

Méthode de gestion et d'analyse de l'information spatiale et temporelle maritime

Anne-Clémence DUVERGER*, Cyril RAY*

*Institut de Recherche de l'École Navale
École Navale - CC 600
29240 Brest Cedex 9, FRANCE
prenom.nom@ecole-navale.fr

Résumé. Le suivi, l'analyse et la compréhension des mobilités maritimes par essence libre de circulation est un élément majeur de nos sociétés modernes tant elles témoignent des enjeux commerciaux, sociétaux ou géopolitiques qui existent entre les pays. Les volumes de données décrivant ces mobilités maritimes sont croissants, fluctuant et l'extraction et la gestion des connaissances au profit, par exemple, du suivi des pêches, des échanges commerciaux ou des actes illicites, etc. nécessitent une démarche d'étude complémentaire entre ingénierie de l'information géographique, intelligence artificielle et sciences humaines. Dans ce contexte, notre objectif est de développer des capacités pluridisciplinaires d'analyse et d'interprétabilité des mobilités maritimes (quels sont les navires, quelles sont leurs caractéristiques, etc.), d'analyse de leurs comportements maritimes (où se déplacent-ils, à quelle vitesse, etc.) et d'analyse de leurs situations maritimes (dans quel contexte se déplacent-ils, pourquoi, quel est le contexte international, etc.). Dans cet article, nous proposons une méthodologie d'analyse relevant de l'intelligence géospatiale et présentons l'architecture de gestion des données construites. Les travaux sont illustrés par une étude préliminaire des enjeux numériques maritimes de la zone Indo-pacifique.

Introduction

Nous avons assisté à une territorialisation continue de l'espace maritime (Miossec, 2014). En effet, les trajets des navires forment aujourd'hui de véritables « *autoroutes de la mer* » (Royer, 2012). Ces routes maritimes relient principalement l'Europe, l'Asie orientale et l'Amérique du Nord, néanmoins, les routes nord-sud ou sud-sud ne cessent de progresser. Le transport maritime assure aujourd'hui plus de 80 % des échanges mondiaux (Fremont et Fremont-Vanacore, 2014). Les profondeurs des océans sont aussi soumises à ce phénomène d'expansion des espaces anthropisés : l'océan est devenu un véritable « *front-pionnier* » (Miossec, 2014) (ZEE¹ en litige, exploitation de ressources naturelles, pose de câbles sous-marins...), engendrant ainsi de nouvelles conflictualités qui s'accompagnent d'un besoin accru de gestion de l'information maritime (qui est par nature spatiale et temporelle).

1. Zone économique exclusive

Pourtant, l'intelligence géospatiale² s'est jusqu'à présent peu saisie des problématiques liées à l'environnement maritime. Pour le géographe Philippe Boulanger, cette dernière se définit « *comme un processus de fusion de données géolocalisées et géo-référencées à partir d'une diversité de sources qu'elles soient militaires ou civiles* » (Boulanger, 2019a). De plus, elle est une discipline globale, qui permet la synthèse de différents champs de compétences, tant dans le domaine des sciences humaines : géopolitique, économie, sociologie, histoire, géographie, anthropologie, que dans celui des technologies (Morel et Boulanger, 2016).

Le CICDE (2021)³ précise que l'intelligence géospatiale se caractérise par l'utilisation de données spatio-temporelles : « *données (...) géolocalisées, géoréférencées et préférentiellement horodatées (...)* » et est basée sur des « *procédés de calcul complexes* » en vue de procéder à des analyses spatiales. En effet, avec l'avènement d'un volume toujours plus massif de données (*Big Data* en anglais), les algorithmes deviennent incontournables pour le traitement de ces dernières. L'analyse spatiale maritime n'échappe pas à ce constat.

En effet, si les données spatio-temporelles des navires, notamment les données du système AIS⁴, sont précieuses pour comprendre les activités maritimes, celles-ci sont croissantes et volumineuses. L'agence européenne pour la sécurité maritime (EMSA), par exemple, annonce traiter environ 50 millions de messages de position par jour. Cela demande d'optimiser la mise en place de procédés adaptés à l'ensemble du processus d'exploitation de ces données spatio-temporelles : de leur stockage jusqu'à leur représentation en passant l'examen critique de leur apport à l'analyse des situations maritimes.

Dans nos travaux nous cherchons à poser les bases d'une méthodologie générique relevant de l'intelligence géospatiale maritime. Dans ce contexte, nous reviendrons premièrement sur la genèse de l'intelligence géospatiale maritime tout en participant à sa formalisation (section 1), puis nous décrirons la représentation des données spatio-temporelles maritimes *via* des graphes et des hexagones (section 2). La section 3 présente les mécanismes de traitement et de stockage des données ainsi que quelques algorithmes réalisés. Enfin une étude de cas relevant de l'intelligence géospatiale sera décrite (l'analyse de la vulnérabilité des câbles sous-marins en Indo-Pacifique) (section 4).

1 L'intelligence géospatiale et les données spatio-temporelles maritimes

1.1 L'émergence d'une intelligence géospatiale maritime

L'intelligence géospatiale est le fruit d'un retour, depuis les années 1990, d'une géographie militaire intégrant tous les composants géographiques et ce sous l'effet de l'importance donnée au renseignement et de l'émergence des « NTIC »⁵ (Boulanger, 2006). En effet, l'information géographique repose aujourd'hui sur des principes nouveaux, liés à l'emploi des hautes technologies, à l'immédiateté des informations (Boulanger, 2011).

« *La Géo intelligence traite fondamentalement de théâtres terrestres* » (Prazuck, 2022) car elle a pris son essor durant les guerres du Koweït (1990-1991), de Bosnie (1991-1995), du

2. Dans le sens anglo-saxon du terme intelligence ; i.e. renseignement.

3. Centre interarmées de concepts, de doctrines et d'expérimentations

4. Automatic Identification System

5. Nouvelles technologies de l'information et de la communication

Kosovo (1999), d'Irak (2003) et d'Afghanistan (2001) (Boulanger, 2020). Ce n'est que plus récemment qu'elle a été appliquée aux opérations maritimes. Dans ce domaine, l'Inde est pionnière ; en 2011, elle crée le *National Command Control Communication and Intelligence*⁶. De plus, depuis 2007 des conférences dédiées à l'intelligence géospatiale sont organisées (sur le modèle de celles de l'USGIF⁷) ; dans le cadre du partenariat stratégique entre ce pays et les États-Unis (Boulanger, 2019b). La dernière, intitulée « *Indo-Pacifique Geointelligence 2023* » organisée en juin 2023 à New Delhi, réunissait des responsables gouvernementaux, des militaires, des experts des domaines du spatial, géospatial, de l'industrie et des universitaires.

