Démarche d'analyse de l'intégrité d'un système de localisation de navires

Clément Iphar, Aldo Napoli, Cyril Ray

Démarche d'analyse de l'intégrité d'un système de localisation de navires

Clément IPHAR[1], Aldo NAPOLI[1], Cyril RAY[2]

1. CRC, MINES ParisTech, PSL Research University
Rue Claude Daunesse, F-06560 SOPHIA ANTIPOLIS, FRANCE
{clement.iphar ; aldo.napoli}@mines-paristech.fr

2. Institut de Recherche de l'Ecole Navale (IRENav)
BRCM Brest, F-29240 BREST, FRANCE
cyril.ray@ecole-navale.fr

RESUME. Le système d'identification automatique (AIS) est un système électronique permettant aux navires d'envoyer et de recevoir des messages de positionnement. Ces messages sont utilisés à des fins diverses, telles que la prévention des abordages ou la surveillance du trafic. Cependant les messages envoyés via ce système contiennent des erreurs et subissent des falsifications. Cet article propose une méthode d'analyse des données transmises via ce système, méthode basée sur l'étude de l'intégrité des données contenues au sein même des messages, et ce à plusieurs niveaux de profondeur dans un message et entre les messages, le but étant d'attribuer à chaque message un coefficient de confiance relatif à aux déterminations de l'intégrité des données en son sein.

ABSTRACT. The Automatic Identification System (AIS) is an electronic system which allows the vessels to send and receive positioning messages. Those messages are used for several purposes such as boarding prevention or traffic monitoring, amongst others. However, the messages are subject to errors and they undergo falsifications. This paper proposes a methodology for the analysis of data broadcasted via this system, based on the integrity of data contained within the messages themselves, this being done at several levels of depth in one lone message and between different messages, the purpose being to assign to each message a coefficient revealing the degree of confidence computed by the proposed integrity-based method.

MOTS-CLES : AIS ; falsification de données ; intégrité des données ; détection d'anomalies.

KEYWORDS: AIS ; data falsification ; data integrity ; anomaly detection.

1. Introduction

Carrefour des enjeux internationaux, l'espace maritime est soumis à une activité de plus en plus intense. Cette croissance des mobilités maritimes a conduit à la généralisation de systèmes de suivi de navigation. Parmi ces systèmes, on peut citer les radars marins, des satellites optiques ou radars à synthèse d'ouverture. L'un de ces systèmes est le Système d'Identification Automatique (AIS).

Des mécanismes de malversation et des pratiques de navigation sont inévitablement apparues récemment pour contourner, falsifier, exploiter de tels systèmes de surveillance dans l'intérêt des contrevenants. C'est la découverte de ces malversations, notamment au travers de l'étude des données et de leur intégrité au sein des messages AIS qui constitue le cœur de notre démarche de recherche.

2. Principes et limite du système d'identification automatique

2.1. Un système pour améliorer la sécurité de la navigation

L'utilisation du système AIS a été instaurée en 2000 par l'Organisation Maritime Internationale, pour "Tous les navires de jauge brute supérieure ou égale à 300 engagés dans des voyages internationaux, les navires de charge de jauge brute supérieure ou égale à 500 et les navires à passagers, quel que soit leur tonnage'' (IMO, 2004). Le système utilise des transpondeurs pour envoyer des messages par ondes VHF (Very High Frequency) aux navires alentours et aux stations côtières.

2.2. Les défauts du système

Trois principaux types d'actions menant à une perte de qualité des données peuvent être recensés : l'erreur, la falsification et le piratage. Les informations erronées transmises peuvent être fausses, incomplètes, impossibles selon la norme ou impossibles physiquement. La falsification est le fait de dégrader volontairement un message par la modification d'une valeur authentique ou l'arrêt volontaire et non motivé de l'envoi des messages, fait dans le but de tromper l'extérieur. D'après (Harati-Mokhtari et al., 2007), environ 50% des messages contiendraient des données erronées et 1% des navires diffusent des données falsifiées. Par ailleurs, le piratage de message est effectué par un acteur extérieur par la création de faux messages et leur transmission sur les fréquences AIS (Balduzzi et al., 2014).

3. Une méthode basée sur l'intégrité des données pour la découverte de falsifications

Notre recherche se concentre sur les problématiques basées sur le message, en portant un intérêt tout particulier à l'information présente au sein même de chacun

des messages AIS et à sa qualité. Ainsi, la notion d'intégrité de la donnée est utilisée à des fins d'analyse (Iphar et al., 2015).

