

---

# Analyse de similarité de trajectoires d'objets mobiles suivant le même itinéraire

**Laurent ETIENNE\*** — **Thomas DEVOGELE\*** — **Alain BOUJU\*\***

\* *Institut de recherche de l'École Navale (IRENAV)  
Lanvéoc-Poulmic, CC 600, 29240 BREST Cedex 9  
prenom.nom@ecole-navale.fr*

\*\* *Laboratoire d'informatique et d'imagerie industrielle  
17042 La Rochelle Cedex 1  
abouju@univ-lr.fr*

---

*RÉSUMÉ. De plus en plus d'objets mobiles sont équipés de capteurs qui facilitent la surveillance de leur évolution et permettent la réception, le stockage et la visualisation en temps réel de leur position. Actuellement, la fouille de données dans des historiques de déplacements volumineux offre des perspectives intéressantes dans le domaine de l'étude du comportement de ces objets. A partir de résultats d'analyses statistiques sur des ensembles de trajectoires, des patrons spatio-temporels (routes types et couloirs associés) ont été définis. En s'appuyant sur ces patrons, il est possible de détecter en temps réel des situations inhabituelles et d'en qualifier le comportement. Dans cet article, cette méthode est appliquée aux déplacements de navires dans un espace maritime ouvert afin de détecter des comportements inhabituels (retard, hors route, etc.).*

*ABSTRACT. More and more mobile objects are now equipped with sensors allowing real time monitoring of their movements which can be stored in spatio-temporal databases. Nowadays, data mining on a huge quantity of stored positions allows to analyse the behaviour of those mobile objects. From results of statistical analysis on sets of trajectories, spatio-temporal patterns can be defined (main routes and spatio-temporal channels). Using these patterns, unusual situations can be detected in real time. Furthermore, mobile object behaviour can be defined by comparing a trajectory with a main route. In this article, this technique is applied to ship movements in an open maritime area in order to detect unusual behaviours such as being ahead of schedule or delayed, veering to the left or to the right of the main route.*

*MOTS-CLÉS : trajectoires d'objets mobiles, similitude, route type, couloir spatio-temporel, comportement anormal.*

*KEYWORDS: mobile objects trajectories, similarity, main route, spatio-temporal channel, outliers.*

---

## 1. Introduction

Aujourd'hui, de plus en plus d'objets mobiles sont équipés de systèmes permettant de transmettre de manière précise des informations sur leur déplacement. Ces différents capteurs génèrent un volume important de données stockées dans des bases de données spatio-temporelles (*BDST*) afin de permettre à des applications de les analyser. Grâce à ces informations, il est possible de suivre en temps réel l'évolution de ces objets mobiles. La fouille de données de ces *BDST* offre de nouveaux débouchés concernant l'analyse des trajectoires de ces objets mobiles et la détection de comportements inhabituels. De tels outils d'analyse de trajectoires couplés à une visualisation appropriée offrent à des opérateurs de surveillance la possibilité de se focaliser sur certaines trajectoires considérées comme anormales. Le volume de données à traiter en temps réel étant de plus en plus important, cette aide est indispensable. L'application de ces outils au contexte maritime a pour objectif d'augmenter la sécurité du trafic dans des zones denses. Ainsi, les comportements inhabituels de navires peuvent être soulignés. Pour définir ceux-ci, il est nécessaire au préalable de déterminer différents critères définissant les trajectoires. Ceux-ci sont définis manuellement par un opérateur ou calculés à partir de l'analyse statistique de données passées. L'objectif de cette étude est d'analyser l'historique d'un ensemble de trajectoires afin d'en déduire des comportements types. Ceux-ci permettent de qualifier une trajectoire et donc le comportement du navire.

La première section de cet article introduit un état de l'art concernant le suivi d'objets mobiles. Dans la seconde section, nous présentons une méthode d'extraction et de filtrage de trajectoires d'objets mobiles disposant d'un même itinéraire à partir d'une base de données spatio-temporelles. La troisième section traite de l'analyse statistique spatiale et temporelle d'un groupe de trajectoires. Enfin, la quatrième section expose les différents outils d'analyse qualitative spatiale et temporelle. Une étude de cas qui porte sur les navires à passagers en rade de Brest illustre les différentes sections.

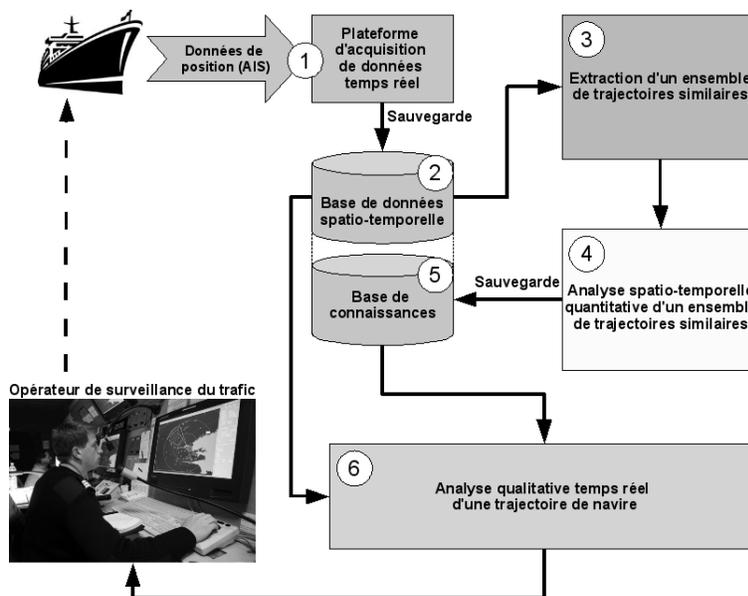
## 2. Etat de l'art concernant le suivi d'objets mobiles

Le suivi d'objets mobiles est utilisé couramment dans de nombreux domaines tels que la migration des animaux (Mefteh *et al.*, 2008), le suivi de phénomènes météorologiques (Lee *et al.*, 2008), les mouvements de foules, de piétons (Knorr *et al.*, 2000), les déplacements de véhicules (automobiles, aéronefs, navires...) (Wang *et al.*, 2006; Baud *et al.*, 2007; Kharrat *et al.*, 2008). Ce suivi d'objets mobiles peut être couplé avec des systèmes d'analyse ayant pour objectif d'optimiser les performances d'un système (comme par exemple les applications d'optimisation du fret pour des véhicules). Aujourd'hui, la gestion par des bases de données d'objets mobiles ouvre des perspectives inédites. En effet, les déplacements de ces objets peuvent être analysés sur des espaces et des durées importants. Les bases de données spatio-temporelles permettent de connaître précisément pour chaque objet mobile, l'historique de ses déplacements. L'analyse de ces déplacements offre la possibilité de déduire des comportements types de déplacements d'objets similaires.

La détection des comportements anormaux d'objets mobiles parmi un ensemble de trajectoires est une problématique de recherche active. Différentes approches s'intéressent à la problématique de détection de trajectoires (ou de portions de trajectoires) anormales au sein d'un ensemble de données de positions. Elles peuvent être classées en fonction de la méthode utilisée pour détecter des données marginales basées sur la distribution (Barnett *et al.*, 1994), la distance (Knorr *et al.*, 1998; Knorr *et al.*, 1999; Knorr *et al.*, 2000; Ramaswamy *et al.*, 2000; Shim *et al.*, 2003; Lee *et al.*, 2008) ou la densité (Gaffney *et al.*, 1999; Breunig *et al.*, 2000; Aggarwal *et al.*, 2001; Papadimitriou *et al.*, 2003; Kharrat *et al.*, 2008; Lee *et al.*, 2008). Les différents travaux s'appuyant sur une analyse de distance se fondent sur une somme pondérée d'un ensemble de valeurs traduisant les vitesses et les directions (moyennes, minimales et maximales) ainsi que les positions de départ et d'arrivée de chaque trajectoire. Les trajectoires sont ainsi analysées de manière globale. Des variations locales sont plus difficilement détectables avec cette méthode. L'analyse basée sur la densité définit un paramètre nommé « local outlier factor » (LOF) pour chaque position d'une trajectoire représentant la densité locale par calcul de la distance au plus proche voisin des autres trajectoires de l'ensemble. Cette technique permet de détecter les positions les plus éloignées des autres pour chaque position d'une trajectoire. Cependant, cela nécessite un nombre important de calculs de distance du plus proche voisin. Lee présente une approche basée sur la partition d'un ensemble de trajectoires et la détection des parties de trajectoires qui n'ont pas assez de voisins (Lee *et al.*, 2008). Ces dernières sont considérées comme inhabituelles. Cette approche à l'avantage de détecter localement des parties de trajectoires inhabituelles à partir de critères de densité et de distance. Cependant la dimension temporelle n'est absolument pas prise en compte.