En effet, l'intelligence géospatiale a pénétré le monde académique, élargissant de ses champs de recherches : santé (Gehlen et al., 2019), protection de l'environnement (Abu Sari et al., 2018), gestion de crise (Markogiannaki et al., 2020), aménagement du territoire (Parambil et al., 2021). En France, les études relevant de l'intelligence géospatiale appliquée sont encore rares et dans le domaine maritime, seuls quelques travaux s'en approchent : César Ducruet (2013, 2015, 2016).

1.2 L'intelligence géospatiale maritime, essai de définition

En s'appuyant sur les travaux du CICDE, nous avons identifié les principaux composants caractérisant l'intelligence géospatiale maritime : elle exploite du renseignement multi-sources ; c'est (1) une approche globale qui permet (2) la contextualisation des données issues de toutes sources ; (3) les données sont exploitées dans des SIG⁸ et *via* (4) des procédés d'automatisation de calculs ; elle permet (5) l'analyse prédictive et l'anticipation ; elle est essentielle (6) à la planification ; elle est (7) assimilable à un système d'information, car elle permet de collecter, entreposer, traiter et distribuer de l'information ; enfin, (8) elle structure des flux d'informations provenant d'acteurs privés ou institutionnels.

Rajoutons que l'intelligence géospatiale maritime doit considérer les quatre dimensions de ce milieu : la surface, la colonne d'eau, le plancher océanique et l'espace, en raison de son degré d'anisotropie⁹ moins important que celui de l'espace terrestre ; il existe moins de structures de surface sur lesquelles se baser pour caractériser une situation.

Ainsi, l'intelligence géospatiale maritime s'appuie sur différents capteurs générant des informations issues des quatre dimensions de ce milieu (cf. figure 1). Sur les mers et océans (surface), les navires sont dotés de systèmes de localisation : Automatic Identification System (AIS) (destiné à tous les navires), Vessel monitoring system (VMS) qui est système de surveillance des navires de pêche ou encore Voyage Data Recorder (VDR) qui est un enregistreur de données du voyage. En surface, on peut également collecter des données issues du renseignement humain, des caméras des ports et des RADAR¹⁰. D'autres types de données sont disponibles pour l'environnement maritime immergé : données des bathysondes, SONAR¹¹,

6. Ce projet de sécurité maritime, lancé après les attentats de Mumbai (2008), s'appuie sur un système de surveillance qui comprend des radar, caméras, avions, satellites et le système d'identification automatique des navires (AIS). Les données issues de ces capteurs sont collectées par le centre de gestion et d'analyse de l'information de la marine indienne (IMAC).

7. United States Geospatial Intelligence Foundation

8. Système d'information géographique

9. L'anisotropie caractérise un espace orienté, qui s'ordonne selon des axes, qui obéit à des polarisations (Tabarly, 2005) ; il n'existe pas d'espace géographique isotrope.

10. Radio detection and ranging, détection et estimation de la distance par ondes radio

11. Sound Navigation and Ranging, système de navigation et de télémétrie par écho sonore

Intelligence Géospatiale Maritime

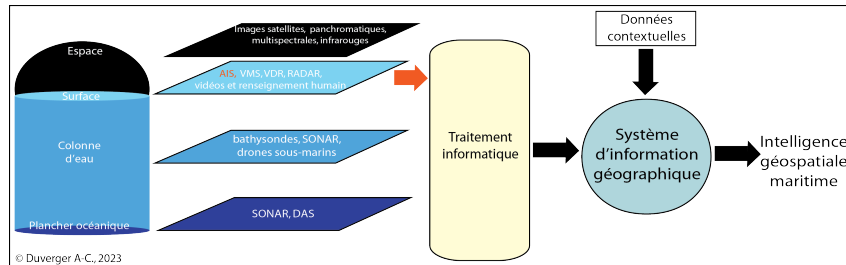


FIG. 1 – L'intelligence géospatiale maritime : milieu physique, données et méthode

des drones sous-marins et DAS¹². Enfin, les satellites se déplaçant au-dessus de l'espace maritime sont une source de données importante.

Au regard des sources d'informations disponibles et en se basant sur les caractéristiques physiques du milieu maritime, l'intelligence géospatiale maritime peut se définir comme une méthode d'analyse qui permet d'acquérir des connaissances par la fusion de toutes ou une partie des données acquises dans les quatre dimensions du milieu maritime. Certaines d'entre elles sont particulières au milieu maritime, notamment les données AIS.

Ces dernières représentent un volume d'informations important. Ainsi, elles nécessitent des systèmes de stockage adaptés et doivent être traitées par des algorithmes pour faciliter leur représentation. Il est ensuite possible de les transférer dans un SIG, afin de les spatialiser et de procéder à la production de cartes. D'autres types de représentations (graphiques et schémas) permettent d'enrichir la compréhension de la dimension spatiale et temporelle de l'information.

Rajoutons que l'intégration de données contextuelles relatives au milieu maritime permet d'enrichir l'analyse tout facilitant l'interprétation des données de localisation de navires, car celles-ci sont partielles et pas toujours fiables : erreurs, falsification, usurpation des messages (Iphar et al., 2015).

2 Représentation des données spatio-temporelles maritimes

2.1 Des données massives

Notre méthodologie d'analyse pour l'intelligence géospatiale a été implémentée et testée avec un jeu de données contenant les localisations de navires dans le monde durant l'année 2019. Ces données sont obtenues grâce au système AIS (*Automatic Identification System*). L'AIS est un dispositif destiné à la prévention des collisions en mer mais également permettant un suivi en temps réel navires¹³. Les navires qui en sont équipés (environ 400 000 à l'échelle mondiale) transmettent régulièrement leur position (quelques secondes) ainsi que leur identité et leurs informations de voyage (quelques minutes)¹⁴.

12. Distributed Acoustic Sensing, système de détection acoustique distribuée.

13. Accessible par exemple via le portail Web Marine Traffic : <https://www.marinetraffic.com>.

14. Numéro MMSI (Marine Mobile Service Identity), numéro IMO (identifiant de l'Organisation maritime internationale), nom, type de navire, dimensions, tirant d'eau, position (latitude et longitude), vitesse sur le fond, cap, cap sur le sol, rayon de giration, dangerosité de la cargaison, port de destination, temps estimé d'arrivée).

Le jeu de données intègre deux versions : les messages AIS d’origines (environ 1,7 Go par jour) et les données AIS prétraitées, sous-échantillonnées à 5 minutes mais agrégées avec les données de la *Lloyds* (environ 6,5 Go par jour). Ce deuxième format constitue notre base de travail principale car elle apporte 37 champs complémentaires aux données AIS ; e.g. propriétaire, opérateur, année de construction du navire, etc. Le volume de données rend difficile la fouille des données historiques et l’extraction d’informations pertinentes par un utilisateur. Nous avons mis en oeuvre deux approches d’abstraction des données AIS historiques ; une indexation hexagonale et un graphe hiérarchiques.

