



**MINISTÈRE
DES ARMÉES**

*Liberté
Égalité
Fraternité*



Génération de jeux de données pour la Cybersécurité des systèmes navals

Rétrospective sur une thèse et mise en perspective



Clet Boudéhenn — Docteur, Ingénieur en Cybersécurité à DGA-MI

Mais d'abord, c'est qui celui la ?

Clet Boudéhenn

Docteur en informatique
Ingénieur en cybersécurité

2018 – 2022

Doctorat (c'était cool mais dur !)

Génération de données pour l'analyse et la détection d'anomalies dans les systèmes cybernétiques navals

16 déc. 2022

Soutenance (c'était court mais bien !)

Aujourd'hui

Expert à la DGA-MI (c'est chouette !)

Ingénieur cybersécurité. Spécialisé en Systèmes embarqués à la DGA-MI

Aujourd'hui

Enseignant (Intervenant Extérieur) à l'ENSTA (ça aussi !)

Formateur en « Cybersécurité des systèmes embarqués et de navigation »

Sommaire

1

Le contexte (pendant la thèse)

Numérisation du navire
et de son écosystème

2

Naval Cyber Range

Champ de tir cyber
numérique & physique
d'un navire

3

Génération de données

Démonstrateurs,
expériences réalisés,
systèmes de captation
développés

4

Détection d'anomalies

IDS / NIDS · TKEO ·
cinématique · Machine
Learning

5

Mise en perspective

Menaces actuelles,
normes, nouvelles
technologies : ce qui a
évolué depuis 2022

Mars 2021 :

Et si une cyberattaque pouvait générer ce type d'incident ?

Obstruction du Canal de Suez

(Pendant 6 jours) par le porte-Conteneurs « Ever Given »

(Source : CNES Geo Image)

PARTIE 1

LE CONTEXTE (pendant la thèse)

2018 – 2022 : Numérisation du navire et de son écosystème



Le navire, un système de systèmes cybernétiques

Le navire moderne est un ensemble de systèmes complexes et interconnectés

IT

Systèmes informatiques

Serveurs, réseaux, Wi-Fi, bureautique, administration, ...

OT

ICS / SCADA

Automates (PLC), supervision, propulsion, énergie, ...

NAV

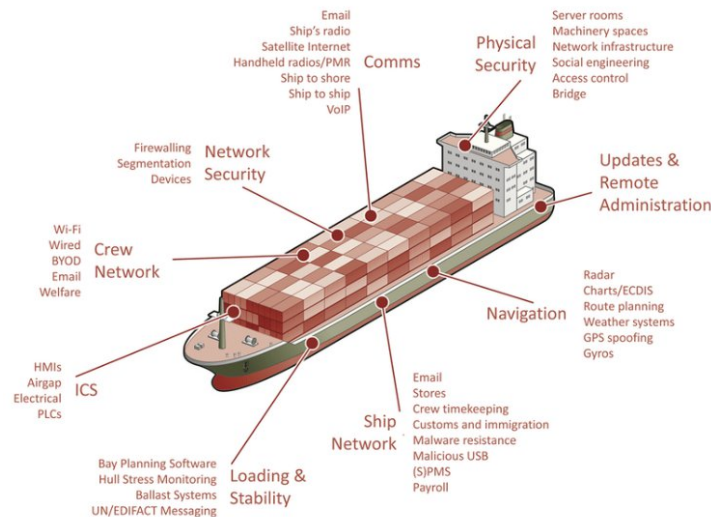
Systèmes de navigation

GNSS, AIS, ECDIS, Radar, VDR

EXT

Liens extérieurs

SatCom, 4G, maintenance à distance par l'armateur ou prestataire

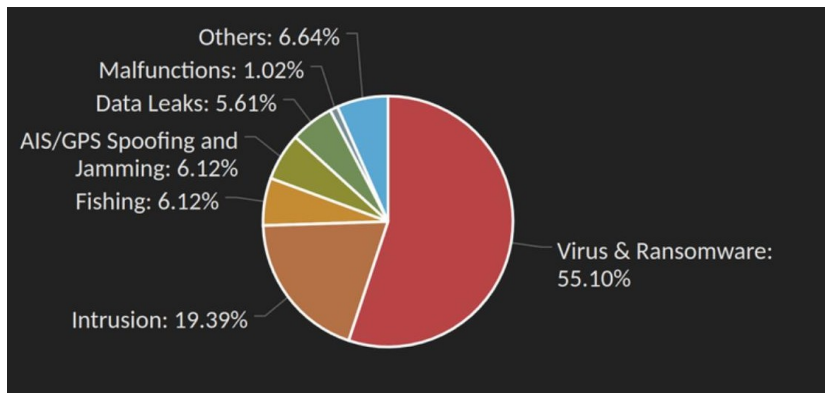


Mileski et al. (2018), *Cyberattacks on ships: a wicked problem approach* ; Caponi & Belmont (2015), *Maritime cybersecurity: a growing threat goes unanswered*.