Cet article s'intéresse plus particulièrement au cadre maritime et à l'analyse des trajectoires de navires dont le trafic ne cesse d'augmenter. Le système AIS (*Automatic Identification System*) permet le suivi en temps réel des navires à l'aide de capteurs GPS et de systèmes de transmission sans fil (VHF) (IMO, 2007). Les navires à passagers, les transports internationaux, d'un fort tonnage ou transportant des produits dangereux sont d'ores et déjà équipés de transpondeurs AIS. Ces équipements sont de plus en plus répandus. Ils sont employés essentiellement dans le cadre de la sécurité maritime à des fins d'identification, de visualisation temps réel et de détection des risques de collision. Cependant, ils sont limités par la portée de communication du transmetteur VHF (environ 30 km). Qui plus est, ces systèmes ne disposent pas de capacité d'analyse des trajectoires de navires. C'est pourquoi, l'objectif principal de cet article est de proposer des outils d'analyse spatiale et temporelle des trajectoires de navires. Partant du principe que des navires d'un même type suivant le même itinéraire ont des comportements similaires et suivent une route optimisée, il est intéressant d'analyser les trajectoires de ces navires afin d'en déduire des patrons spatio-temporels permettant de qualifier leur comportement. La combinaison des systèmes de suivi et d'analyse temps réel est fort utile en termes d'aide à la décision dans le cadre de la surveillance du trafic. Ces systèmes répondent aux problématiques de sécurité inhérentes aux déplacements relatifs des navires (croisements, collisions) tout en permettant la détection de comportements inhabituels (éviter de collision, panne moteur, chan-

gement de route pour porter assistance, etc.) et la prédiction de déplacements futurs. Les travaux de recherche décrits dans cet article suivent les recommandations de l'International Maritime Organization concernant l'e-navigation pour une meilleure sécurité en mer et la protection de l'environnement (IMO, 2007). De plus, ceux-ci sont complémentaires des récents travaux visant à prendre en compte les statistiques d'accidents (Degré, 2007), les objets mobiles ayant des comportements illicites (Morel *et al.*, 2008) ou les données météorologiques (Richard, 2007) en termes d'analyse de trajectoires.



**Figure 1.** Démarche méthodologique générale

La démarche méthodologique générale proposée afin d'identifier les positions inhabituelles est résumée par la figure 1. Cette analyse se décompose en deux parties. Dans un premier temps, une fouille de données est réalisée. Elle s'appuie sur une base de données spatio-temporelles (figure 1 étape 2) contenant des positions des navires recueillies en temps réel à l'aide d'une plateforme d'acquisition (figure 1 étape 1). Tout d'abord, les trajectoires suivant un même itinéraire sont extraites de la *BDST* (figure 1 étape 3). Une analyse statistique permet alors de définir des patrons spatio-temporels pour chaque itinéraire (figure 1 étape 4). Ces patrons sont mémorisés dans une base de connaissances (figure 1 étape 5). Dans un deuxième temps, pour chaque nouvelle position d'un navire obtenue en temps réel, il est possible d'y associer un itinéraire en fonction de l'historique des positions du navire, sa position actuelle ainsi que la destination fournie. Dans le domaine maritime, cette information est transmise par l'AIS du navire. Ainsi, la comparaison de la position avec les patrons associés à l'itinéraire suivi permet de qualifier en temps réel le comportement du navire (figure 1

étape 6). Les sections suivantes de cet article présentent les différentes phases de cette démarche.

### 3. Extraction de trajectoires dans une base de données spatio-temporelles

Afin d'analyser un historique de trajectoires d'objets mobiles, il est nécessaire de sauvegarder l'évolution spatiale de ces objets en fonction du temps. Pour cela, il est possible d'utiliser une base de données spatio-temporelles, constituée d'un ensemble de données discrètes ayant des caractéristiques spatiales et temporelles (Güting, 1994; Mouza *et al.*, 2000). Ces *BDST* disposent d'outils permettant d'effectuer des requêtes sur ces ensembles de données portant sur des critères spatiaux et temporels. Des requêtes plus complexes de type OLAP (Wan *et al.*, 2007) sont également envisageables. Formellement, une position d'un objet mobile  $o$  est composée de coordonnées spatiales et d'une estampille temporelle correspondant à la date à laquelle l'objet mobile était à cette position (temps absolu). De même, la trajectoire  $T_o$  d'un objet mobile  $o$  est définie comme une suite de positions  $P_{oj}$  ordonnée temporellement  $T_o = (P_{o1}, \dots, P_{oj}, \dots, P_{on})$ .

La base de données spatio-temporelles étudiée dans l'exemple maritime contient 1 005 navires ainsi que 4 821 447 positions recueillies depuis Mai 2007 dans la mer d'Iroise. Chaque position est associée à un navire. La *BDST* offrant des outils d'extraction spatiale, il est intéressant de les utiliser pour discriminer des objets mobiles évoluant dans différents espaces. Ainsi, il est possible de définir des zones spatiales ( $Z$ ) représentées par des surfaces géométriques ( $Z.g$ ) ayant un intérêt particulier (zone portuaire, chenal d'accès). Une position  $P_o$  est incluse dans une zone  $Z_i$  si ses coordonnées géométriques sont incluses dans la géométrie surfacique  $Z_i.g$  de la zone.

#### 3.1. Définition d'un graphe de zones

Afin de pouvoir sélectionner un groupe de trajectoires suivant un même itinéraire, il est nécessaire de préciser la notion d'itinéraire grâce à l'élaboration d'un graphe de zones. L'espace maritime est un espace ouvert qui permet de se déplacer d'une zone à une autre, en traversant des zones de passage (par exemple, aller de Brest à Camaret *via* le goulet de Brest). La notion d'itinéraires de zones va maintenant être formalisée. Les zones sont liées à des lieux remarquables par exemple, un port, un détroit, un dispositif de séparation de trafic (*DST*). Ces zones d'intérêt peuvent être définies manuellement par l'opérateur en fonction de différents critères tels que la réglementation (zones d'attente, *DST*), la géographie (isthmes, détroits, goulets), l'économie (ports). Pour aller d'une zone à une autre, un graphe orienté est défini. Les zones correspondent aux sommets de ce graphe de zones ( $G_Z$ ). Un itinéraire peut ainsi être défini comme une suite ordonnée de zones reliées deux à deux par des arcs, correspondant à une chaîne du graphe de zones. Un itinéraire étant composé d'au moins un arc, il dispose par conséquent d'une zone de départ  $Z_D$  et d'arrivée  $Z_A$ .

Une trajectoire  $T_o$  est considérée comme suivant un itinéraire  $I$  passant par des sommets du graphe  $G_Z$  si elle respecte les conditions suivantes :

Soit un itinéraire  $I = (Z_D, \dots, Z_i, \dots, Z_A)$

Soit une trajectoire  $T_o = (P_{od}, \dots, P_{oj}, \dots, P_{oa})$

La trajectoire  $T_o$  suit l'itinéraire  $I$  si :

$$\forall Z_i \in I, \exists P_{oj} \in T_o, P_{oj} \subset Z_i \quad [1]$$

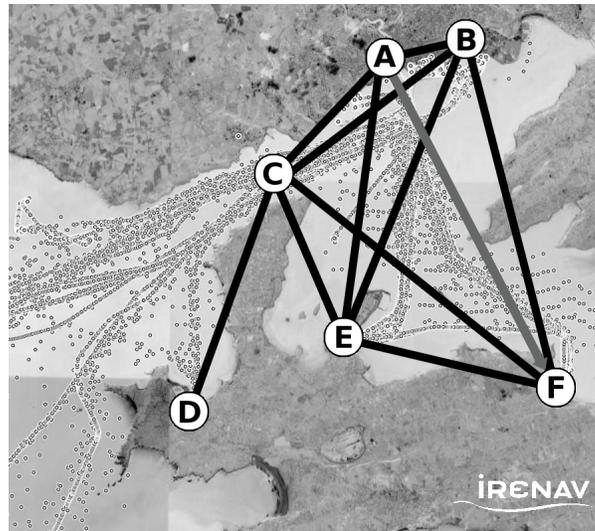
$$\forall P_{oj} \in T_o \wedge P_{oj} \subset Z_l, \forall P_{ok} \in T_o \wedge P_{ok} \subset Z_m, Z_l <_I Z_m \rightarrow P_{oj} < P_{ok} \quad [2]$$

$$\forall P_{oj} \in T_o \wedge P_{oj} \subset Z_i \rightarrow Z_i \in I \quad [3]$$

$$P_{oj} \subset Z_D \rightarrow P_{oj} = P_{od} \quad [4]$$

$$P_{oj} \subset Z_A \rightarrow P_{oj} = P_{oa} \quad [5]$$

En d'autres termes, pour chaque zone de l'itinéraire  $I$ , il existe au moins une position  $P_o$  de la trajectoire  $T_o$  dans cette zone [1] respectant la relation d'ordre temporelle définie précédemment [2]. Aucune position  $P_o$  de la trajectoire  $T_o$  n'appartient à une zone du graphe autre que celles de l'itinéraire [3]. Seule la première position  $P_{od}$  de la trajectoire appartient à la zone de départ de l'itinéraire  $Z_D$  [4]. De même, seule la dernière position  $P_{oa}$  de l'itinéraire appartient à la dernière zone de l'itinéraire  $Z_A$  [5]. Au vu de ces définitions, il est important de noter que pour toute trajectoire intersectant une zone, une position de cette trajectoire doit être incluse dans la zone. Ainsi, en fonction de l'emprise géographique de la zone, de la vitesse de déplacement de l'objet mobile et de sa fréquence d'émission, il est potentiellement nécessaire de suréchantillonner les trajectoires.



