diferentes metodologías que se pueden aplicar en el algoritmo de detección de anomalías.
- La presentación de los resultados, es importante facilitar la interpretación de las alertas que se obtengan.

A partir de todos ellos se señalan las principales características y posibles dificultades a los que se tendrán que enfrentar los diseñadores de los sistemas de detección de anomalías en el ámbito marítimo. Con el fin de que las conozcan desde el principio, capacitándoles para tomar todas las decisiones que sean conveniente en el diseño de una buena solución a este problema.

Palabras clave: Detección de anomalías, rutas marítimas, navegación, predicción de trayectorias, Inteligencia Artificial

1. Introducción

1.1. Contexto

El transporte marítimo a lo largo de historia ha sido un motor de progreso, desarrollo y prosperidad. En nuestros días sigue siendo fundamental para el avance de la economía, habiendo tomado más relevancia aún ante el proceso de globalización que se ha producido en las últimas décadas a nivel mundial.

La importancia para la humanidad de tráfico marítimo es evidente y con el fin de que se desarrolle de la manera más eficiente y segura, en 2002 se desarrolla un sistema de identificación automática de embarcaciones denominado AIS¹ de obligada implantación en numerosas embarcaciones.

La expansión del sistema AIS, ha puesto a disposición de las autoridades competentes una fuente más de información para supervisar el tránsito marítimo, prever las situaciones de riesgo y combatir las actividades ilegales que se desarrollan en los mares.

1.2. Objetivo

El principal objetivo del trabajo es estudiar la viabilidad de aplicar sistemas automatizados basados en inteligencia artificial para la detección de anomalías en el tráfico marítimo que permitan identificar las actividades ilegales que se llevan a cabo.

Pretende aglutinar todos principales factores que sean de especial aplicación en este ámbito concreto, los cuales no deberían pasarse por alto en el diseño de los futuros sistemas de detección de anomalías, salvo que así se decida por motivos justificados.

2. Resultados y discusión

El diseño de un sistema que detecta anomalías en el tráfico marítimo es un proceso complejo en el que hay que tener en cuenta múltiples y muy diferentes opciones. Por ello se ha llevado a cabo un amplio estado del arte en que se han incluido los principales y más relevantes trabajos sobre este ámbito.

Estos trabajos se han clasificado en dos grandes grupos, muy diferenciados entre sí, de acercamientos a la hora de definir la solución con que resolver la detección de anomalías en el tráfico marítimo.

El primero de ellos se centra en detectar patrones definidos por expertos en navegación marítima que denotan comportamientos ilícitos o situaciones de riesgo, mientras que el segundo en vez de centrarse en casos concretos, busca los comportamientos que se alejan de los frecuentes y más habituales que son los que se considerarían anómalos.

En la Figura 1 se muestran todos los estudios incluidos en el estado del arte clasificados en estos dos grandes grupos, comparando sus principales características.

¹ AIS: Automatic Identification System.