Une menace qui a changé d'échelle

(pendant la thèse) : Cyberattaques recensées 1980-2021

Aujourd'hui — quelques chiffres



Le leurrage / brouillage AIS-GPS ne représentaient que **~6 %** des incidents recensés....

612

incidents cyber recensés en 2023 par le M-CERT dans le secteur maritime (dont 206 portuaires)

> 10 000

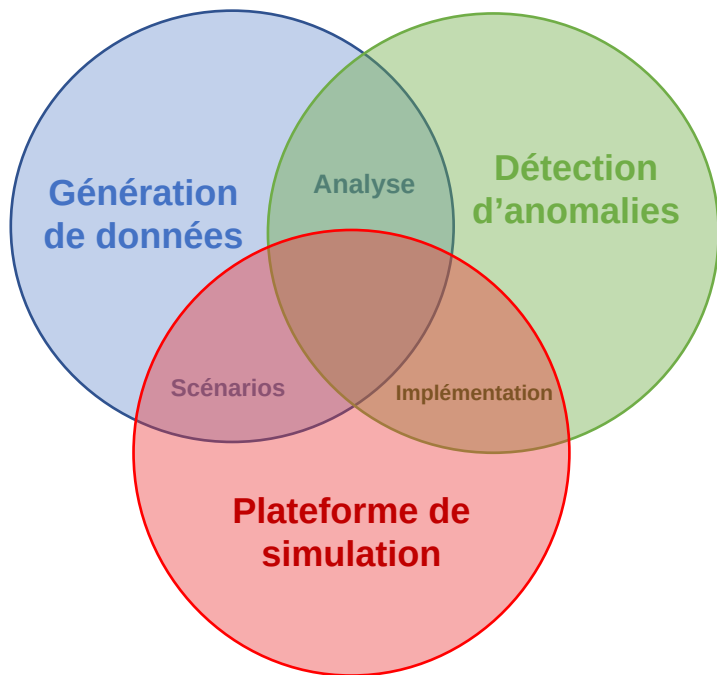
navires affectés par des interférences GNSS au 2e trimestre 2025

~ 3 000

navires perturbés dans le Golfe et le détroit d'Ormuz en juin 2025

Ben Farah et al. (2022), Information 13(1) ; France Cyber Maritime / M-CERT, Panorama 2023 ; GPSPatron, Maritime GNSS Interference Worldwide (2025).

Les différentes questions de recherche



QR1

Comment générer des données représentatives et mener des attaques réalistes pour valider les travaux ?

QR2

Quels sont les marqueurs significatifs des attaques générées ?

QR3

Comment détecter les anomalies dans ce contexte maritime ?

QD1

Quels outils développer pour faciliter la génération de données et la détection ?

Anomalie : « observation non conforme au comportement attendu » (Chandola et al., 2009)

PARTIE 2

NAVAL CYBER RANGE

Champ de tir cyber numérique & physique d'un navire



Naval Cyber Range :