**Figure 2.** Graphe de zone associé à la BDST

L'itinéraire représenté sur la figure 2 par l'arc (A,F) (Brest Arsenal → Lanvéoc Ecole Navale) du graphe de zone  $G_Z$  sera distinct de celui représenté par la chaîne (A,E,F) (Brest Arsenal → Ile Longue → Lanvéoc Ecole Navale). Le graphe étant incomplet, tous ses sommets ne sont pas reliés directement entre eux par une arête. Par exemple en partant de Brest (sommets B), pour atteindre Camaret (sommets D), il est nécessaire de passer par le goulet de la rade de Brest (sommets C). De même, le graphe est orienté. Les itinéraires aller/retour peuvent être différents et les dispositifs de séparation de trafic dus aux règles de navigation sont ainsi modélisés. Les multiples points représentés sur la figure 2 correspondent aux positions enregistrées de navires. Des zones de navigation plus denses sont observables. Elles correspondent aux itinéraires empruntés par les navires.

### 3.2. Extraction d'un groupe homogène de trajectoires

Les notions de trajectoires et d'itinéraires ayant été formalisées, les critères d'extraction de trajectoires suivant un itinéraire vont être détaillés. L'objectif de cette section est d'extraire de la  $BDST$  les trajectoires d'un même type  $T$  d'objets mobiles suivant le même arc  $A$  d'un itinéraire  $I$ . Cet ensemble est appelé groupe homogène de trajectoires d'objets mobiles de même type suivant le même arc d'un itinéraire ( $GHT_{AIT}$ ).

Ainsi, le premier critère de sélection étant le type des objets mobiles, dans l'exemple, seuls les navires de type « navire à passagers » sont sélectionnés, soit 5 navires parmi 1 005. Ce critère de sélection ayant pour objectif d'obtenir une population homogène d'objets mobiles à analyser pouvant être n'importe quel critère *a priori* (par exemple, espèce d'animaux sur une période précise de migration, navire marchand avec chargement et prise en compte de la marée, déplacement d'un véhicule à une heure de pointe). Cette première sélection d'une population homogène d'objets mobiles nous permet d'obtenir des résultats statistiques satisfaisant au vu du nombre important de trajectoires analysées. Le second critère de sélection est basé sur un critère géographique, la première position de la trajectoire doit nécessairement se trouver dans la première zone de l'itinéraire, de même, la dernière position de la trajectoire doit être incluse dans la dernière zone de l'itinéraire. De plus, pour chaque trajectoire, il ne doit exister qu'une seule position de départ appartenant à la zone de départ (de même pour la zone d'arrivée). Finalement, le dernier critère employé est temporel, certains des objets mobiles pouvant effectuer ces itinéraires périodiquement. Il est donc nécessaire d'affecter ces différentes positions de départ et d'arrivée à des trajectoires distinctes.

Considérant la durée  $D_{Z_i Z_j}$  approximative nécessaire à ce type d'objet mobile pour réaliser le trajet entre chaque arc  $Z_i Z_j$  de l'itinéraire, il est possible de sélectionner des positions uniques de départ et d'arrivée de trajectoires pour chaque arc de l'itinéraire. En effet, si un objet même mobile met un temps  $D_{Z_i Z_j}$  pour effectuer le trajet entre les deux zones  $Z_i$  et  $Z_j$  et qu'il réalise des allers-retours, chaque point de départ d'une nouvelle trajectoire de ce même objet mobile est donc séparé

d'environ  $2 * D_{Z_i Z_j}$ . Par conséquent, le point d'arrivée  $P_{oa}$  de la trajectoire est forcément compris dans l'intervalle entre  $t_{Pod}$  et  $t_{Pod} + 2 * D_{Z_i Z_j}$ . Ainsi, une position de départ  $P_{od}$  est considérée comme valide s'il n'existe aucune autre position  $P_{oj}$  d'estampille temporelle  $t_{P_{oj}}$  incluse dans la zone de départ  $Z_D$  dans l'intervalle de temps  $t_{P_{oj}} + D_{Z_i Z_j}$ . De même, une position d'arrivée  $P_{oa}$  est considérée comme valide s'il n'existe aucune autre position  $P_{ok}$  d'estampille temporelle  $t_{P_{ok}}$  incluse dans la zone d'arrivée  $Z_A$  dans l'intervalle de temps  $t_{P_{ok}} - D_{Z_i Z_j}$ . Finalement, une position de départ  $P_{od}$  peut être appariée à une position d'arrivée  $P_{oa}$  si  $t_{Pod} < t_{P_{oa}} \wedge t_{P_{oa}} < (t_{Pod} + 2 * D_{Z_i Z_j})$ . Ceci permet d'extraire de la base de données spatio-temporelles toutes les positions constituant la trajectoire de l'objet mobile sur l'intervalle compris entre la position  $P_{od}$  et  $P_{oa}$ . Finalement, pour chacune de ces positions composant la trajectoire, il est nécessaire de vérifier la propriété [3] en s'assurant qu'aucune d'elles n'est incluse dans une zone du graphe  $G_Z$  n'appartenant pas à l'itinéraire  $I$ . L'ensemble des trajectoires ainsi reconstitué compose le  $GHT_{AIT}$  à analyser.

---

**Algorithme 1** Groupe homogène de trajectoires
 

---

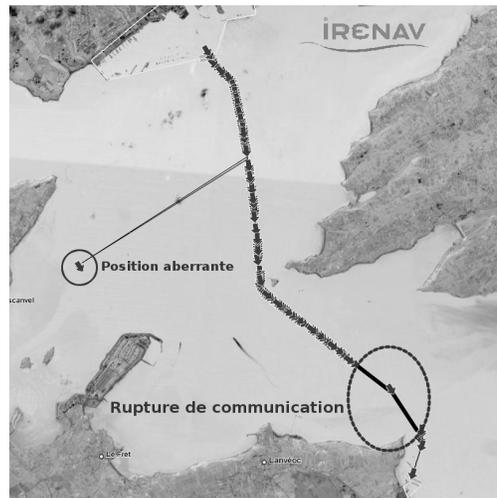
- 1: **Algorithme**  $GHT_{AIT}$ (Itinéraire  $I$ , Arc  $A$ , Type d'objet mobile  $T$ )
  - 2: Soit  $Z_D$  la zone de départ de l'arc  $A$
  - 3: Soit  $Z_A$  la zone d'arrivée de l'arc  $A$
  - 4: Soit  $D$  la durée estimée de l'arc ( $Z_D, Z_A$ )
  - 5: **Rechercher** l'ensemble des positions des objets de type  $T$  :  $EP_d \subset Z_D$
  - 6: **Rechercher** l'ensemble des positions des objets de type  $T$  :  $EP_a \subset Z_A$
  - 7: **Sélectionner** parmi  $EP_d$  le sous-ensemble  $EVP_d$  tel que :
  - 8:  $EVP_d \subset EP_d \wedge \forall P_{di} \in EVP_d, \nexists P_{dj} \in EVP_d, t_{P_{di}} < t_{P_{dj}} \wedge t_{P_{dj}} < (t_{P_{di}} + 2 * D)$
  - 9: **Sélectionner** parmi  $EP_a$  le sous-ensemble  $EVP_a$  tel que :
  - 10:  $EVP_a \subset EP_a \wedge \forall P_{ai} \in EVP_a, \nexists P_{aj} \in EVP_a, t_{P_{ai}} < t_{P_{aj}} \wedge t_{P_{aj}} < (t_{P_{ai}} - 2 * D)$
  - 11: **pour** chaque  $P_{di} \in EVP_d$  **faire**
  - 12:   **si**  $\exists P_{ja}, dure(P_{id}, P_{ja}) < 2 * D$  **alors**
  - 13:     Reconstituer la trajectoire  $T_k$  entre  $P_{id}$  et  $P_{ia}$
  - 14:     **si**  $\forall P_k \in T_k \wedge P_k \subset Z_i \Rightarrow Z_i = Z_D \vee Z_i = Z_A$  **alors**
  - 15:       **Ajouter**  $T_k$  au  $GHT_{AIT}$
  - 16:     **fin si**
  - 17: **fin si**
  - 18: **fin pour**
- 

L'algorithme 1 reprend de façon ordonnée l'ensemble des critères de sélection d'un  $GHT_{AIT}$ . Dans le cadre de notre exemple, cette méthode d'extraction de données a permis de définir 554 trajectoires de navires à passagers suivant l'itinéraire Brest Arsenal  $\rightarrow$  Lanvéoc Ecole Navale représenté par l'arc (A,F) sur la figure 2. L'exemple choisi n'utilise qu'un seul arc du graphe de zones afin de simplifier la selection et l'appariement des positions de départ et d'arrivée des trajectoires. Chaque arc d'un

itinéraire étant traité indépendamment. Le temps CPU nécessaire à l'exécution de cette étape appliqué à notre exemple comportant 4 800 000 positions est de 17 306 ms<sup>1</sup>.