2.2 Représentations hexagonales des données maritimes

Nous avons basé une partie des algorithmes et cartographies sur un système d’indexation géospatiale hiérarchique hexagonal (selon le modèle Uber H3) qui rend la solution originale à plus d’un titre par rapport aux travaux existants (Maslek Elayam et al., 2022). Les hexagones sont utilisés dans plusieurs traitements en remplacement des coordonnées géographiques précises des navires ou pour construire une connaissance sur le temps long (statistique) de l’espace maritime.

Ainsi, l’espace maritime est divisé en cellules hexagonales d’une résolution donnée (la résolution est paramétrable). L’hexagone est une structure de données adaptée car permettant de stocker des statistiques maritimes historiques telles que le nombre de navires ayant traversé chaque hexagone pendant une certaine période (densité de l’hexagone) et leur vitesse minimale, maximale et moyenne. Ces statistiques et bien d’autres liées à la trajectoire fournissent aux algorithmes des informations utiles pour vérifier les comportements réguliers et anormaux.

La précision et la pertinence des statistiques calculées dépendent de plusieurs facteurs tels que la taille des données sur zone, la résolution de l’hexagone et l’intervalle de temps (microscopique : par heure, mésoscopique : par jour ou macroscopique : par semaine ou par mois).

Au-delà des bénéfices algorithmiques de cette approche basée sur les hexagones, elle a démontré un avantage certain dans la réduction des volumes de données et dans l’accroissement des performances computationnelles.

2.3 Représentations par modèle de graphe

Afin d’identifier des récurrences et certaines anomalies dans les trajets des navires nous avons opté pour la définition d’un modèle de graphe (Maslek Elayam et al., 2021). Dans ce dernier, les nœuds représentent les origines et destinations et les points saillants d’une trajectoire de navire. Les arcs sont les flux entre ces nœuds. Le modèle de graphe conçu est hiérarchique, ainsi il peut être construit à partir de données AIS historiques (cf. Figure 2). Des fonctions d’agrégation permettent d’abstraire le graphe spatio-temporel à différentes échelles spatiales et temporelles. La dimension spatiale (i.e. les positions des navires) est réalisée sous forme d’une indexation sémantique par grille régulière et hiérarchique (i.e. avec les hexagones H3).

L’objectif de cette structure de données est double. Premièrement cela permet l’exploration et la compréhension du réseau maritime ainsi que l’étude de son évolution dans le temps. Deuxièmement cela permet de rechercher et d’analyser le cycle de vie d’un navire ou de manière agrégée d’un type de navires ou encore d’un pavillon.

Deux métriques sur graphe ont été implémentées ; Elles permettent l’analyse du réseau maritime et de fait favorisent la compréhension de cycles de vie des navires, leurs régularités

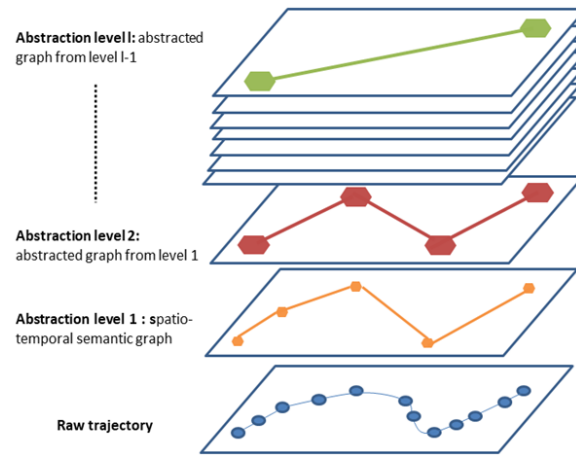


FIG. 2 – Modèle de graphe hiérarchique

ou comportements inhabituels au regard du trafic mondial. Premièrement, l’accessibilité des nœuds. L’accessibilité correspond aux facilités disponibles pour atteindre un nœud en prenant en compte le coût dû au franchissement de la séparation spatiale avec les autres nœuds. Nous avons ainsi proposé, implémenté et évalué une accessibilité hiérarchique compatible avec le modèle de graphe. Deuxièmement, nous nous sommes intéressés à l’évolution dynamique du graphe. Pour cela nous avons défini une mesure de similarité pour évaluer les modifications du graphe dans le temps. La mesure prend en compte la hiérarchie du modèle de graphe et la différence de granularité temporelle entre les intervalles de temps évalués.

3 Gestion des données spatio-temporelles maritimes

Les informations de localisation des navires sont indispensables à une démarche relevant de l’intelligence géospatiale maritime. Dans ce contexte, la gestion des données historiques est aussi importante que l’intégration des flux de données temps-réel. Afin de gérer, et de requêter ces données nous avons développé une chaîne de traitement permettant de gérer le volume et de les mettre à disposition pour les différents cas d’usage. Nous avons opté pour une architecture basée sur la suite ELK composée de Elastic Search (base de données NoSQL), Logstash (utilitaire de collecte de log) et Kibana (Outils de supervision, d’exploration et d’analyse de base de données) qui alimente des développements en Python et une base de données graphe conçue pour l’analyse du réseau maritime et le cycle de vie des navires (cf. Figure 3).

Dans cette chaîne de traitement, les données de localisation historiques (cf. section 2.1) et temps-réel sont intégrées dans la base ELK. Similairement les données dites externes utiles à la construction des cas d’usage (zones géographiques, conditions environnementales, informations documentaires, etc.) peuvent être intégrées à la base de données. Enfin nous y associons une base de données portuaire mondiale (environ 21 000 entités) comprenant la localisation du port, son nom et son extension spatiale sous forme hexagonale issus du clustering de données AIS historiques (Ménard et al., 2021). Du point de vue technique, la suite ELK est instan-

