Méthodologie d'extraction de signatures issues des signaux AIS

Erwan Alincourt, Cyril Ray, Pierre-Michel Ricordel, Delphine Daré-Emzivat et Abdel Boudraa
erwan.alincourt@gmail.com
cyril.ray@ecole-navale.fr
pierre-michel.ricordel@ssi.gouv.fr
delphine.dare@ecole-navale.fr
abdel.boudraa@ecole-navale.fr

IRENav / ANSSI

Résumé. L'AIS (Automatic Identification System) est un système de balises de localisation pour bateaux particulièrement vulnérable à l'injection de fausses données. Chargé notamment de transmettre identité et position des navires entre eux, les conséquences matérielles et humaines d'une telle attaque peuvent être considérables. Avec l'avènement des outils de radio logicielle, la menace grandit et il y a un réel besoin d'outils en mesure de détecter ces malversations. Cet article cherche à déterminer s'il est possible de détecter ces attaques en analysant le signal AIS reçu.

1 Introduction

Depuis 1960, le trafic maritime ne cesse de croître. Ainsi, entre 1960 et 2000, il a été multiplié par cinq en volume, et depuis le début des années quatre-vingt-dix, son taux de croissance annuel est en augmentation. Cette forte augmentation du trafic engendre de nombreuses difficultés et en particulier des risques de fortune de mer de plus en plus importants. C'est dans ce contexte que l'AIS a été créé et rendu obligatoire par l'OMI pour les navires de commerce. Il s'agit d'un système automatisé de transpondeurs radio fonctionnant en VHF capables de transmettre régulièrement l'identité, la position, la route et la vitesse d'un navire aux navires qui lui sont proches et à des centres à terre.

Massivement déployé, il est également de plus en plus employé par les navires de plaisance. De plus, avec le développement de plates formes de collecte et de diffusion par Internet des informations AIS, de nouveaux besoins apparaissent. Il s'agit en particulier d'un besoin fort de discrétion, que ce soit pour des motivations légitimes (protection du secret commercial ou de la vie privée) ou illégitimes (dissimulation d'activités illégales).

2 Contexte

C'est dans ce contexte qu'il est possible d'observer de plus en plus de coupure des transpondeurs, voire de falsifications des données AIS [9]. La possibilité de créer et d'émettre facilement des trames arbitraires a été démontrée en 2013 [1]. Cette menace d'une falsification externe doit être prise au sérieux. Devant le délai nécessaire à l'évolution de la norme qui pourrait prendre en compte des mesures de sécurité robustes comme la cryptographie (l'AIS a mis 8 ans à être généralisé aux navires de commerce), plusieurs approches sont développées afin de détecter de telles attaques : l'analyse comportementale [6], le croisement des données AIS avec un radar [4], la triangulation du signal [7], la radiogoniométrie ou encore l'effet Doppler (uniquement sur certains satellites AIS). Chacune de ces méthodes a des inconvénients, en particulier la nécessité de disposer de plusieurs capteurs (plusieurs AIS ou AIS et radar) pour être efficaces.

Ainsi, pour être en mesure d'offrir une première capacité de détection avec un seul récepteur, ou compléter les méthodes précédentes, l'idée de cet article est de déterminer si l'exploitation d'informations relatives à la couche physique [2] peut être utilisée afin d'attribuer un niveau de confiance à l'information reçue.

Dans un premier temps, la cohérence entre la puissance de signal reçue et la distance entre l'émetteur et le récepteur donne une première information. Dans un second temps, il s'agit de rechercher des éléments d'identification de l'émetteur [8] et de les comparer avec une référence. En effet, le nombre de marques et modèles présents sur le marché et les défauts (même minimes) de fabrication, voire les différences d'installation et de positionnement de l'antenne [5] permettent de penser qu'il peut être possible d'identifier une famille de transpondeurs, voire un transpondeur particulier.

Ces deux techniques combinées doivent permettre de détecter une falsification d'identité, de position ou une tentative de spoofing.

3 Description fonctionnelle de l'AIS

L'AIS permet aux navires et aux centres de surveillance du trafic maritime d'échanger automatiquement des informations relatives à leurs positions, leur identité et leurs intentions, mais aussi des messages textes, binaires ou encore météorologiques. C'est un système composé de plusieurs types de transpondeurs radio qui échangent entre eux en utilisant deux fréquences VHF : 161.975 MHz et 162.025 MHz. Il est prévu pour fonctionner en cellules locales d'environ 80 à 100 km de diamètre.

Ces signaux utilisent la modulation GMSK (Gaussian Minimum Shift Keying). Enfin, la couche liaison utilise plusieurs variantes de multiplexage temporel TDMA (Time Division Multiple Access) pour accéder au réseau. Ce système à référence commune (le temps universel coordonné) divise une minute en 2250 slots de 26,67 ms chacun. Chaque trame AIS sera émise sur un (ou plusieurs) de ces slots.