### 3.3. Filtrage des trajectoires erronées

Une fois le  $GHT_{AIT}$  extrait de la base de données, il est nécessaire d'effectuer un premier filtrage des données pouvant amener un biais dans notre analyse. Les trajectoires contenant des ruptures de communication comme présenté sur la figure 3 doivent être filtrées. Ces ruptures de communication sont détectées à l'aide d'un seuil temporel maximum entre deux transmissions. Si la durée entre deux positions consécutives d'une trajectoire est supérieure à ce seuil, la trajectoire est alors supprimée du  $GHT_{AIT}$  à analyser. De même, certaines trajectoires peuvent comporter des positions aberrantes (comme indiqué sur la figure 3) dues à un dysfonctionnement du système de géolocalisation ou de transmission. Ces positions sont détectées lors du calcul des données de cap et de vitesse en comparant la vitesse calculée avec la vitesse maximale pouvant être atteinte par un objet mobile du type étudié. Les trajectoires contenant des positions aberrantes sont également retirées du  $GHT_{AIT}$  étudié.



**Figure 3.** *Suppression des données erronées*

Dans le cadre de notre exemple, parmi les 554 trajectoires, 506 trajectoires ont été conservées après filtrage des trajectoires erronées. Le nombre de trajectoires constituant le  $GHT_{AIT}$  après filtrage étant par conséquent très largement suffisant pour la génération par analyse statistique de la route type et du couloir spatio-temporel présentés dans la suite de cet article.

1. Mesure réalisée sur un PC disposant de deux CPU Intel Core2 6600 à 2,4 GHz et de 2 Go de mémoire.

### 3.4. Recalage spatial

Afin de disposer de trajectoires dont les positions de départ et d'arrivée sont indépendantes du moment d'émission, les parties de la trajectoire à l'intérieur des zones de départ et d'arrivée doivent également être filtrées. Sans ce filtrage, des biais importants peuvent être mesurés sur le patron spatio-temporel défini dans les étapes suivantes. Dans notre exemple, le nuage des positions de départ d'un  $GHT_{AIT}$  est représenté sur la figure 4(a). Il induit un écart entre les positions les plus éloignées pouvant atteindre jusqu'à 200 m. Les nouvelles positions de départ sont calculées en interpolant la position qui intersecte la ligne de départ virtuelle (représentant la frontière de la zone  $Z_D$ ) et chaque trajectoire du  $GHT_{AIT}$ . Il en est de même pour la zone d'arrivée. L'estampille temporelle de cette nouvelle position de départ est interpolée en fonction de la vitesse initiale de la position de départ. Le résultat du recalage spatial appliqué à notre exemple pour la zone de départ de l'itinéraire Brest → Lanvéoc, est illustré par la figure 4(b).

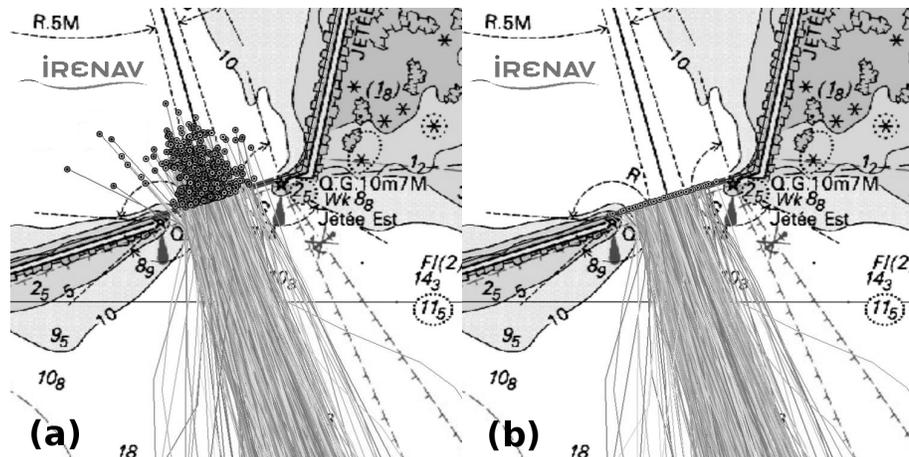


Figure 4. Recalage spatial des trajectoires

### 3.5. Filtrage de Douglas et Peucker spatio-temporel

Une fois le recalage spatial effectué, afin d'optimiser les temps de calcul, les trajectoires sont simplifiées en utilisant un filtrage basé sur une technique initialement proposée par Douglas et Peucker (Douglas *et al.*, 1973). Différentes variantes de cet algorithme sont comparées dans les travaux de (Wu *et al.*, 2007). Dans le cadre de cette étude, l'algorithme de Douglas et Peucker spatio-temporel (Bertrand *et al.*, 2007) est utilisé. L'objectif de ce filtrage est de ne conserver que les positions significatives représentant la trajectoire à savoir des changements de vitesse ou de cap. Pour effectuer cette simplification, la plus grande distance  $d_{max}$  entre les positions de la trajectoire

$P_i$  et leurs projections spatio-temporelles  $P'_i$  sur la droite reliant les positions de départ  $P_d$  et d'arrivée  $P_a$  est calculée. Pour illustrer l'algorithme 2, la figure 5 représente une trajectoire dont les positions  $P_d$  et  $P_a$  sont respectivement les positions de départ et d'arrivée. Le point  $P'_i$  correspond à l'interpolation de la position  $P_i$  sur la droite  $[P_dP_a]$  au temps  $t_{P_i}$  en utilisant la vitesse  $V_d$  de la position de départ  $P_d$ . Si la distance  $d$  entre  $P_i$  et  $P'_i$  est supérieure à un seuil fixé, la position la plus éloignée est conservée. La trajectoire est alors divisée en deux à cette position et l'algorithme est appliqué de façon récursive aux deux sous-parties. Si la distance  $d$  est inférieure au seuil fixé, seules les positions  $P_d$  et  $P_a$  sont conservées. L'algorithme 2 présente en détail les différentes étapes de cette simplification de trajectoire. Celui-ci permet également de supprimer les erreurs locales dues par exemple à des imprécisions des appareils de mesure (Bertrand *et al.*, 2007).

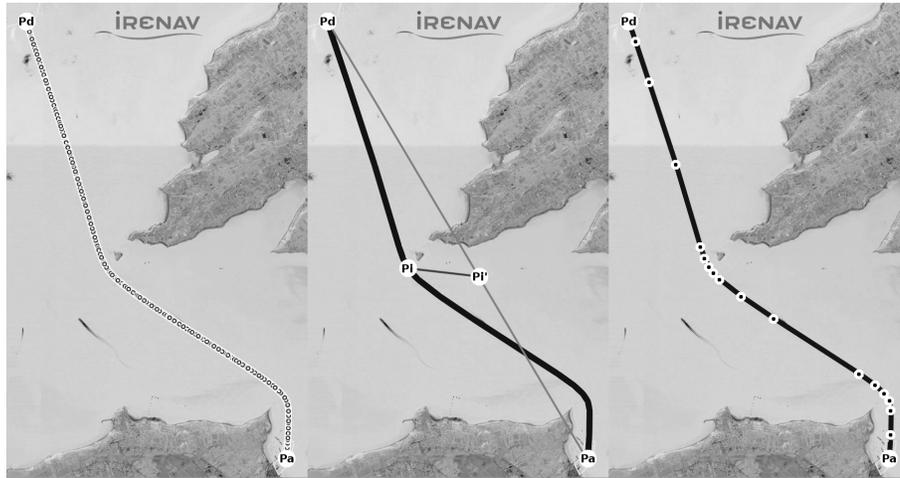
---

**Algorithme 2** Douglas et Peucker spatiotemporel
 

---

- 1: **Algorithme** Douglas\_Peucker\_ST(Trajectoire  $Tr$ )
  - 2: Soit  $[P_dP_a]$  la droite reliant les positions de départ  $P_d$  et d'arrivée  $P_a$  de  $Tr$
  - 3: Soit  $V_d$  la vitesse de l'objet mobile en  $P_d$
  - 4: **pour** chaque Position  $P_i$  de  $Tr$  **faire**
  - 5:   Calculer la position  $P'_i$  interpolée sur la droite  $[P_dP_a]$  à l'instant  $t_{P_i}$  à partir du point  $P_d$  à la vitesse  $V_d$
  - 6:   Calculer la distance  $d$  entre  $P_i$  et  $P'_i$
  - 7:   **si**  $d > d_{max}$  **alors**
  - 8:      $d_{max} = d$
  - 9:      $P_{max} = P_i$
  - 10:   **fin si**
  - 11: **fin pour**
  - 12: **si**  $d_{max} > seuil$  **alors**
  - 13:   Scinder  $Tr$  à la position  $P_{max}$  en deux sous-trajectoires  $Tr_1$  et  $Tr_2$
  - 14:   Concaténer *Douglas\_Peucker\_ST*( $Tr_1$ ) et *Douglas\_Peucker\_ST*( $Tr_2$ )
  - 15: **sinon**
  - 16:   Ne conserver que les positions  $P_d$  et  $P_a$  de la trajectoire  $Tr$
  - 17: **fin si**
- 

Dans le cadre de notre exemple, l'application du filtrage de Douglas et Peucker spatio-temporel au  $GHT_{AIT}$  a permis de réduire le nombre nécessaire de positions pour représenter efficacement les trajectoires de 104 201 à 16 110 positions soit un taux de compression de 84,54 % pour un seuil de 10 m (correspondant à la précision d'un GPS). Le temps CPU nécessaire à l'exécution des étapes de filtrage du  $GHT_{AIT}$  appliqué à notre exemple est de 934 ms.