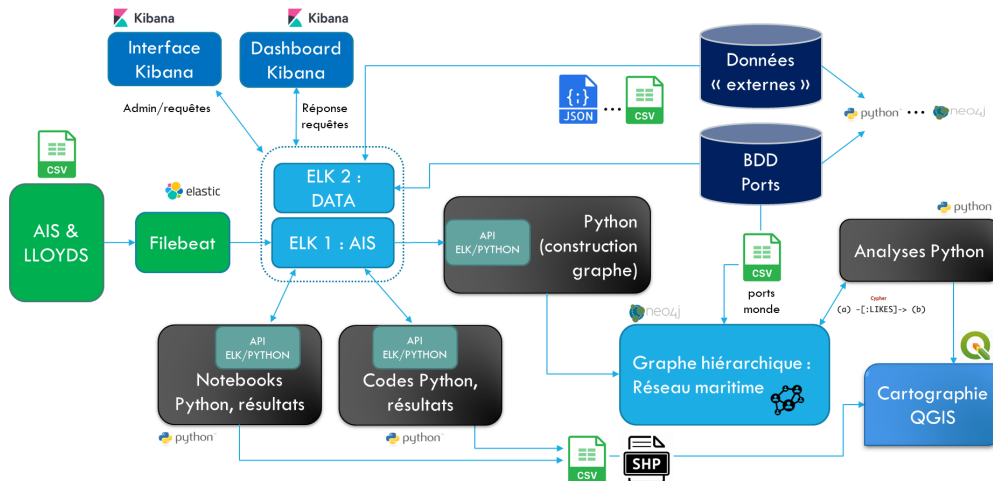


FIG. 3 – Chaîne de traitement pour l'intelligence géospatiale maritime

ciée (sans répliquas) dans 3 conteneurs où chaque conteneur dispose de 64Go de mémoire, 30 processeurs, un volume de données et 30 *shards*.

Le modèle de graphe hiérarchique a été mis en œuvre et testé sur les données réelles de trajectoires maritimes. L'implémentation combine plusieurs outils tels que Python, la bibliothèque MovingPandas¹⁵, l'index H3 d'Uber, le système de gestion de base de données de graphes Neo4j (Community Edition disponible sous licence GPLv3) et son langage d'interrogation Cypher (Maslek Elayam et al., 2022). L'algorithme fonctionne en trois grandes étapes. Après une étape de pré-traitement (eg. suppression d'éventuels doublons à vitesse nulle), les trajectoires sont extraites à l'aide de MovingPandas. Les trajectoires sont ensuite annotées avec les index des hexagones traversés puis compression pour ne retenir que les points remarquables des trajectoires (i.e. point tournants, origines et destinations). Ces points sont ensuite agrégés avec HDBSCAN pour définir les noeuds du graphe. Cette structure peut ensuite être requêtée et les résultats (cycle de vie des navires, sous-graphe, mesure de centralité, etc.) peuvent être utilisés par l'analyste pour enrichir ou annoter sa production cartographique.

Kibana est un plug-in de visualisation de données développé pour permettre la création de visualisations à partir de bases de données Elasticsearch. Kibana est conçu pour être utilisé à partir d'un navigateur web, et permet tant d'explorer des jeux de données de manière approfondie à l'aide d'une interface graphique facile d'utilisation que de créer des visualisations utiles à l'interprétation ou pour compléter les cartes produites à l'aide du logiciel QGIS. Kibana permet notamment de créer des dashboards automatisés, qui peuvent aussi bien inclure les visualisations statistiques habituelles d'un dashboard (histogrammes, série temporelles, graphes, "donuts", etc...) que des cartes intégrant de l'information géospatiale. La figure 4 illustre un exemple simple dans lequel une carte présente certains navires en mer de Chine et trois graphiques présentent les nombres de présence de ces navires et leur répartition durant l'année.

La partie analyse est pour l'essentiel réalisée en Python, notamment au travers de *notebooks* permettant l'extraction, le traitement est l'analyse des données nécessaires aux diffé-

15. <https://github.com/anitagraser/movingpandas>.

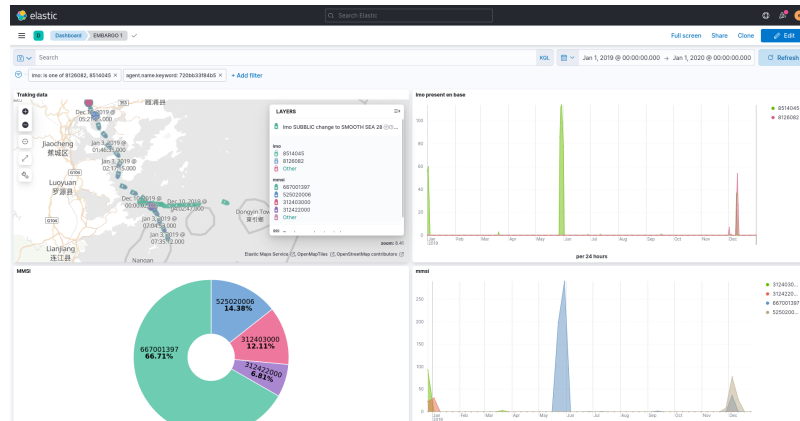


FIG. 4 – Exemple de dashboard Kibana

rentes études. Les développements des algorithmes, méthodes et études réalisés sont guidés par le besoin d'analyse. Actuellement nous avons mis en oeuvre quelques développements communs à plusieurs études sur le prétraitement des données AIS, l'exploitation des hexagones, etc. D'un point vu algorithmique nous avons également réalisé quatre fonctionnalités basées sur les données AIS historiques :

Analyse des flux maritimes : afin de connaître le flux maritime journalier, par exemple, port à port ; mer à mer ou zone économique à zone économique, nous avons conçu un algorithme générique permettant de quantifier les entrées et les sorties dans une zone et applicable sur n'importe quel espace maritime. L'algorithme développé permet de produire des tableaux de bord des flux maritimes entre zones géographiques. Le résultat est généralement présenté sous forme tabulaire ou de *chord diagram* en fonction des besoins (Figure 7).

Le changement d'identité : l'algorithme développé vise à détecter les disparitions et ré-apparitions de navires accompagnées d'un changement d'identité (susceptible d'être illicite). Pour cela, lorsqu'un navire cesse d'émettre, nous avons développé une expansion hexagonale multi-cibles (multi-threadé d'un point de vue informatique) avec une loi d'évolution gaussienne adaptée au type de navire. La couverture hexagonale s'étend avec le temps. L'apparition de navires dans ces bulles d'expansion déclenche la recherche de falsification d'identité (Figure 8).

L'étude de la densité et de la couverture AIS : un des freins à la parfaite étude de la situation maritime de surface est la présence de zone d'ombres susceptibles de masquer des activités maritimes et qui rend difficile l'identification de la malversation du système (Salmon et al., 2016). Nous avons ainsi conçu et implémenté deux algorithmes. Le premier algorithme agrège les informations historiques de localisation pour produire des cartes de densité (Figure 6). Le résultat est un ensemble d'hexagones (de taille paramétrable), chacun annoté avec des informations statistiques (nombre de navires, vitesse moyenne, etc.). En complément un algorithme d'apprentissage (basé sur XGBoost) pour la prédiction de la couverture AIS d'un récepteur a été conçu. L'algorithme utilise les données AIS historiques couplées aux données

environnementales collectées par le programme Européen Copernicus. Les données sont alignées spatialement et temporellement grâce aux hexagones H3. Les cartes résultantes (période paramétrable) sont également des cartes d'hexagones.