Les messages AIS sont divisés en champs, et ces champs peuvent prendre différentes valeurs en accord avec les spécifications techniques du système. De plus, ces données peuvent être de types différents : numérique représentant un identifiant (Numéro OMI), numérique représentant une quantité ou une grandeur physique (Latitude), numérique représentant un choix dans une liste (statut de navigation), textuel (nom du navire), de type date (Date d'arrivée estimée) ou de type binaire.

3.1. Détermination de l'intégrité des messages AIS

En prenant en considération les données en provenance de tous les messages AIS, notre proposition de démarche se base sur quatre moyens de tester l'intégrité de ces données. La première voie consiste en le contrôle de l'intégrité interne à un message : la conformité aux spécifications techniques de chaque champ de chaque message est étudiée, ainsi que l'intégrité au sein des champs de ce seul message. La deuxième voie consiste en le contrôle d'intégrité de l'information entre plusieurs messages : les messages de même type ont les mêmes champs, et il est donc possible de comparer les valeurs de ces champs et de déterminer leur intégrité entre différents messages, par ailleurs la comparaison et la détermination de l'intégrité de la donnée AIS entre champs de différents messages est également possible, en effet, même si les données proviennent de différents messages, du fait de l'existence de champs similaires et liés dans différents messages, l'intégrité de certaines informations est vérifiable. La troisième voie est une méthode statistique pour évaluer la similarité d'un message avec les messages d'un jeu de données (en partant du principe que dans un groupe de messages, issus d'un même navire ou d'une même aire géographique, les valeurs de certains champs auront tendance à se ressembler). La quatrième voie consiste en l'utilisation de bases de données externes afin de recouper des informations (par exemple la base des données des navires de pêches de l'Union Européenne étant disponible librement en ligne, il est aisé de vérifier certaines informations envoyées par les navires, telles que nom ou longueur).

3.2. Cas d'études

Les méthodes de détermination d'intégrité et les méthodes statistiques vues précédemment nécessitent un jeu de données qui doit être sélectionné afin de pourvoir à une étude donnée, telles que proposées dans cette partie. Nous avons sélectionné deux études principales et deux études secondaires, qui sont basées pour les premières sur la station et le MMSI (numéro unique d'identifiant de navire), et pour les secondes sur une zone et la route. Les hypothèses proposées ici sont que des messages sélectionnés selon l'une de leurs caractéristiques (même station de réception, même émetteur) auront tendance à se ressembler, et donc que certaines anomalies, dans ce contexte, auront tendance à émerger.

Par leur nature, les études basées sur la station et sur le MMSI peuvent être effectuées aussi bien à la volée qu'avec des données archivées alors que les études basées sur une zone et la route ne peuvent s'effectuer que basées sur des données archivées à des fins d'analyse.

Afin de réaliser cette étude, nous avons développé un parser AIS opérationnel, adapté du parser AISMESSAGE disponible en ligne. Par ailleurs, l'architecture de base de données a été réalisée, construite avec une base de données relationnelle utilisant une architecture PostgreSQL/PostGIS. Le processus d'analyse en cours de développement utilise la base de données à l'aide de scripts Python qui extraient les informations issues des traitements présentés dans la sous-partie précédente.

4. Conclusion

Les vulnérabilités de l'AIS, qui conduisent à la présence d'erreurs et de falsifications au sein des messages que le système transmet aux autres navires et aux systèmes de surveillance, sont présentées dans ce papier. C'est l'intégrité en tant que dimension caractéristique de la qualité de la donnée que nous vérifions par une approche basée sur le message et plus particulièrement sur les informations contenues au sein des champs de données des messages. Les travaux en cours se concentrent sur le développement d'un système fondé sur une architecture de base de données, et l'usage de bases de données contextuelles (de type météorologiques, ou registres de pêche). Ces travaux pourront en outre se concentrer sur la définition d'alertes en fonction de l'analyse des mouvements, des risques potentiels engendrés et des éléments d'intégrité issus des analyses proposées dans cet article.

Remerciements :

Ces travaux sont financés par l'ANR sous la référence ANR-14-CE28-0028 dans le cadre du projet DéAIS, cofinancés par la DGA, et labélisés par le Pôle Mer Bretagne Atlantique et le Pôle Mer Méditerranée.

Références

Balduzzi M, Pasta A., Wilhoit K. (2014). A security evaluation of AIS automated identification system. *Actes de la 30th annual computer security applications conference 2014*, La Nouvelle-Orléans, USA.

Harati-Mokhtari A., Wall A., Brooks P., Wang J. (2007). Automatic Identification System (AIS): a human factors approach. *Journal of Navigation.* Vol 60(3).

IMO (2004). International Maritime Organization. *International convention for the safety of life at sea*.

Iphar C, Napoli A, Ray C. (2015). Detection of false AIS messages for the improvement of maritime situational awareness. *Actes de la conférence OCEANS'15 Washington,* USA