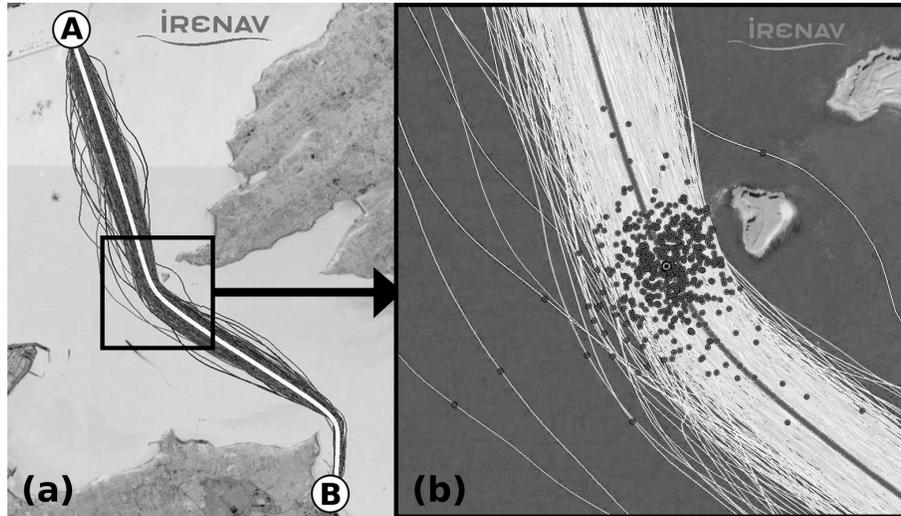
**Figure 5.** *Algorithme de Douglas et Peucker spatio-temporel*

#### 4. Analyse statistique quantitative des trajectoires

##### 4.1. Calcul de la route type

Une fois le  $GHT_{AIT}$  obtenu et filtré, le calcul de la route type suivie par les trajectoires peut être effectué. L'objectif de cette étape est d'obtenir par analyse statistique spatiale une route type d'un itinéraire pour les navires d'un même type. Afin d'éviter les distorsions spatiales introduites par des vitesses légèrement différentes des trajectoires composant le  $GHT_{AIT}$ , il est nécessaire de normaliser temporellement chaque trajectoire du  $GHT_{AIT}$ .

Pour réaliser cette normalisation, trois estampilles temporelles sont affectées à chaque position d'une trajectoire. Les positions initiales des trajectoires disposent toutes d'une estampille temporelle représentant le temps absolu ( $t_A$ ) au format UTC auquel l'objet mobile était à cette position. Connaissant le temps absolu de départ de la trajectoire, il est possible d'affecter à chaque position un temps relatif ( $t_R$ ) représentant la durée écoulée depuis la position de départ de la trajectoire. Enfin, un temps relatif normalisé ( $t_{RN}$ ) est affecté à chaque position du  $GHT_{AIT}$  par le calcul suivant. En premier lieu, la durée médiane  $D_{med}$  du  $GHT_{AIT}$  est calculée. Puis, une fonction de normalisation temporelle est appliquée à toutes les trajectoires du  $GHT_{AIT}$  afin que chaque trajectoire commence à un temps  $t_0$  et finisse au même temps normalisé  $t_{med} = t_0 + D_{med}$ . Au vu des données exploitées, le choix de la durée médiane du  $GHT_{AIT}$  permet d'obtenir une route type moins sujette aux perturbations dues aux trajectoires ayant un comportement inhabituel. Le temps CPU nécessaire à l'exécution de la normalisation temporelle du  $GHT_{AIT}$  appliqué à notre exemple est de 33 ms.



**Figure 6.** Calcul de la route type, nuages de positions au même temps normalisé

Une fois le  $GHT_{AIT}$  temporellement normalisé, la route type est élaborée. Pour chaque position de chaque trajectoire du  $GHT_{AIT}$ , les positions interpolées des autres trajectoires du  $GHT_{AIT}$  sont calculées. Cette seconde étape du calcul de la route type génère des ensembles de positions interpolées à différents temps normalisés. Ces ensembles de positions composent des nuages de positions.

En sélectionnant les valeurs médianes des latitudes et longitudes des différents nuages de positions, des positions médianes pour chaque temps normalisé sont obtenues. Ces positions médianes ordonnées à l'aide de leur estampille temporelle composent la trajectoire représentant la route type de l'itinéraire. Finalement, un filtrage de la route type est réalisé afin de ne conserver que les positions significatives. L'utilisation des positions médianes pour la génération de la route type permet d'éviter un biais pouvant être lié à des trajectoires anormales ayant un écart très important (contrairement à une approche basée sur la moyenne). Une fois calculée, cette route type représente le comportement médian des objets mobiles d'un même type suivant le même arc d'un itinéraire. Il est ainsi possible de comparer une trajectoire d'objet mobile (de même type et itinéraire) avec la route type et d'en déduire des similitudes ou différences de position, orientation, vitesse sur tout ou partie de cette trajectoire. Le temps CPU nécessaire à la génération de la route type appliqué à notre exemple est de 14 462 ms.

L'algorithme 3 synthétise cette phase d'élaboration de la route type. Dans l'exemple illustré par la figure 6a, la route type représentée en blanc est calculée à partir d'un  $GHT_{AIT}$  composé des 506 trajectoires tracées en noir. Le nuage de points

de la figure 6b représente l'ensemble des positions ayant un même temps normalisé, le point cerclé de noir correspond à la position médiane de l'ensemble des positions.

---

**Algorithme 3** Calcul de la route type
 

---

**Préconditions :**

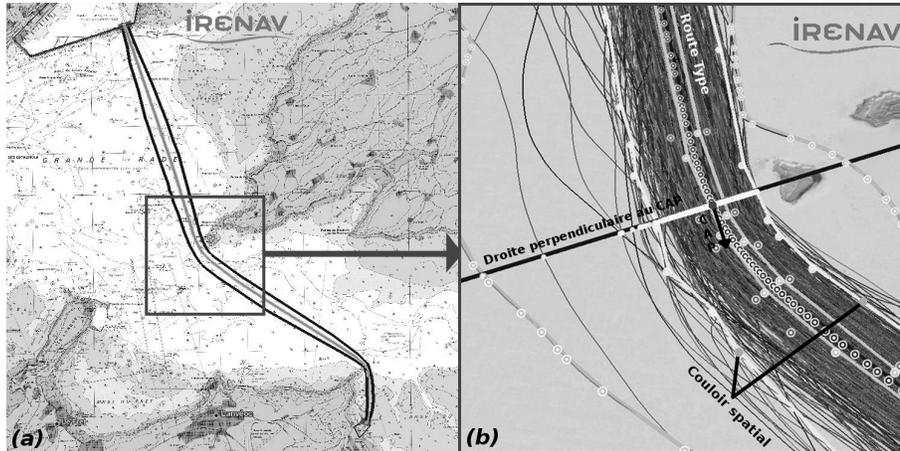
```

1: pour chaque trajectoire  $T_r$  du  $GHT_{AIT}$  faire
2:   Suppression des trajectoires erronées
3:   Recalage spatial des trajectoires
4:   Douglas_Peucker_ST(Trajectoire  $T_r$ )
5:   Normalisation temporelle des trajectoires à une durée médiane  $t_{med}$ 
6: fin pour
7:
8: Algorithme Calcul_Route_Type( $GHT_{AIT}$ )
9: pour chaque trajectoire  $T_{rm}$  du  $GHT_{AIT}$  faire
10:  pour chaque position  $P_{im}$  de  $T_{rm}$  faire
11:    Soit  $t_{ni}$  le temps normalisé de la position  $P_{im}$ 
12:    pour toutes les autres trajectoires  $T_{rn}$  du  $GHT_{AIT}$  faire
13:      Interpoler l'ensemble des positions  $EP_{jn}$  à  $t_{ni}$ 
14:      Calculer la position médiane  $P_{med}$  de  $EP_{jn}$  à  $t_{ni}$ 
15:      Ajouter  $P_{med}$  d'estampille temporelle  $t_{ni}$  à la route type  $RT$ 
16:    fin pour
17:  fin pour
18: fin pour
19: return Douglas_Peucker_ST(Trajectoire  $RT$ )
  