La prédiction de trajectoires : dans ce travail nous avons souhaité explorer la pertinence d'algorithmes d'apprentissage pour la prédiction de trajectoires représentées à l'aide d'hexagones H3. Nous avons opté pour une approche originale qui transforme les indexes de la séquence d'hexagones d'une trajectoire en caractère de langue naturelle (le chinois en l'occurrence). Nous avons ensuite utilisé le modèle RoBERTa (Robustly Optimized BERT Pretraining Approach) sur ces textes pour faire de la prédiction de trajectoires. La prédiction est une séquence d'hexagones (Figure 5).



FIG. 5 – Exemple de suivi de route courbe par le modèle prédictif. Jaune : hexagones donnés au modèle, rouge : trajectoire réelle, vert : prévision

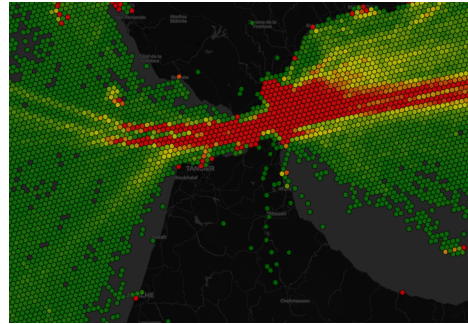


FIG. 6 – Représentation hexagonale de la densité normalisée du trafic maritime au Détroit de Gibraltar sur un mois (janvier 2016)

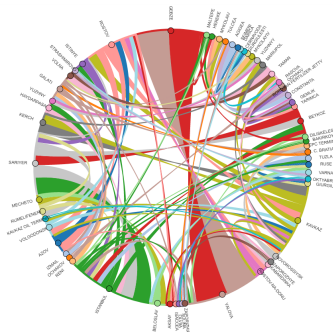


FIG. 7 – Flux maritimes entre les ports de la Mer noire sur une journée (juin 2021)

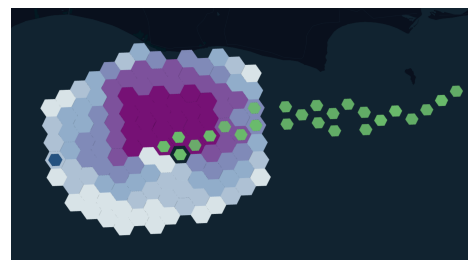


FIG. 8 – Estimation de l'expansion hexagonale au bout de 21 minutes après la disparition d'un cargo (trajectoire verte) en mer de Chine (hexagone de résolution 8 : arrête de 461 mètres)

4 Étude des enjeux des câbles sous-marins en zone Indo-Pacifique

Les rapports de la National Geospatial Intelligence Agency (NGIA) attribuent à l'intelligence géospatiale maritime un nombre restreint d'objets d'étude : cartographies nécessaires à la navigation, planification des opérations maritimes (NGIA, 2006) et sécurité durant la navigation (NGIA, 2018). Pourtant, au regard de la littérature spécialisée, nous avons identifié huit domaines de l'intelligence géospatiale maritime (cf. Figure 9) : les activités de pêches illégales, la sécurité environnementale (lutte contre les pollutions sauvages, préservation des ressources marines), les incidents maritimes, les migrations illégales, le trafic de stupéfiants, la piraterie, le terrorisme et enfin le développement du numérique (interaction entre les activités humaines et les infrastructures de communication). C'est ce dernier axe que nous proposons de développer dans cette section : les enjeux relatifs aux câbles sous-marins.

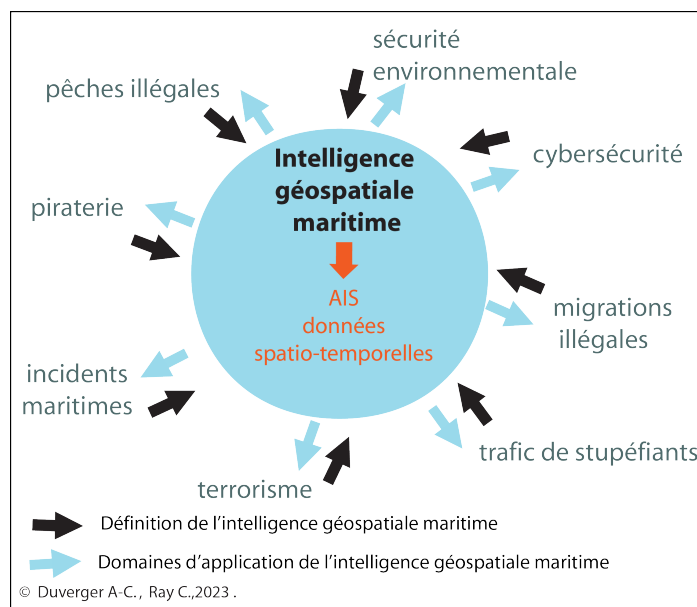


FIG. 9 – Les domaines d'analyse de l'intelligence géospatiale maritime

4.1 Objet et zone d'étude

Les principales artères de communication sont au nombre environ de 450 (Morel, 2023) et permettent à 95% / 97% de nos communications de transiter (Bivaud, 2016). Les vulnérabilités potentielles de ces infrastructures sont de deux ordres : des cyber-vulnérabilités (de type électronique, radio, ou sur l'attaque des protocoles réseaux tel que BGP) et des vulnérabilités physiques (Galland, 2010). En effet, leur endommagement peut conduire à la déstabilisation d'un État ou d'une région puisque l'ensemble de nos activités sont aujourd'hui dépendantes

de réseaux numériques : services administratifs ou bancaires, secteur des transports, du commerce, de la santé, de l'enseignement ou de la recherche (Cattaruzza, 2021).

Nous abordons ce thème sur la zone Indo-Pacifique en raison de deux faits : son caractère hautement maritime et le basculement des enjeux géopolitiques mondiaux vers cette partie du monde. Cependant, ce vaste espace fait l'objet de délimitations variables. Dans le cas de la France, ces dernières adoptent une vision périmétrique à partir de ses sols (Lechevry, 2019). Ainsi, en considérant les espaces d'outre-mer comme les sièges de la projection française en Indo-Pacifique, nous pouvons diviser cet espace en quatre zones : l'Indo-Pacifique occidental, central, austral et oriental. L'étude actuellement menée se concentre sur sa partie orientale, car l'état de la recherche en géographie a montré que les transports maritimes ont peu été analysés dans cette région¹⁶.