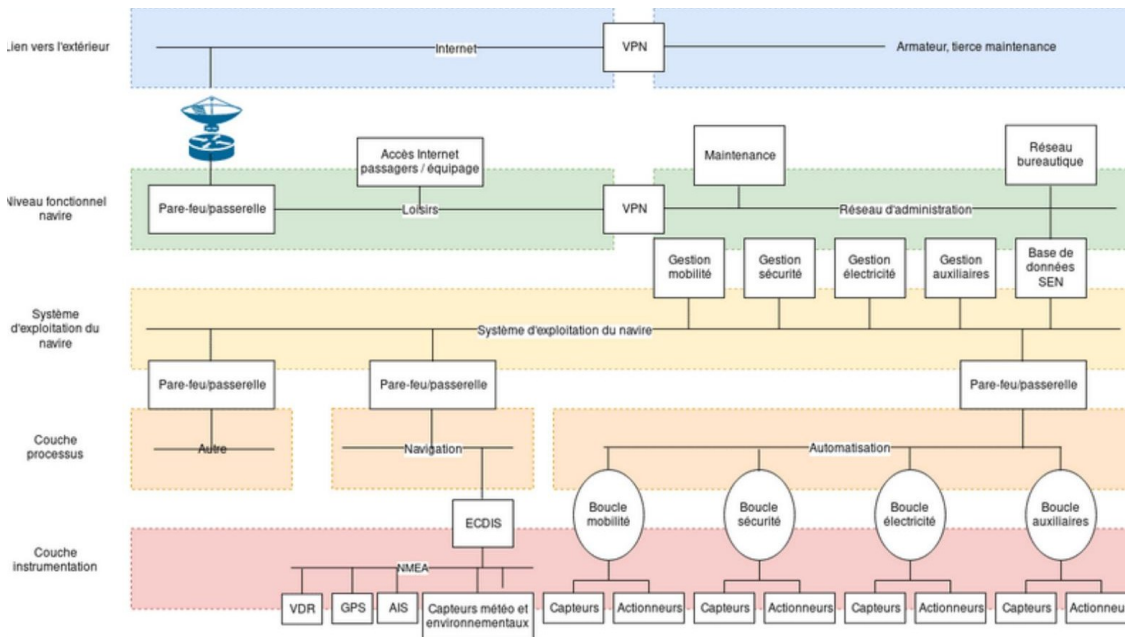


plate-forme reproduisant les systèmes cybernétiques d'un navire :

- **4 boucles industrielles** : propulsion, sécurité, électricité, auxiliaires (automates Siemens S7, capteurs, actionneurs)
- **Passerelle de navigation** : simulateur Bridge Command, ECDIS OpenCPN, GPS, AIS, VDR
- **Architecture en couches** : de l'instrumentation aux liens extérieurs
- **Sondes réseau et capteurs** + corrélation des informations & des événements (SIEM via ELK)

Naval Cyber Range :

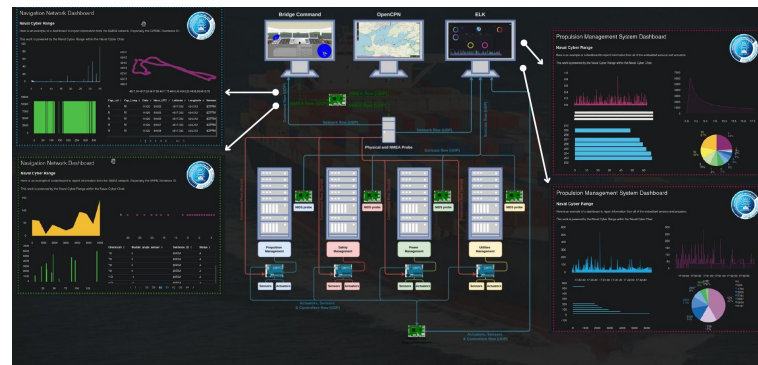
Un environnement réaliste pour rejouer des attaques et produire des données labellisées



Boucles industrielles



Passerelle de navigation



Supervision, sondes et tableaux de bord

Et aujourd'hui alors ?

1

Testbeds virtuels open source

MaCySte (Univ. de Gênes, SoftwareX 2023) : pont intégré (ECDIS, radar), automates, protocoles NMEA / Modbus / ASTERIX, SIEM ; entièrement conteneurisé et reproductible (docker, katharà, ...).

2

Jumeaux numériques

Couplage numérique-physique pour les navires autonomes (ex. banc du Marine Cybernetics Lab, NTNU, 2025) : Les jumeaux numériques cyber deviennent incontournables.

3

Écosystème opérationnel

Centres de réponse dédiés (France Cyber Maritime / M-CERT) et panoramas annuels de la menace : la donnée d'incident devient disponible.

Lecture : ce que nous faisons avec de vrais automates se fait aujourd'hui largement en virtuel : on gagne en reproductibilité et en partage, mais le biais de simulation que nous avons identifié en 2022 reste la limite n°1.

Longo et al. (2023), MaCySte: A virtual testbed for maritime cybersecurity, SoftwareX 23 ; arXiv:2505.06787 (2025) ; france-cyber-maritime.eu