```

---

#### 4.2. Calcul du couloir spatial

L'observation de la figure 6 permet de constater que les objets mobiles d'un même type effectuant un même itinéraire se déplacent globalement sur une route type. L'espace dans lequel ces objets évoluent étant ouvert, il est normal de constater de légers écarts spatiaux. Ceux-ci doivent être distingués d'écarts plus importants considérés comme inhabituels. Ces écarts inhabituels représentent un certain pourcentage du  $GHT_{AIT}$ . Afin de séparer les comportements normaux des situations inhabituelles, un couloir spatial est calculé à l'aide d'une analyse statistique des positions des trajectoires du  $GHT_{AIT}$  vis-à-vis de la route type. Pour ordonner par distance les positions des trajectoires du  $GHT_{AIT}$  par rapport à la route type, la droite perpendiculaire au cap ( $DPC$ ) de chaque position de la route type est utilisée. Les points d'intersection entre la  $DPC$  et chaque trajectoire du  $GHT_{AIT}$  sont classés de part et d'autre de la route type par ordre de distance. Les trajectoires composées des positions correspondant au neuvième décile de chaque côté de la route type permet d'obtenir le couloir spatial de navigation. Le choix de ce décile statistique permet d'obtenir un couloir au sein duquel la majorité des objets mobiles suivant cet itinéraire circulent. De plus, celui-ci conditionne le nombre de trajectoires considérées comme étant anormales. Ce paramètre pouvant être adapté en fonction de la sensibilité et du volume



**Figure 7.** Calcul du couloir spatial

de données disponibles. De même que pour la route type, l'utilisation d'une approche médiane/décile plutôt que moyenne/variance permet d'éviter que la génération du couloir spatio-temporel soit perturbée par des trajectoires ayant un écart à la route type très important. Le temps CPU nécessaire à la génération du couloir spatial appliqué à notre exemple est de 1 887 ms.

---

**Algorithme 4** Calcul du couloir spatial

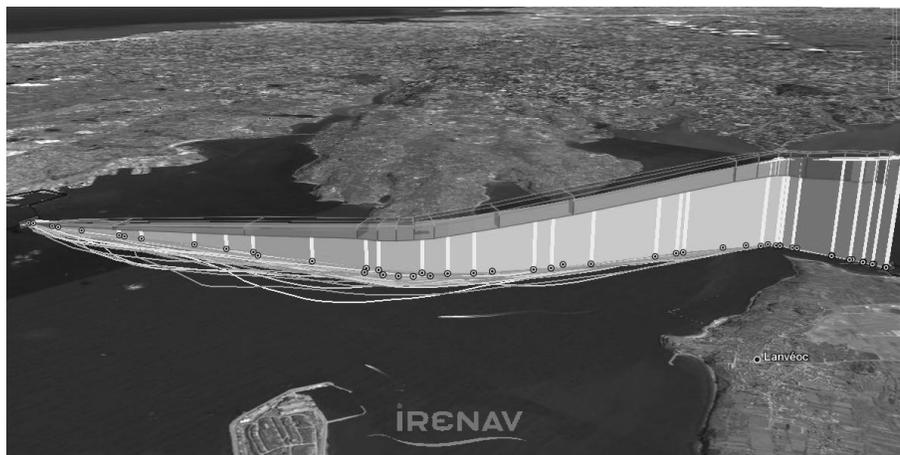
---

- 1: **Algorithme** Calcul\_Couloir\_Spatial( $GHT_{AIT}$ )
  - 2: **pour** chaque position  $P_i$  de la route type  $RT$  **faire**
  - 3:   Calculer la droite  $[DP]$  perpendiculaire au cap en  $P_i$
  - 4:   **pour** chaque trajectoire  $Tr_j$  du  $GHT_{AIT}$  **faire**
  - 5:     Calculer le point d'intersection  $P'_i$  entre  $Tr_j$  et  $[DP]$
  - 6:     **si**  $P'_i$  est à droite de  $P_i$  **alors**
  - 7:       Mémoriser  $P'_i$  dans le tableau  $A_{droite}$
  - 8:     **sinon**
  - 9:       Mémoriser  $P'_i$  dans le tableau  $A_{gauche}$
  - 10:    **fin si**
  - 11:   **fin pour**
  - 12:   Classer les tableaux  $A_{droite}$  et  $A_{gauche}$  par ordre de distance à  $P_i$
  - 13:    $Pfi_{droite}$  = neuvième décile du tableau  $A_{droite}$
  - 14:    $Pfi_{gauche}$  = neuvième décile du tableau  $A_{gauche}$
  - 15:   Ajouter  $Pfi_{droite}$  à la trajectoire frontière  $Tr_{droite}$  du couloir spatial
  - 16:   Ajouter  $Pfi_{gauche}$  à la trajectoire frontière  $Tr_{gauche}$  du couloir spatial
  - 17: **fin pour**
-

La figure 7 illustre le calcul de ce couloir pour notre exemple. Ainsi, il est possible d'identifier pour chaque  $GHT_{AIT}$  les positions inhabituelles. Il faut noter que la distance entre la route type et le couloir de droite et de gauche n'est pas identique. Effectivement, dans un espace ouvert, il est plus facile pour un objet mobile de dévier vers l'extérieur lorsqu'il tourne. De même, la largeur de ce couloir fournit une indication concernant l'écartement des trajectoires vis-à-vis de la route type. Dans notre exemple, comme le montre la figure 7, cette largeur est plus étroite au départ, à l'arrivée et dans les virages. Par contre, dans les lignes droites, celle-ci a tendance à augmenter. L'algorithme 4 synthétise les différentes étapes de calcul du couloir spatial.

#### 4.3. Calcul du couloir temporel

Sachant désormais si un objet mobile se situe dans le couloir spatial de l'itinéraire qu'il est censé suivre, il est intéressant de calculer si celui-ci est dans les temps par rapport à la route type. De manière similaire au couloir spatial, il est possible de calculer un couloir temporel permettant d'analyser le retard ou l'avance d'un objet mobile. Pour cela, lors du calcul du couloir spatial, les temps relatifs (durée depuis le départ) des positions d'intersection avec les droites perpendiculaires au cap de la route type sont prises en compte. Les temps relatifs de chacune de ces positions sont ensuite classés par ordre croissant. La sélection du neuvième décile pour les positions en avance et en retard par rapport au temps normalisé de la position de la route type permet d'obtenir des bornes temporelles de notre couloir pour chaque position de la route type. Le temps CPU nécessaire à la génération du couloir spatial appliqué à notre exemple est de 1 889 ms. Il faut noter que les écarts temporels entre une position de la route type et les bornes associées du couloir temporel ne sont pas identiques. Effectivement, il est plus facile pour un objet mobile de prendre du retard que de l'avance.



**Figure 8.** Couloir spatio-temporel

La combinaison du couloir spatial et temporel génère une suite de quadrilatères irréguliers dans un espace 3D dont l'axe z représente le temps. La liaison de ces quadrilatères génère un couloir spatio-temporel représenté pour notre cas d'étude sur la figure 8 par un polyèdre allongé de couleur foncée. Visuellement, ce couloir spatio-temporel est utilisable pour détecter des trajectoires d'objets mobiles inhabituelles n'étant pas contenues dans celui-ci. L'algorithme 5 résume les différentes étapes du calcul du couloir temporel.

---

**Algorithme 5** Calcul du couloir temporel
 

---

```

1: Algorithme Calcul_Couloir_Temporel( $GHT_{AIT}$ )
2: pour chaque position  $P_i$  de la route type  $RT$  faire
3:   Calculer la droite  $[DP]$  perpendiculaire au cap en  $P_i$ 
4:   pour chaque trajectoire  $Tr_j$  du  $GHT_{AIT}$  faire
5:     Calculer le point d'intersection  $P'_i$  entre  $Tr_j$  et  $[DP]$ 
6:     si  $P'_i$  se trouve dans le couloir spatial alors
7:       Soit  $tr_{P_i}$  le temps relatif en  $P_i$ 
8:       Soit  $tr'_{P_i}$  le temps relatif en  $P'_i$ 
9:       si  $tr'_{P_i}$  est en avance par rapport à  $tr_{P_i}$  alors
10:        Mémoriser  $P'_i$  dans le tableau  $A_{avance}$ 
11:      sinon
12:        Mémoriser  $P'_i$  dans le tableau  $A_{retard}$ 
13:      fin si
14:    fin si
15:  fin pour
16:  Classer les tableaux  $A_{avance}$  et  $A_{retard}$  par ordre croissant de temps relatif
17:   $trfi_{droite}$  = neuvième décile du tableau  $A_{avance}$ 
18:   $trfi_{gauche}$  = neuvième décile du tableau  $A_{retard}$ 
19:  Mémoriser  $trfi_{avance}$  et  $trfi_{retard}$  pour chaque  $P_i$ 
20: fin pour

```

---

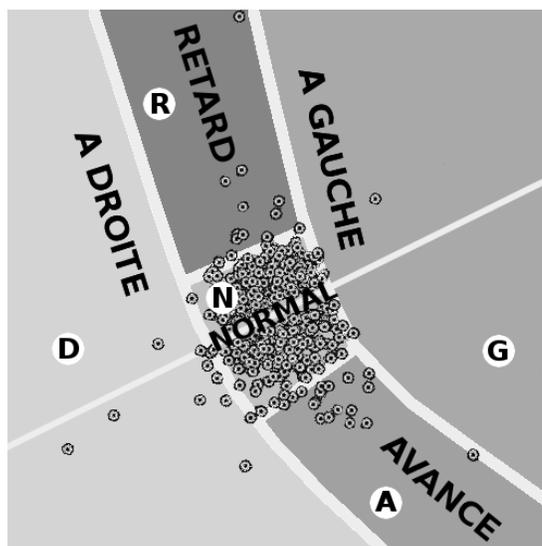
Pour chaque itinéraire et chaque type d'objet mobile, les couloirs de navigation spatio-temporels sont sauvegardés dans une base de connaissances (figure 1.5) afin d'être réutilisés pour une analyse temps réel. Le temps CPU global nécessaire à la génération du couloir spatio-temporel appliqué à notre exemple est de 36 511 ms. Ce temps étant fortement dépendant du nombre de positions conservées pour chaque trajectoire du  $GHT_{AIT}$  ainsi que pour la route type après filtrage de Douglas et Peucker spatio-temporel.