TRABAJO		REFERENCIA	CLASIFICACIÓN	SUPERVISADO	ENTRENAMIENTO	PARAMETRICO	DIMENSIONES	METODOLOGIA	TIEMPO REAL
TÍTULO									
Vessel Trajectories Outliers	3	Detección de Eventos	No supervisado	SI	SI	Varias	Rules	SI	
Loitering with intent—Catching the outlier vessels at sea	4		Supervisado	SI	SI	Varias	Probability Model	SI	
Extracting rules from expert operators to support situation awareness in maritime surveillance	21		Supervisado	SI	NO	Varias	Queries/Rules	SI	
Maritime anomaly detection and threat assessment	12		Supervisado	SI	NO	Varias	Probability Model	NO	
Crisis: Integrating ais and ocean data streams using semantic web standards for event detection	8		Supervisado	SI	NO	Varias	Queries/Rules	SI	
Abstracting and reasoning over ship trajectories and web data with the Simple Event Model (SEM)	7		Supervisado	SI	NO	Varias	Queries/Rules	SI	
Event Recognition for Maritime Surveillance	10		Supervisado	SI	NO	Varias	Queries/Rules	SI	
Mining maritime traffic conflict trajectories from a massive AIS data	23		No supervisado	NO	SI	Varias	Clustering	NO	
Online Event Recognition from Moving Vessel Trajectories	11		Supervisado	SI	NO	Varias	Queries/Rules	SI	
Semantic modelling of ship behavior in harbor based on ontology and dynamic bayesian network	13		Supervisado	SI	NO	Varias	Probability Model	NO	
Open data for anomaly detection in maritime surveillance	9		Supervisado	SI	NO	Varias	Queries/Rules	SI	
A complex event processing approach to detect abnormal behaviours in the marine environment	5		Supervisado	SI	NO	Varias	Rules	SI	
Detecting search and rescue missions from ais data	6		No supervisado	NO	NO	Varias	Rules	SI	
A multi-task deep learning architecture for maritime surveillance using ais data streams	19	Detección de Anomalías	No Supervisado	SI	NO	Varias	Neural Network	SI	
Valor atipico por cuartiles - Four Techniques for Outlier Detection	2		No Supervisado	No	SI	Una	Statistical	NO APLICA	
Z-Score - Four Techniques for Outlier Detection	2		No Supervisado	SI	SI	Varias	Gaussian Model	NO APLICA	
DBSCAN - Four Techniques for Outlier Detection	2		No Supervisado	SI	SI	Varias	Clustering	NO APLICA	
Isolation forrest - Four Techniques for Outlier Detection	2		No Supervisado	NO	SI	Varias	Rules	NO APLICA	
Seafaring TRANsport ANomaly Detection, STRAND	1		No Supervisado	SI	NO	Varias	Rules	SI	
A network abstraction of multi-vessel trajectory data for detecting anomalies	22		No Supervisado	SI	SI	Varias	Clustering	SI	
Statistical analysis of motion patterns in data: Anomaly detection and motion prediction	14		No Supervisado	SI	NO	Varias	KDE	SI	
Maritime anomaly detection using gaussian process active learning	15		Supervisado	SI	NO	Varias	Gaussian Model	NO	
Improving maritime anomaly detection and situation awareness through interactive visualization	16		Supervisado	SI	NO	Varias	Gaussian Mixture Model (GMM)	NO	
Anomaly detection in maritime data based on geometrical analysis of trajectories	17		No Supervisado	NO	SI	Varias	Rules	SI	
Associative learning of vessel motion patterns for maritime situation awareness	18		No Supervisado	SI	NO	Varias	Neural Network	SI	
Vessel pattern knowledge discovery from ais data: A framework for anomaly detection and route prediction	20		No Supervisado	NO	NO	Varias	Clustering + Probability Model	SI	

Figura 1. Tabla comparativa de los diferentes trabajos incluidos en el estado del arte.

De todos ellos se extraen numerosos aspectos fundamentales que no deberían pasarse por alto en el desarrollo de un sistema de detección de anomalías en el tráfico marítimo

2.1. Fuentes de información

El primer aspecto a analizar es la cantidad y calidad de fuentes de información disponibles, siendo algunas de las más representativas; señalización AIS, información meteorológica, información

portuaria, cartografía e información geográfica, información de rutas marítimas, imágenes de drones, satélite, información aduanera.

Antes de seleccionar las fuentes a utilizar hay que tener en cuenta que cada una de las que se incluya en el sistema aumenta las dimensiones a tratar y con ello, la complejidad del mismo.

2.1.1. Sistema AIS

Representa una base muy sólida, por no decir fundamental para analizar el comportamiento de las embarcaciones gracias a que facilita registros periódicos con su geolocalización, además de más parámetros cinemáticos y de otra naturaleza.

Pero este sistema de comunicaciones, ni es infalible, ni toda la información que aporta se puede considerar igual de válida, pudiendo llegar parte de ella generar confusión. Algunos de los campos remitidos en la señalización AIS deben ser actualizados a mano por los tripulantes, como es el caso del “status”, el cual no es fiable, un número muy elevado de tripulaciones no lo actualizan durante la travesía.

2.1.2. Información meteorológica

La información meteorológica es otra de las fuentes de información que puede tener más sentido tener en cuenta ya que puede ejercer una importante influencia en la navegación de las embarcaciones.

2.1.3. Imágenes

El importante desarrollo que tiene la IA en lo correspondiente al tratamiento de imágenes en otros ámbitos sería fácilmente extrapolable al tránsito marítimo en cuanto exista una forma de obtener imágenes de manera económica y fluida.

2.2. Transformación, preprocesado y filtrado de la información

Una vez se hayan seleccionado las fuentes de información para el sistema de detección de anomalías hay que analizar el formato de los datos, porque es bastante posible que requieran de un procesado o transformación previa a su utilización.