4.2 Traitements algorithmiques, indexations hexagonales des données spatio-temporelles et données contextuelles

Afin d'analyser la vulnérabilité des câbles sous-marins pour la partie orientale de l'Indo-Pacifique, plusieurs traitements algorithmiques sont nécessaires (cf. figure 6). Le premier s'appuie sur la dimension temporelle des données AIS ; il s'agit de les filtrer selon une période précise. Pour le présent cas d'étude, des données AIS sont sélectionnées afin d'effectuer des corrélations avec le dernier état des lieux du réseau de câbles (datant de 2018). Notons que cette analyse est centrée sur les territoires français de la zone.

1. Premièrement un filtrage temporel est réalisé ; le mois de juillet 2019 dans notre cas (un mois hors de la période cyclonique ; cette dernière pouvant impacter les activités des navires).
2. La seconde étape vise à sélectionner les données d'une zone géographique précise, qui peut être définie selon deux variables : les coordonnées géographiques et par le biais des délimitations des mers et océans du monde (données issues du *Flanders Marine Institute*).

Éventuellement, en fonction des cas d'usage, une étape de prétraitement des données est possible (par exemple ; ne conserver que les navires en mouvement). Ensuite, les données AIS sont intégrées dans des hexagones Uber H3. Ces derniers concernent les zones à proximité des câbles, permettant de créer « *un profil de risque le long du câble* » (Doan et al., 2016). Pour cette étude à l'échelle régionale, des hexagones d'une résolution de 4 ont été choisis (arête d'une longueur moyenne de 22,6 km) (Maslek Elayam et al., 2022). Ainsi, au regard de l'échelle de la carte : 1/400 km, nous obtenons un niveau de détail suffisant, tout en agrégeant un volume important d'informations.

3. La troisième étape consiste à agréger dans les hexagones les navires présents, selon leur type (cargos, navire militaires, remorqueurs, voiliers...).
4. La quatrième étape agrège les positions des navires câbliers. Elle est basée sur la liste de ces navires établie par l'*International Cable Protection Committee*.

16. Soit les espaces maritimes compris entre le 40^{ème} parallèle nord et le 60^{ème} parallèle sud ; entre le 100^{ème} parallèle est et le 60^{ème} parallèle ouest.

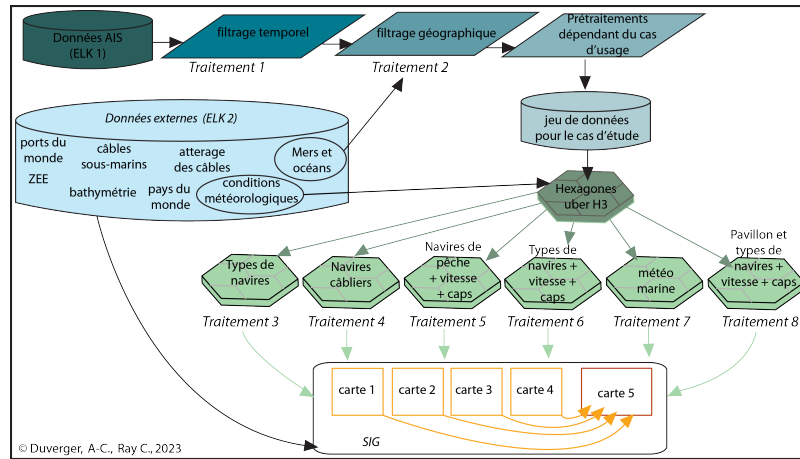


FIG. 10 – L'intelligence géospatiale appliquée à l'analyse des câbles sous-marins en Indo-Pacique

5. La cinquième étape permet de construire des hexagones contenant seulement des informations relatives aux navires de pêche (nombre, vitesse et caps : cap vrai et cap suivi sur le fond).
6. La sixième étape vise à obtenir des renseignements sur les divers types de navires, mais également sur leur vitesse et leurs caps.
7. Ensuite, un autre traitement génère dans les hexagones des données météorologiques : vagues (hauteur) et vents (vitesse) selon leur direction (données produites par le programme *Copernicus* de l'Union européenne).
8. Enfin, le huitième traitement de ces données spatio-temporelles permet de construire des hexagones contenant à la fois les types de navires présents, leurs pavillons ainsi que leur vitesse et leurs caps.

Afin d'enrichir l'analyse, des données contextuelles sont nécessaires : celles des câbles sous-marins et leurs terminaux (produites par *Telegeography*¹⁷) concernent l'année 2018. Les mers et océans du monde ; des données constituées par la *Flanders Marine Institute* (2018). Les informations relatives aux ZEE ont aussi été produites par la *Flanders Marine Institute* et la *Vlaams Instituut voor de Zee* (2018). Les frontières administratives des pays (*World Food Programme*, 2022) sont aussi indispensables à l'analyse. Les données relatives aux ports et zones stationnaires mobilisées sont produites par des chercheurs ayant fusionné quatre bases de données : celle du *World Port Index* (*World Food Programme*) ; les mouillages de *Global Fishing Watch* et la base de ports de l'IFREMER¹⁸ (Menard et al., 2021). Concernant la bathymétrie, deux jeux de données sont utilisés : celui de la *North American Cartographic Information Society* (2009) et celui de l'OHI et de la COI¹⁹.

17. Une société spécialisée dans le marché des télécommunications

18. Institut français de recherche pour l'exploitation de la mer

19. Organisation hydrographique internationale et Commission océanographique intergouvernementale

4.3 Projets cartographiques et étude de cas

Les traitements algorithmiques générant l'indexation hexagonale des données AIS (sous la forme de fichier Shapefile) ainsi que l'ajout de données ouvertes dans le SIG, permettront la réalisation de cartes relatives à la vulnérabilité des câbles sous-marins :

- densité des navires et vulnérabilité des câbles sous-marins.
- activités de pêche et vulnérabilité des câbles sous-marins.
- ancrage des navires et vulnérabilité des câbles sous-marins.
- actes de malveillance et vulnérabilité des câbles sous-marins.

Puis, un indice de vulnérabilité des câbles sera créé. Celui-ci sera représenté sous la forme de diagrammes radiaux. Ces derniers comporteront six dimensions et seront associés à chaque hexagone. Les différentes dimensions des diagrammes sont construites en associant des valeurs statistiques connues (par exemple, 44 % des incidents sur les câbles sont liés aux activités de pêche et 14 % à des problèmes d'ancrage des navires) (ICPC, UNEP, WCMC, 2009) à des informations relatives aux navires présents et à leur environnement (densité de navires, bathymétrie, météorologie marine, nombre de câbles par pays et temps d'intervention des navires câblés). Ces dernières varient selon la zone d'étude considérée, rendant cette méthode d'identification de zones à risques pour les câbles sous-marins, applicable à d'autres zones géographiques.