## 5. Analyse qualitative

La base de connaissances ayant été renseignée, il est désormais possible de l'employer pour qualifier les positions sauvegardées dans la  $BDST$  ainsi que celles obtenues en temps réel.

### 5.1. Analyse qualitative des positions sauvegardées

Pour chaque position d'une trajectoire d'un  $GHT_{AIT}$ , il est possible de la comparer à la position de la route type ayant le même temps relatif ainsi qu'au couloir spatio-temporel préalablement calculé pour ce  $GHT_{AIT}$ .



**Figure 9.** Couloir spatio-temporel à un temps relatif défini

Comme indiqué sur la figure 9, une position comparée au couloir spatial peut être considérée comme : « sur la route », « à droite » ou « à gauche » de la route type en fonction du couloir spatial. De même, si la position est sur la route, la comparaison de la position avec le couloir spatio-temporel permet de qualifier cette position comme étant « dans les temps », « en avance » ou « en retard ». Ces 5 zones spatio-temporelles permettent de qualifier la position d'une trajectoire à un temps relatif. Cependant, il n'est pas possible d'indiquer si une position qui ne se trouve pas dans le couloir spatial est en avance ou en retard, en effet l'objet mobile peut en sortant de la route, soit prendre un raccourci ou bien faire un détour. Ainsi, dans l'exemple représenté sur la figure 9, la position (*G*) associée à l'itinéraire Brest → Lanvéoc pour un navire de type « navire à passagers » est « à gauche » du couloir. Cependant, il n'est pas possible de qualifier temporellement cette position pour indiquer si elle se trouve en avance ou en retard. Ainsi, seule la position (*N*) est considérée comme étant normale. Les autres positions (*A*, *D*, *G*, *R*) sont considérées comme anormales et sont respectivement qualifiées comme étant « en avance », « à droite », « à gauche » et « en retard ».

## 5.2. Analyse qualitative temps réel

La figure 10 reprend de façon générale le processus de génération des routes types et des couloirs spatio-temporels pour chaque  $GHT_{AIT}$ . A partir de données de positions obtenues en temps réel ainsi que de l'historique des positions précédentes sauvegardées dans la  $BDST$  (figure 10 étape 2), la plateforme d'analyse temps réel (figure 10 étape 6) doit appairer une position à un itinéraire dont le couloir spatio-temporel a été précédemment calculé (figure 10 étape 4) et sauvegardé dans une base de connaissances (figure 10 étape 5).

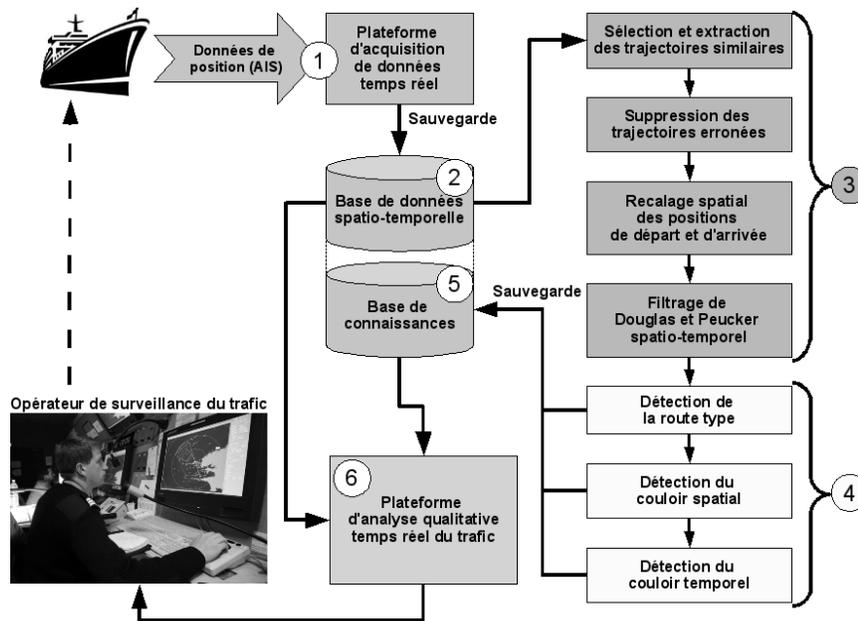


Figure 10. Schéma général détaillé

Cet appariement s'effectue en fonction de la position actuelle de l'objet mobile, de l'historique de sa trajectoire ainsi que des informations concernant sa destination prévue. Une fois l'itinéraire sélectionné, les dernières positions connues du navire sont comparées au couloir spatio-temporel en fonction du temps relatif du navire depuis le départ de l'itinéraire et de leur position géographique par rapport à la route type et au couloir spatio-temporel. Ces informations qualitatives concernant la trajectoire du navire sont fournies à l'opérateur de surveillance du trafic. Ainsi, pour une position considérée par le système comme anormale, l'opérateur peut se focaliser sur cette trajectoire parmi un ensemble important d'autres trajectoires de navires. De plus, ce module d'analyse est capable, une fois l'itinéraire apparié, d'indiquer la destination du navire ainsi que le temps nécessaire pour arriver à destination par comparaison avec les données de la route type.

## 6. Conclusion

Cette étude s'est intéressée à la problématique particulière du déplacement d'objets mobiles dans un espace ouvert. Elle a été appliquée plus particulièrement au domaine maritime comme exposé dans notre exemple basé sur un jeu conséquent de données réelles. Après avoir défini les notions d'itinéraires et de trajectoires suivant un itinéraire, une plateforme complète d'analyse qualitative temps réel de trajectoire a été proposée. Une fouille de données spatio-temporelles a permis d'extraire des groupes homogènes de trajectoires d'objets mobiles de même type suivant un même itinéraire. Pour chaque groupe homogène de trajectoires d'objets mobiles de même type suivant le même arc d'un itinéraire ( $GHT_{AIT}$ ), une route type a été générée à partir d'une analyse statistique. Des couloirs spatio-temporels ont également été calculés et mémorisés dans une base de connaissances.

Ainsi, l'étude statistique d'un  $GHT_{AIT}$  permet d'obtenir des informations concernant les comportements de ces objets. Grâce aux couloirs spatio-temporels, un premier niveau d'analyse qualifie simplement des trajectoires sur des critères spatiaux et temporels. Un prototype fonctionnel permettant de qualifier une trajectoire suivant un itinéraire a été implémenté et appliqué sur un jeu de données conséquent. La section traitant de l'appariement en temps réel d'une trajectoire à un itinéraire a été modélisée mais n'a cependant pas encore été implémentée dans ce prototype. Il serait désormais intéressant de valider celui-ci en le mettant à disposition d'opérateurs de surveillance du trafic maritime basés dans les CROSS (Centres régionaux opérationnels de surveillance et de sauvetage). Ces opérateurs peuvent avoir à surveiller jusqu'à plus de 250 navires affichés simultanément sur leurs écrans de contrôle. Cet outil couplé à un système de visualisation adapté pourrait leur permettre de se focaliser sur les trajectoires de navires ayant un comportement considéré par le système comme anormal et ainsi diminuer leur charge cognitive. De plus, une extension possible du système pourrait être la sélection ou la suppression manuelle par l'opérateur des trajectoires considérées comme anormales ce qui permettrait un recalcul des outils de qualification des trajectoires par apprentissage.