2.3. Sistemas basados en la detección de eventos y patrones de comportamiento

De los dos grandes bloques de metodologías comentados al inicio del capítulo, este es el primero de ellos, el que se basa en el reconocimiento de patrones predefinidos por expertos como sospechosos de delatar comportamientos ilícitos o ilegales. Los principales patrones que incluyen los trabajos revisados son:

- Pérdida señal AIS: Pérdida intencionada o no intencionada de la señalización AIS de una embarcación.
- Detención de un barco un tiempo largo de tiempo.
- Si una embarcación hace un giro repentino sin justificación.
- Dos barcos se detienen en alta mar muy próximos entre sí.
- Si un barco tiene un encuentro con un barco de menor tamaño.
- Dos barcos que van en paralelo (o navegan muy próximos) al menos una cierta distancia y/o tiempo.

- Más aún si un barco que navega próximo a otro durante un tiempo y justo después de estar lo más próximos entre sí uno cambia de dirección significativamente.
- Dos buques que van uno hacia el otro y tras el encuentro uno o los dos realiza un giro.
- Si una embarcación navega a velocidades muy elevadas.
- Una embarcación que transite a una velocidad muy reducida.
- Una embarcación que abandone una ruta habitual para adentrarse a una zona en la que habitualmente no hay tráfico.
- Si una embarcación que habitualmente navega entre dos puertos, de repente va a un tercero al que no iba antes.
- Si un barco no ha dejado registros AIS abandonando o entrando en puerto pero sí figura en los registros portuarios correspondientes.
- Si un barco que figura en los registros de un puerto como que está atracado en él, deja registros AIS en otra localización.
- Si el destino de una embarcación indicado en el sistema AIS, no está registrado en la agenda del puerto.
- Si un barco entra en puerto sin haberle informado previamente.
- Si en vez de entrar a puerto un barco en un horario previsto, entra otro justo en ese intervalo.
- Si una embarcación no ha dejado señalización dentro de puerto, pero sí figura que haya solicitado un práctico.
- Si el tiempo estimado de llegada registrado en las comunicaciones AIS, no coincide con el real de llegada a puerto.

Bastantes de las metodologías de este tipo, detección de eventos, optan por dividir la región marítima que van a tratar en múltiples celdas con objeto de subdividir las trayectorias de las embarcaciones en trozos más cortos sobre los que buscar anomalías con mayor facilidad o bien para aplicar patrones concretos característicos de cada zona concreta.

Queda demostrado que las dimensiones concretas que se elijan para estas celdas afecta directamente a la capacidad de detección de los algoritmos.

2.4. Sistemas basados en la detección de comportamientos anómalos

El segundo grupo de trabajos incluidos en el estado del arte es el que tiene que ver con detectar las embarcaciones que se salen del comportamiento normal, las que tienen algo diferenciador respecto a la mayoría en su comportamiento. Se da por sentado que el gran grueso de las embarcaciones no está inmerso en actividades ilegales y se pretende detectar las que realizan esas actividades justo por las diferencias que tendrá su navegación frente a las que llevan a cabo las tareas habituales y cotidianas.

Los trabajos recopilados en el estado del arte abordan la detección de comportamientos anómalos con una variedad de técnicas diferentes, distribuciones de Gauss, *Kernel Density Estimation*, redes neuronales, agrupamiento, etc.

Siendo la carencia de datos etiquetados una de las principales dificultades a las que se enfrentan y que intentan solventar de diferente forma.

- Retroalimentando el sistema a través de visualizaciones interactivas que validen las detecciones de forma que el sistema se vaya refinando constantemente.
- Que un grupo de expertos etiquete la información manualmente, lo que resultaría extraordinariamente lento y caro por la dificultad que presenta.
- Mediante aplicaciones en Internet que permiten a los propios tripulantes subir las navegaciones que efectúen y etiquetarlas, de forma que se genere una base de datos de información etiquetada con la que trabajar.
- A través de aprendizaje por transferencia, esto es, a partir de unos pocos ejemplos etiquetados, generar y etiquetar otros.

Otra de las singularidades que hay que tener en cuenta es que cierto tipo de embarcaciones se desenvuelven en el mar de diferente manera, por ello algunos estudios se centran en algún tipo de barco concreto o aplican el mismo modelo de forma separada según la naturaleza de éstos.

Son de especial interés los trabajos que utilizan redes neuronales para predecir las trayectorias futuras de las embarcaciones además de detectar anomalías. A partir de modelar los comportamientos habituales de los barcos con estos sistemas se puede detectar cuando alguno se aleja de ellos incluso predecir su trayectoria futura.