Enfin, nous nous concentrerons sur une rupture de câble avérée : le 10 mai 2023 aux Îles Salomon. Une étude de cas sélectionnée, car elle correspond à une situation dominante : la majorité des endommagements de câbles sont liés à des activités de pêche (ICPC, UNEP, WCMC, 2009). De plus, les Îles Salomon sont un territoire insulaire²⁰. En ce sens, elles entretiennent des caractéristiques communes avec les collectivités françaises²¹ de la zone qui, du fait de leur multi-insularité, sont généralement plus vulnérables aux pannes d'Internet liées aux câbles sous-marins, car elles n'ont pas accès à des réseaux câblés terrestres denses (Bueger et al., 2022).

Cette situation d'endommagement de câble permettra de prévenir des événements similaires. Pour cela, nous nous baserons sur le comportement du navire impliqué, sur ses caractéristiques de circulation et l'environnement maritime dans lequel il évoluait au moment de l'incident. Nous travaillons à partir de données AIS relevant de trois filtres temporels différentes :

- les 12 mois précédents l'incident nous renseigneront sur le comportement routinier du navire (mouillages privilégiés, trajets réguliers, zones de pêche...);
- l'heure qui a précédé et celle qui a suivi la rupture du câble permettront de décrire avec précision les conditions de survenue de celle-ci;
- enfin, le mois suivant l'incident est considéré afin de rendre compte des événements ayant engendré l'arrivée tardive du câblé²².

Pour apprécier l'environnement maritime du navire concerné, nous nous appuyons sur la localisation des câbles sous-marins; ce jeu de données comprend des informations indispensables à une approche géopolitique (date d'installation et propriétaire du câble). Les données

20. Environ 1000 îles (MACBIO, 2019) dont 164 sont habitées (SINSO, 2009).

21. La Nouvelle-Calédonie compte environ 141 îles dont 15 sont peuplées (Duverger, 2021); la Polynésie française comprend 121 îles dont 72 sont habitées (ISPF, 2015).

22. Ce dernier n'est arrivé sur place que le 20 juin (DCmag, 2023).

relatives aux ZEE et aux frontières des pays nourrissent aussi ce type d'analyse, car l'extrémité d'un câble se situe généralement dans d'autres pays (O'Malley, 2019).

Les zones d'atterrages sont également intégrées, car si elles sont proximité les unes des autres, il existe un risque accru de défaillance dû à une même perturbation (Bueger et al., 2022). Ainsi, on pourra caractériser la gravité de l'événement. Les données relatives aux ports et zones stationnaires permettent d'analyser les interactions entre les situations statiques et l'événement. Nous exploitons également des la bathymétrie pour enrichir notre examen de la situation du câble. Les données de vitesse du vent et de hauteur des vagues (selon leur direction) le 10 mai 2023 sont intégrées, afin d'approcher l'influence des conditions météorologiques sur la trajectoire du navire. Enfin, cette étude demande de recourir à des hexagones de résolution 11 (arrête d'une longueur moyenne de 24.9 m) (Maslek Elayam et al., 2022), pour rendre compte de la densité de navires à une grande échelle, et ce, pour le filtrage temporel proche de l'heure de l'incident.

Conclusion

Nous avons mis en évidence qu'en raison du basculement des enjeux géopolitiques mondiaux vers la zone Indo-Pacifique, l'intelligence géospatiale doit aujourd'hui se tourner vers l'analyse de l'espace maritime. Pour cela elle dispose de diverses sources d'informations associées aux quatre dimensions physiques du milieu maritime, notamment les données AIS. Cependant, en raison de leur caractère massif (géoréférencées, horodétées et enrichies d'informations statistiques), ces dernières appellent à penser des modes de gestions, de représentation et d'analyse spécifiques, nécessitant des compétences qui relèvent aussi bien des sciences de la donnée, de la géomatique que de la géographie.

Ainsi, du point de vue de la gestion et de l'analyse de ces données spatio-temporelles maritimes, nous avons montré l'intérêt de la construction d'une infrastructure basée sur la suite Elastic Search associée à des développements en Python. Ils permettent de procéder à la création de graphes et d'algorithmes originaux indispensables à la représentation et à la compréhension des mobilités et flux maritimes. Par ailleurs, l'étude de la vulnérabilité des câbles sous-marins confirme le potentiel de l'indexation hexagonale pour spatialiser des informations diverses (densité, vitesses, pavillons des navires...).

Enfin, nous avons pu constater que du fait des détournements inhérents à la nature de ces données, les algorithmes développés permettent d'affiner les analyses en détectant les anomalies et en offrant la possibilité de reconstruire les trajectoires des navires. Cela ouvre de nombreuses perspectives d'études dans divers domaines de l'intelligence géospatiale : du terrorisme à l'examen des pêches illégales.

Références

- Abu Sari, N., M. Y. B. Abu Sari, A. Ahmad, S. Sahib, et F. Othman (2018). Using laper quadcopter imagery for precision oil palm geospatial intelligence (op geoint). *Electronic and Computer Engineering* 10(1), 25–33.
- Bivaud, A. (2016). Les dommages causés aux câbles sous-marins.
- Boulanger, P. (2006). *Géographie militaire*. Carrefours, Ellipses.

- Boulanger, P. (2011). Renseignement géographique et culture militaire. *Hérodote* n°140, 47–63.
- Boulanger, P. (2019a). Le geoint à l’origine d’une nouvelle science de l’information géospatiale ? *Défense* n°200(2-3), pp.32–36.
- Boulanger, P. (2019b). Le geospatial intelligence français : quels enjeux et défis aujourd’hui ? *Défense et stratégie hors-série*(63), 70–76.
- Boulanger, P. (2020). *La géographie reine des batailles*. Perrin, Paris.
- Bueger, C., T. Liebetrau, et J. Franken (2022). Security threats to undersea communications cables and infrastructure – consequences for the eu. Technical report, European Parliament and Bueger Christian and Liebetrau Tobias and Franken Jonas.
- Cattaruzza, A. (2021). Vers une géopolitique numérique. *Constructif* 3(60), 46–50.
- CICDE (juillet, 2021). Renseignement géospatial, doctrine interarmées-dia-2.17. Technical report, Centre interarmées de concepts, de doctrines et d’expérimentations, Ministère des Armées.
- DCmag (Mis en ligne en juin 2023, modifié en août 2023). Le navire a arraché un câble sous-marin en mer de corail, capitaine et propriétaire feront face à la justice. <https://dcmag.fr/le-navire-a-arrache-un-cable-sous-marin-en-mer-de-corail>.
- Doan, H., L. Macnay, A. Savadogo, et K. Smith (2016). Profondeur d’enfouissement des câbles offshore par une approche basée sur le risque. *Journées Nationales de Géotechnique et de Géologie de l’Ingénieur*, 1–8.
- Ducruet, C. (2013). Network diversity and maritime flows. *Journal of Transport Geography* 30, 77–88.
- Ducruet, C. (2015). *Maritime Networks : Spatial Structures and Time Dynamics*, Chapter Co-evolutionary dynamics of ports and cities in the global maritime network, 1950–90, pp. 351–373. Routledge Studies in Transport Analysis.
- Ducruet, C. (2016). La spatialité des réseaux maritimes : Contributions maritimes à l’analyse des réseaux en géographie.
- Duverger, A.-C. (2021). Systèmes de transports et de mobilités inter-îles en nouvelle-calédonie et au vanuatu.
- Fremont, A. et A. Fremont-Vanacore (2014). *Géographie des espaces maritimes*. La documentation française, Dossier n°8104, Paris.
- Galland, J.-P. (2010). Critique de la notion d’infrastructure critique. *Flux* 3(81), 6–18.
- Gehlen, M., M. R. Nicola, E. R. Costa, V. K. Cabral, E. L. de Quadros, C. O. Chaves, R. A. Lahm, A. D. R. Nicolella, M. L. Rossetti, et D. R. Silva (2019). Geospatial intelligence and health analytics : Its application and utility in a city with high tuberculosis incidence in brazil. *journal of Infection and Public Health* 12, 681–689.
- ICPC-UNEP-WCMC (2009). Submarine cables and the oceans : connecting the world. Technical report, United Nations Environment Program and World Conservation Monitoring Centre and International Cable Protection Committee.
- Iphar, C., A. Napoli, et C. Ray (2015). Detection of false ais messages for the improvement of maritime situational awareness. *OCEANS 2015 MTS/IEEE Washington*, 1–7.