L'originalité de cette démarche d'analyse de similitude réside dans l'introduction de métaconnaissances ( $GHT_{AIT}$ , routes types, couloirs spatio-temporels) permettant de qualifier les comportements d'objets mobiles suivant un itinéraire dans un espace ouvert. Cependant, l'appariement en temps réel d'une position à un itinéraire reste une problématique à résoudre. Certains arcs d'un itinéraire pouvant être partagés avec d'autres, l'appariement d'une position en temps réel à un itinéraire est rendue difficile. Le recoupement d'informations telles que l'historique de la trajectoire, l'identifiant de l'objet mobile, la destination prévue, facilitent l'appariement à un itinéraire. Néanmoins, à chaque nouvelle position de l'objet mobile, cet appariement est susceptible d'être modifié. Par analogie avec le graphe de zone, il serait intéressant de découper les trajectoires en une suite de tronçons afin de faciliter l'analyse du comportement d'un navire sur une sous-partie d'une trajectoire ayant des caractéristiques communes (vitesse, cap, taux de rotation...). L'algorithme de Douglas et Peucker spatio-temporel conservant les positions remarquables d'une trajectoire, celles-ci pourraient être uti-

lisées afin de définir les jonctions entre deux tronçons. Chaque position d'un tronçon ayant des propriétés communes, l'analyse et la détection de données anormales sur une période en seraient facilitées (couper un virage, louvoyer).

Qui plus est, afin d'optimiser les temps de calculs, les données de positions de la *B DST* pourraient également être filtrées à l'aide du même algorithme. En effet, 50,04 % du temps CPU nécessaire à la génération des patrons spatio-temporels est lié à l'extraction et au filtrage du *GHT<sub>AIT</sub>*. Finalement, cette analyse utilise un critère de sélection basé principalement sur le type de l'objet mobile et ne prend pas en compte les spécificités dues à l'environnement dans lequel il évolue (les saisons et les marées dans le cadre de notre exemple). Il serait intéressant d'appliquer d'autres critères de sélection du *GHT<sub>AIT</sub>* ainsi qu'enrichir cette analyse en la couplant avec des travaux portant sur la prise en compte de l'environnement (Richard, 2007), des statistiques d'accidents (Degré, 2007) et des comportements illicites (Morel *et al.*, 2008).

## 7. Bibliographie

- Aggarwal C. C., Yu P. S., « Outlier detection for high dimensional data », *SIGMOD*, vol. 30, n° 2, p. 37-46, 2001.
- Barnett V., Lewis T., *Outliers in Statistical Data*, John Wiley & Sons New York, 1994.
- Baud O., El-Bied Y., Honore N., Taupin O., « Trajectory Comparison for Civil Aircraft », *Aerospace Conference, 2007 IEEE*, p. 1-9, 2007.
- Bertrand F., Bouju A., Claramunt C., Devogele T., Ray C., *Web and Wireless Geographical Information Systems*, vol. 4857 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, chapter Web Architecture for Monitoring and Visualizing Mobile Objects in Maritime Contexts, p. 94-105, 2007.
- Breunig M., Kriegel H., Ng R., Sander J., « LOF : identifying density-based local outliers », *ACM SIGMOD Record*, vol. 29, n° 2, p. 93-104, 2000.
- Degré T., « The use of risk concept to characterize and select High Risk Vessels for ship inspections », *12th International Congress of the International Maritime Association of the Mediterranean*, vol. 6, p. 37-49, 2007.
- Douglas D. H., Peucker T. K., « Algorithms for the reduction of the number of points required to represent a digitized line or its caricature », *Cartographica : The International Journal for Geographic Information and Geovisualization*, vol. 10, p. 112-122, 1973.
- Gaffney S., Smyth P., « Trajectory clustering with mixtures of regression models », *KDD '99 : Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, New York, NY, USA, p. 63-72, 1999.
- Güting R. H., « An Introduction to Spatial Database Systems », *VLDB Journal*, vol. 3, p. 357-399, 1994.
- IMO, Development of an E-Navigation strategy, Technical report, International Maritime Organization, 2007.
- Kharrat A., Popa I. S., Zeitouni K., Faiz S., *Clustering Algorithm for Network Constraint Trajectories*, Springer Berlin Heidelberg, chapter Clustering Algorithm for Network Constraint Trajectories, p. 631-647, 2008.

- Knorr E. M., Ng R. T., « Algorithms for Mining Distance-Based Outliers in Large Datasets », *VLDB '98 : Proceedings of the 24rd International Conference on Very Large Data Bases*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, p. 392-403, 1998.
- Knorr E. M., Ng R. T., « Finding Intensional Knowledge of Distance-Based Outliers », *VLDB '99 : Proceedings of the 25th International Conference on Very Large Data Bases*, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, p. 211-222, 1999.
- Knorr E. M., Ng R. T., Tucakov V., « Distance-based outliers : algorithms and applications », *The VLDB Journal*, vol. 8, n° 3-4, p. 237-253, 2000.
- Lee J., Han J., Li X., « Trajectory Outlier Detection : A Partition-and-Detect Framework », *Data Engineering, 2008. ICDE 2008. IEEE 24th International Conference on Data Engineering*, p. 140-149, 2008.
- Mefteh W., Bouju A., Malki J., « Vers un modèle générique des trajectoires », *Manifestation des Jeunes Chercheurs en Sciences et Technologies de l'Information et de la Communication*, 2008.
- Morel M., Napoli A., Littaye A., George J., Jangal F., Alhadeb B., Scapel C., Lebrevelec J., Dejardin D., « Surveillance et contrôle des activités des navires en mer ScanMaris », *Workshop Interdisciplinaire sur la Sécurité Globale*, 2008.
- Mouza C. D., Rigaux P., *Bases de données spatio-temporelles*, Documents et évolution, 2000.
- Papadimitriou S., Kitagawa H., Gibbons P., Faloutsos C., « LOCI : fast outlier detection using the local correlation integral », *Data Engineering, 2003. Proceedings. 19th International Conference on Data Engineering*, p. 315-326, 2003.
- Ramaswamy S., Rastogi R., Shim K., « Efficient algorithms for mining outliers from large data sets », *SIGMOD '00 : Proceedings of the 2000 ACM SIGMOD international conference on Management of data*, ACM, New York, NY, USA, p. 427-438, 2000.
- Richard A., « L'optimisation des transits de navires marchands par l'utilisation des courants de marée », *Canadian Journal of Transportation*, 2007.
- Shim C.-B., Chang J.-W., *Advances in Databases and Information Systems*, Springer, chapter Similar Sub-trajectory Retrieval for Moving Objects in Spatio-temporal Databases, p. 308-322, 2003.
- Wan T., Zeitouni K., Meng X., « An OLAP system for network-constrained moving objects », *SAC '07 : Proceedings of the 2007 ACM symposium on Applied computing*, ACM, New York, NY, USA, p. 13-18, 2007.
- Wang X., Tieu K., Grimson E., « Learning Semantic Scene Models by Trajectory Analysis », *Computer Vision – ECCV 2006*, vol. 3953 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, p. 110-123, 2006.
- Wu Y., Pelot R., *Geomatics Solutions for Disaster Management*, Jonathan Li and Sisi Zlatanova and Andrea G. Fabbri, chapter Comparison of Simplifying Line Algorithms for Recreational Boating Trajectory Dedensification, p. 321-334, 2007.

**ANNEXE POUR LE SERVICE FABRICATION**  
A FOURNIR PAR LES AUTEURS AVEC UN EXEMPLAIRE PAPIER  
DE LEUR ARTICLE ET LE COPYRIGHT SIGNE PAR COURRIER  
LE FICHER PDF CORRESPONDANT SERA ENVOYE PAR E-MAIL

1. ARTICLE POUR LA REVUE :  
*RSTI - ISI – 14/2009. SI et géolocalisation*
2. AUTEURS :  
*Laurent ETIENNE\* — Thomas DEVOGELE\* — Alain BOUJU\*\**
3. TITRE DE L'ARTICLE :  
*Analyse de similarité de trajectoires d'objets mobiles suivant le même itinéraire*
4. TITRE ABRÉGÉ POUR LE HAUT DE PAGE MOINS DE 40 SIGNES :  
*Analyse de similarité de trajectoires*
5. DATE DE CETTE VERSION :  
*27 août 2009*
6. COORDONNÉES DES AUTEURS :
  - adresse postale :
  - \* Institut de recherche de l'École Navale (IRENAV)  
Lanvéoc-Poulmic, CC 600, 29240 BREST Cedex 9  
prenom.nom@ecole-navale.fr
  - \*\* Laboratoire d'informatique et d'imagerie industrielle  
17042 La Rochelle Cedex 1  
abouju@univ-lr.fr
  - téléphone : 02 98 23 40 21
  - télécopie : 02 98 23 38 57
  - e-mail : laurent.etienne@ecole-navale.fr
7. LOGICIEL UTILISÉ POUR LA PRÉPARATION DE CET ARTICLE :  
L<sup>A</sup>T<sub>E</sub>X, avec le fichier de style article-hermes.cls,  
version 1.23 du 17/11/2005.
8. FORMULAIRE DE COPYRIGHT :  
Retourner le formulaire de copyright signé par les auteurs, téléchargé sur :  
<http://www.revuesonline.com>

SERVICE ÉDITORIAL – HERMES-LAVOISIER  
14 rue de Provigny, F-94236 Cachan cedex  
Tél. : 01-47-40-67-67  
E-mail : [revues@lavoisier.fr](mailto:revues@lavoisier.fr)  
Serveur web : <http://www.revuesonline.com>