Una limitación de la gran mayoría de trabajos de este ámbito es que analizan principalmente el comportamiento individual de las embarcaciones aunque algunas anomalías solo pueden detectarse al analizar un grupo de embarcaciones. Comprender los patrones de movimiento colectivo de los barcos puede ser una tarea más fácil que analizar los comportamientos individuales, tarea que no se está explotando en exceso.

Por ultimo pero no menos importante, estos sistemas de detección de anomalías se enfrentan también a la baja interpretabilidad de los resultados que obtienen. Debe de procurarse añadir a los sistemas capacidades de visualización y procesamiento de datos relevantes involucrados en las alarmas que detecten, para facilitar a los agentes marítimos la interpretación de las mismas. En estos casos sistemas de retroalimentación, como los que se han comentado antes, podrían ser de utilidad.

3. Conclusiones

La detección de anomalías en el tráfico marítimo es un problema complejo que se puede abordar de muy diferentes formas, siendo cada una de ellas más o menos eficaz según sean los objetivos que se pretendan cubrir y los recursos disponibles para ello.

Aunque la señalización de las comunicaciones AIS no es la única fuente de información, se ha comprobado que es la principal siendo utilizada en todos los sistemas analizados. Aunque no todos sus campos tienen la misma fiabilidad, se ha demostrado que no existe garantía de que las tripulaciones actualicen convenientemente los que deben modificarse manualmente.

El crecimiento de la actividad marítima junto con el desarrollo de la tecnología AIS ha resultado en la generación de grandes volúmenes de datos todos los días. Se dispone de una gran cantidad de datos, pero su problema radica en que no están etiquetados, dificultando su utilización en algoritmos de aprendizaje supervisado.

Por lo que una gran parte de los trabajos existentes no aprenden de los datos, sino que se basan en el establecimiento de reglas o umbrales definidos de antemano que son indicativos de un conjunto restringido de actividades presuntamente ilícitas que desarrollan las embarcaciones. La definición de

esas reglas o patrones requiere de la implicación de especialistas, lo que es muy costoso tanto en términos económicos, como de tiempo. Además puede resultar en la introducción del sesgo humano.

En cambio las principales dificultades a las que enfrentan los sistemas no supervisados son de diferente índole, como estar limitados al análisis de unas pocas variables o características por la complejidad y necesidades de cómputo que exige aumentar su número, o que se estudie principalmente el comportamiento individual de las embarcaciones y no en su conjunto que es fundamental para la detección de algunas anomalías.

Además, los sistemas no supervisados presentan un problema importante, la dificultad en la interpretabilidad de sus detecciones, con la desconfianza que esto genera en los resultados que presentan estos sistemas ante el personal que los explota. Llegar a entender qué está pasando detrás de una anomalía detectada por un sistema de detección de anomalías llega a ser un desafío para los operadores.

Así que hay que aunque en todos los sistemas de información de detección de anomalías es importante entregar la información de forma ordenada, clasificada y con representaciones visuales que ayuden a interpretar las alarmas que produzcan, en el caso de los no supervisados es fundamental.

Referencias

1. «DESIGN AND IMPLEMENTATION OF A MARITIME TRAFFIC MODELING AND ANOMALY DETECTION METHOD» [En línea]. Available: <https://www.diva-portal.org/smash/get/diva2:833998/FULLTEXT01.pdf> [Último acceso: 7 diciembre 2021]
2. «Four Techniques For Outlier Detection» [En línea]. Available: <https://www.knime.com/blog/four-techniques-for-outlier-detection> [Último acceso: 25 noviembre 2020] Autor/Autores: Maarit Widmann & Moritz Heine [Publicado: 1 de octubre del 2018]
3. «Vessel Trajectories Outliers » [En línea]. Available: https://easychair.org/publications/preprint_open/FbMB [Último acceso: 15 noviembre 2020] Autor/Autores: Tomas Machado, Rui Maia, Pedro Santos and João Ferreira [Publicado: 24 de mayo 2018]
4. «Loitering with intent—Catching the outliervessels at sea» [En línea]. Available: <https://journals.plos.org/plosone/article/file?id=10.1371/journal.pone.0200189&type=printable> [Último acceso: 17 noviembre 2020] Autor/Autores: Jessica H. Ford, David Peel, Britta Denise Hardesty, Uwe Rosebrock, Chris Wilcox [Publicado: 12 de julio 2018]
5. Fernando Terroso-Saenz, Mercedes Valdes-Vela, and Antonio F Skarmeta-Gomez. A complex event processing approach to detect abnormal behaviours in the marine environment. Editorial: Information Systems Frontiers, 18(4):765–780, 2016 [En línea]. Available: https://www.researchgate.net/publication/276532092_A_complex_event_processing_approach_to_detect_abnormal_behaviours_in_the_marine_environment
6. «Detecting search and rescue missions from ais data» [En línea]. Available: <http://www.master-project-h2020.eu/wp-content/uploads/2018/09/ICDE-Tserpes-et-al.pdf> [Último acceso: 8 diciembre 2020]
7. «Abstracting and reasoning over ship trajectories and web data with the Simple Event Model (SEM)» [En línea]. Available: <https://link.springer.com/article/10.1007/s11042-010-0680-2> [Último acceso: 8 diciembre 2020]
8. «Crisis: Integrating ais and ocean data streams using semantic web standards for event detection» [En línea]. Available:

- https://www.researchgate.net/publication/335935262_CRISIS_Integrating_AIS_and_Ocean_Data_Streams_Using_Semantic_Web_Standards_for_Event_Detection [Último acceso: 5 diciembre 2020]
9. «Open data for anomaly detection in maritime surveillance» [En línea]. Available: <https://www.diva-portal.org/smash/get/diva2:832155/FULLTEXT01.pdf> [Último acceso: 7 diciembre 2020]
 10. «Event Recognition for Maritime Surveillance» [En línea]. Available: <http://openproceedings.org/2015/conf/edbt/paper-364.pdf> [Último acceso: 2 diciembre 2020]
 11. «Online Event Recognition from Moving Vessel Trajectories» [En línea]. Available: <https://arxiv.org/pdf/1601.06041.pdf> [Último acceso: 2 diciembre 2020]
 12. «Maritime anomaly detection and threat assessment» [En línea]. Available: https://www.researchgate.net/publication/224218677_Maritime_anomaly_detection_and_threat_assessment [Último acceso: 8 diciembre 2020]
 13. «Semantic modelling of ship behavior in harbor based on ontology and dynamic bayesian network» [En línea]. Available: <https://www.mdpi.com/2220-9964/8/3/107> [Último acceso: 09 diciembre 2020]
 14. «Statistical analysis of motion patterns in data: Anomaly detection and motion prediction» [En línea]. Available: https://www.researchgate.net/publication/4370120_Statistical_Analysis_of_Motion_Patterns_in_AIS_Data_Anomaly_Detection_and_Motion_Prediction [Último acceso: 10 diciembre 2020]
 15. «Maritime anomaly detection using gaussian process active learning» [En línea]. Available: https://www.researchgate.net/publication/261308815_Maritime_anomaly_detection_using_Gaussian_Process_active_learning [Último acceso: 10 diciembre 2020]
 16. «Improving maritime anomaly detection and situation awareness through interactive visualization» [En línea]. Available: https://www.researchgate.net/publication/4370119_Improving_maritime_anomaly_detection_and_situation_awareness_through_interactive_visualization [Último acceso: 10 diciembre 2020]
 17. «Anomaly detection in maritime data based on geometrical analysis of trajectories» [En línea]. Available: https://c4i.gmu.edu/~pcosta/F15/data/fileservers/file/472181/filename/Paper_1570106169.pdf [Último acceso: 9 diciembre 2020]
 18. «Associative learning of vessel motion patterns for maritime situation awareness» [En línea]. Available: <https://fusion.isif.org/proceedings/fusion06CD/Papers/234.pdf> [Último acceso: 5 diciembre 2020]
 19. «A multi-task deep learning architecture for maritime surveillance using ais data streams» [En línea]. Available: <https://arxiv.org/pdf/1806.03972.pdf> [Último acceso: 11 diciembre 2020]
 20. «Vessel pattern knowledge discovery from ais data: A framework for anomaly detection and route prediction» [En línea]. Available: <https://www.mdpi.com/1099-4300/15/6/2218> [Último acceso: 12 diciembre 2020]
 21. «Extracting rules from expert operators to support situation awareness in maritime surveillance» [En línea]. Available: <http://fusion.isif.org/proceedings/fusion08CD/papers/1569106719.pdf> [Último acceso: 12 diciembre 2020]
 22. «A network abstraction of multi-vessel trajectory data for detecting anomalies» [En línea]. Available: <https://zenodo.org/record/2649606#.X9PWTthKg2w> [Último acceso: 2 diciembre 2020]
 23. «Mining maritime traffic conflict trajectories from a massive AIS data» [En línea]. Available: https://www.researchgate.net/publication/331945393_Mining_maritime_traffic_conflict_trajectories_from_a_massive_AIS_data [Último acceso: 26 diciembre 2020].