4 Méthodologie

Une campagne de collecte de trames AIS a été réalisée en rade de Brest. Afin de disposer d'échantillons de la meilleure qualité possible pour cette première étude, un analyseur vectoriel NI PXI-5661 de National Instruments a été utilisé. Les trames décodées et les caractéristiques des signaux associés ont ensuite été intégrées dans une base de données géographique afin de réaliser une étude statistique. Cette collecte a permis d'obtenir 16 enregistrements de 5 minutes, contenant au total environ 10 000 trames exploitables.

Pour l'analyse du signal, cinq paramètres sont mesurés pour chaque trame AIS reçue. Il s'agit d'une part du niveau de puissance reçue, et d'autre part de quatre paramètres relatifs à la forme du signal. Il s'agit du *temps de montée*, du *temps de descente* (temps mis par l'émetteur pour passer de 10 % à 90 % de sa puissance d'émission maximale, respectivement son pendant en fin de trame) du *temps avant modulation* et du *temps après modulation* (temps entre le moment où l'émetteur atteint 50 % de sa puissance d'émission maximale et le moment où l'émetteur commence à moduler, respectivement son pendant en fin de trame).

5 Corrélation entre la puissance du signal et la distance

Dans un premier temps, l'étude se concentre sur la puissance du signal reçu. L'objectif est d'estimer la puissance de l'émetteur de la trame AIS en fonction de sa distance, du niveau de signal reçu, des caractéristiques de la chaîne de réception et du modèle de propagation. C'est une façon de vérifier que la distance entre la position reportée par l'émetteur et la position connue du récepteur est cohérente.

Le calcul de la puissance d'émission à partir de la puissance du signal reçu et de la distance consiste à estimer les pertes qu'a subies le signal sur son trajet entre l'émetteur et le récepteur. La méthodologie mise en oeuvre ici utilise une méthode d'approximation des conditions de propagation basée sur la formule de John Egli [3].

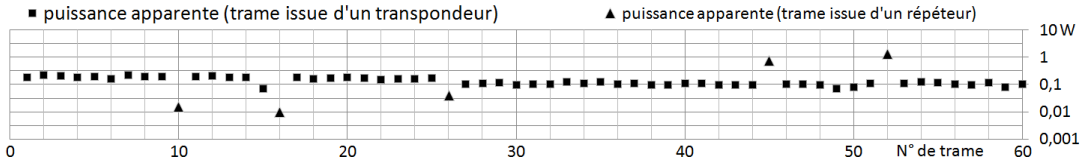

Fig. 1. Évolution, pour un même bâtiment, de la puissance en fonction du temps

Cette technique a permis de mettre en évidence deux méthodologies qui permettent de détecter des anomalies. La première, facile à mettre en oeuvre est basé sur un seuil fixé empiriquement (sur la base de l'analyse d'un nombre important de trames) au-delà duquel la trame reçue est considérée comme suspecte. La seconde, plus complexe à mettre en œuvre, consiste à comparer les trames consécutives d'un même bateau afin de relever un changement brutal de puissance apparente. La figure 1 illustre l'évolution de la puissance apparente pour un même navire. Bien que la distance séparant ce bateau du récepteur augmente, la puissance apparente reste stable. Par contre, pour le répéteur (équipement AIS à terre dont le rôle est de retransmettre certaines trames afin d'en augmenter la portée), la puissance apparente évolue dans le temps et il est possible de distinguer une trame issue du répéteur des trames issues du transpondeur légitime. Ces trames permettent d'illustrer la présence d'un intrus qui transmet une position différente de sa position réelle. Une telle détection pourrait être évitée en modulant la puissance du signal émis, mais cela ne permet de leurrer qu'une seule cible, et nécessite d'être capable de l'identifier et de la suivre.

6 Observations sur les caractéristiques des signaux

Les données utilisées sont également les trames de position. Les mesures devenant délicates et source d'erreur en deçà d'un certain niveau de réception, les données reçues avec un niveau de signal inférieur à -45 dBm ont été écartées. Ainsi, le corpus étudié se compose de 3646 trames.

La figure 2 est une représentation sous forme de boîte à moustaches du *temps avant modulation* pour les messages 1 (reports de positions de transpondeurs de classe A). Elles sont regroupées par navire sur l'axe des abscisses (dont la légende a été anonymisée), et la valeur en seconde du *temps avant modulation* est représentée sur l'axe des ordonnées. Les navires sont ici triés par ordre croissant de numéro identifiant (Maritime Mobile Service Identity). Elle permet d'observer de fortes disparités de valeur d'un navire à l'autre d'une part, et le bon regroupement des valeurs

pour chaque navire d'autre part. L'exploitation brute de ce seul paramètre ne permet pas une identification certaine, mais cette constatation valide l'hypothèse de l'existence de différences importantes entre transpondeurs.