- ISPF (2015). Polynésie française en bref. Technical report, Institut de la Statistique de la Polynésie française.
- Lechevry, C. (2019). La place des outre-mer océaniens dans la politique indo-pacifique de la France. *Revue de Défense Nationale* 8(823), 21–27.
- MACBIO (2019). Marine atlas maximizing benefits for Solomon Islands. Technical report, Marine and Coastal Biodiversity Management in Pacific Island Countries, Deutsche Gesellschaft für Internationale Zusammenarbeit and GRID-Arendal and International Union for Conservation of Nature and Republic of Solomon Islands and Secretariat of the Pacific Regional Environment Programme.
- Markogiannaki, O., A. Karavias, D. Bafi, D. Angelou, et I. Parcharidis (2020). A geospatial intelligence application to support post-disaster inspections based on local exposure information and on co-seismic dingsar results : the case of the durres (albania) earthquake on november 26, 2019. *Natural Hazard* 103, 3085–3100.
- Maslek Elayam, M., G. Kerhoas, V. Lambert, C. Ray, et A. Ménard (2022). On the interest of hexagonal abstraction of maritime information. In *OCEANS 2022, Hampton Roads, Hampton Roads, VA, USA*.
- Maslek Elayam, M., C. Ray, et C. Claramunt (2021). Modèle de graphe pour l’analyse des structures de trajectoires maritimes. In *21ème édition de la conférence Extraction et Gestion des Connaissances (EGC)*.
- Maslek Elayam, M., C. Ray, et C. Claramunt (2022). A hierarchical graph-based model for mobility data representation and analysis. *Data and Knowledge Engineering* 141, 102054.
- Ménard, A., V. L. de Cursay, C. Ray, M. E. M. Charles Guenois, et M. Dréau (2021). Construction d’une méta-base de ports à l’échelle mondiale. *SAGEO-2021, poster*, 1–70.
- Miossec, A. (2014). *Géographie des mers et océans*. Presses universitaires de Rennes.
- Morel, C. (2023). *Les câbles sous-marins*. Biblis, CNRS Edition, Paris.
- Morel, E. et P. Boulanger (2016). *Géographie et guerre. De la géographie militaire au Geospatial Intelligence en France (XVIIIe-XXIe siècle)*, Chapter La géolocalisation et le GEOINT comme outil d’analyse sociétales et géopolitique, pp. 189–197. Bulletin de la Société de géographie.
- Ménard, A., V. Lambert de Cursay, C. Ray, C. Guenois, M. Maslek Elayam, et M. Dréau (2021). Construction d’une méta-base de ports à l’échelle mondiale. In *Conférence internationale de Géomatique et d’Analyse Spatiale. SAGEO’21*.
- NGIA (avril 2018). Geospatial intelligence (geoint) basic doctrine. Technical report, National Geospatial Intelligence Agency, National System for Geospatial Intelligence.
- NGIA (septembre 2006). Geospatial intelligence (geoint) basic doctrine. Technical report, National Geospatial Intelligence Agency, Office of Geospatial-Intelligence Management.
- O’Malley, S. (2019). Vulnerability of south korea’s undersea cable communications infrastructure. a geopolitical perspective. *Korea Observer* 50(3), 309–330.
- Parambil, N. V., S. Krishnan, et M. Firoz (2021). *Geo-intelligence for Sustainable Development*, Chapter Geo-intelligence-Based Approach for Sustainable Development of Peri-Urban Areas : A Case Study of Kozhikode City, Kerala (India), pp. 35–52. Springer.

- Prazuck, C. (2022). *Le Geoint et les océans*. dans : Boulanger Philippe (Dir.), Géographie, Geoint et opérations, Bulletin de liaison des membres de la Société de géographie, Hors-série, pp. 82-89.
- Royer, P. (2012). *Géopolitique des Mers et des Océans*. PUF, Paris.
- Salmon, L., C. Ray, et C. Claramunt (2016). Continuous detection of black holes for moving objects at sea. pp. 1–10.
- SINSO (2009). Basic tables and census description. Technical report, Solomon Islands National Statistical Office.
- Tabarly, S. (Mis en ligne en 2005, modifié en mars 2023). Isotropie et anisotropie. <http://geoconfluences.ens-lyon.fr/glossaire/isotropie-anisotropie-1>.

Remerciement

Ce travail est réalisé dans le cadre du projet national GEOINT financé par l'Agence d'Innovation de la Défense (AID).

Summary

Monitoring, analyzing and understanding maritime mobility is a major concern of modern society, as it reflects the commercial, societal and geopolitical stakes that exist between countries. The volumes of data describing these maritime mobilities are growing and fluctuating, and the extraction and management of knowledge for the benefit of, for example, fisheries monitoring, trade, illegal acts, etc. requires a complementary approach combining geographic information engineering, artificial intelligence and human sciences. In this context, our aim is to develop multi-disciplinary capabilities for analyzing and interpreting maritime mobilities (what are the ships, what are their characteristics, etc.), analyzing their maritime behaviors (where do they move, at what speed, etc.) and analyzing their maritime situations (in what context do they move, why, what is the international context, etc.). In this article, we propose an analysis methodology and present the data management architecture constructed. The work is illustrated by a preliminary study of the maritime digital challenges of the Indo-Pacific region.