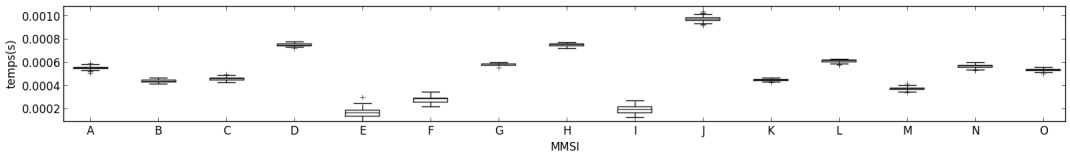

Fig. 2. Répartition du *temps avant modulation* par transpondeur de classe A

La figure 3 représente cette fois le *temps après modulation* pour les transpondeurs de classe A. Il est possible de constater une moins grande répartition des valeurs d'un transpondeur à l'autre, même si certains se détachent nettement. Cette information peut donc venir en complément du *temps avant modulation* qui est plus discriminant.

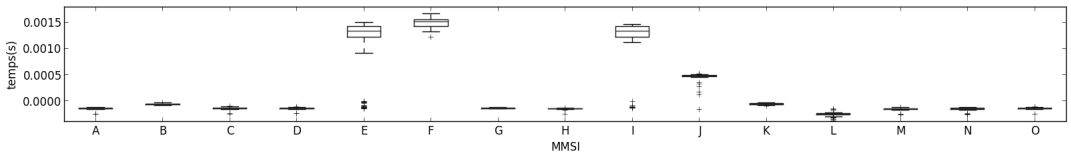

Fig. 3. Répartition du *temps après modulation* par transpondeur de classe A

Bien que non représentées ici, les deux autres caractéristiques étudiées (*temps de montée* et *temps de descente*) présentent une répartition des valeurs similaires. Au bilan, l'étude de ces seuls paramètres permet d'identifier de façon unique seulement certains bateaux dans l'échantillon analysé. Cette étude ouvre la voie à d'autres perspectives. Un premier objectif est de rechercher d'autres caractéristiques signantes en se basant sur un jeu de données beaucoup plus significatif. Par ailleurs il nous semble judicieux de rechercher à évaluer le taux de succès d'une identification sur une combinaison de paramètres ou encore d'évaluer d'autres méthodologies de qualification : forme de la courbe de distribution, mesure des delta d'une trame à la suivante, etc.

7 Conclusion

Ce travail a permis de montrer que l'étude des caractéristiques du signal est une approche intéressante si l'on souhaite donner un niveau de confiance aux données reçues, voire détecter une anomalie ou une attaque.

Certaines caractéristiques intrinsèques d'une installation radio peuvent en effet être mesurées à partir du signal reçu, et ces caractéristiques sont stables dans le temps. Vérifier la cohérence entre les signaux reçus et ses caractéristiques en disposant d'une base de connaissances et en vérifiant la stabilité de ces paramètres d'une trame à l'autre peut permettre de détecter une attaque.

D'autre part, l'analyse de la puissance du signal reçu peut également permettre de détecter certaines attaques. Si la puissance du signal reçu est manifestement incompatible avec la distance entre les transpondeurs ou si l'évolution de cette puissance est incohérente avec leurs cinématiques, l'anomalie peut être détectée.

Références

1. Marco Balduzzi, Alessandro Pasta, and Kyle Wilhoit. A security evaluation of ais automated identification system. In *Proceedings of the 30th Annual Computer Security Applications Conference*, pages 436–445. ACM, 2014.
2. Boris Danev, Davide Zanetti, and Srdjan Capkun. On physical-layer identification of wireless devices. *ACM Computing Surveys (CSUR)*, 45(1) :6, 2012.
3. John J Egli. Radio propagation above 40 mc over irregular terrain. *Proceedings of the IRE*, 45(10) :1383–1391, 1957.
4. Fotios Katsilieris, Paolo Braca, and Stefano Coraluppi. Detection of malicious ais position spoofing by exploiting radar information. In *Information Fusion (FUSION), 2013 16th International Conference on*, pages 1196–1203. IEEE, 2013.
5. Philipp Last, Martin Hering-Bertram, and Lars Linsen. How automatic identification system (ais) antenna setup affects ais signal quality. *Ocean Engineering*, 100 :83–89, 2015.
6. Giuliana Pallotta, Michele Vespe, and Karna Bryan. Traffic knowledge discovery from ais data. In *Information Fusion (FUSION), 2013 16th International Conference on*, pages 1996–2003. IEEE, 2013.
7. Francesco Papi, Dario Tarchi, Michele Vespe, Franco Oliveri, Francesco Borghese, Giuseppe Aulicino, and Antonio Vollero. Radiolocation and tracking of automatic identification system signals for maritime situational awareness. *IET Radar, Sonar & Navigation*, 9(5) :568–580, 2014.
8. McKay D Williams, Michael A Temple, and Donald R Reising. Augmenting bit-level network security using physical layer rf-dna fingerprinting. In *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*, pages 1–6. IEEE, 2010.
9. Windward. Ais data on the high seas : an analysis of the magnitude and implication of growing data manipulation at sea, 2014.