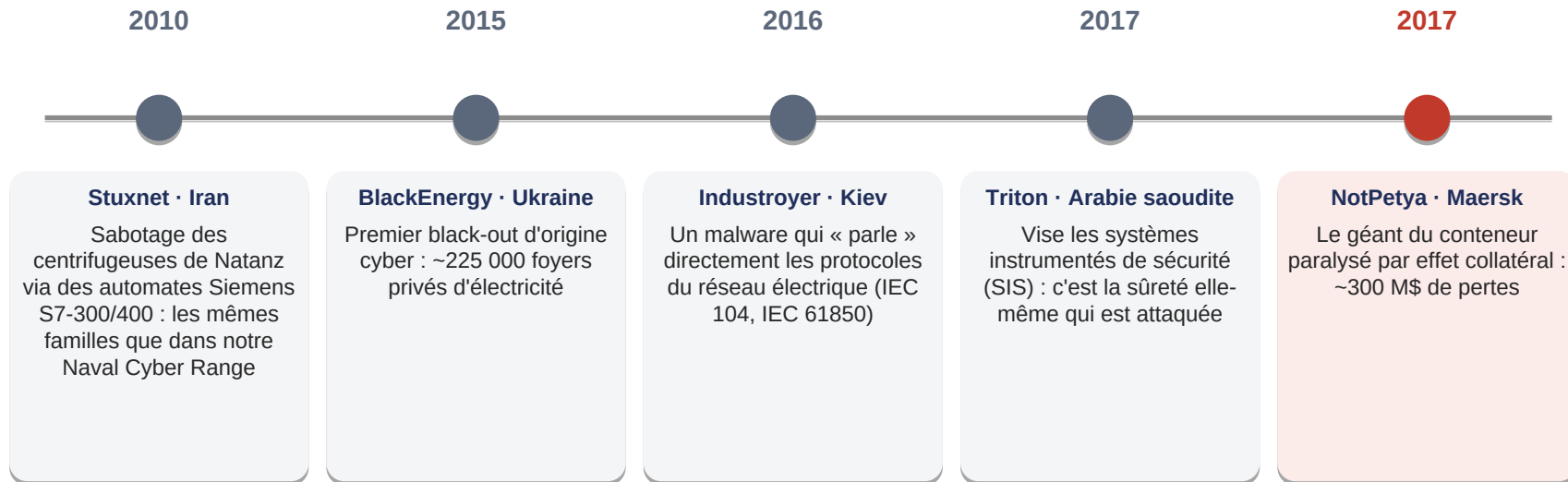
PARTIE 3

GÉNÉRER DES DONNÉES

Des attaques réalistes, des données représentatives



Quelques exemples d'attaques sur les ICS

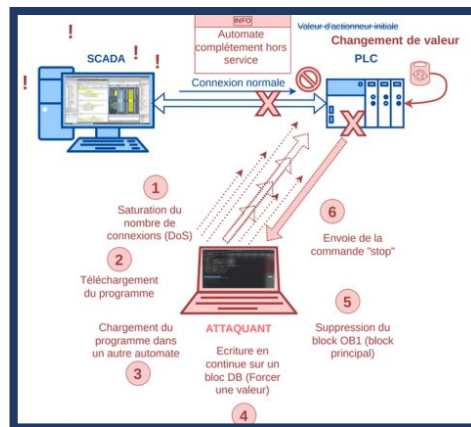
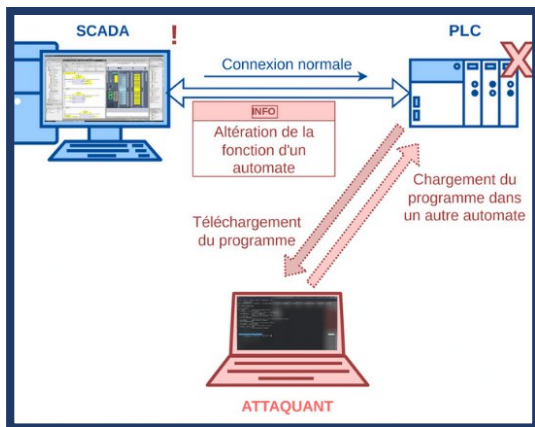


Un automate est fait pour fonctionner 24h/24, 7j/7 et 365j/an. Il a été conçu pour être sûr et fiable, pas sécurisé.

Hemsley & Fisher (2018), A history of cyber incidents and threats involving industrial control systems, CIP, Springer.

Démonstrateur AEGIS :

Démonstrateur de génération d'attaques sur automates — Attack and Exploit Generator on Industrial control Systems

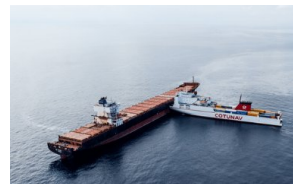


Type d'attaques rejouées

- Reconnaissance, brute-force, contournement d'authentification
- Injection de commandes (ex. « PLC-stop »), lecture/écriture mémoire
- Déni de service, Man-in-the-Middle, altération de firmware

Conséquences imaginées

- Sabotage, manœuvrabilité compromise, échouement, collision



Automates Siemens S7-300 / S7-400 ; références : Stuxnet, BlackEnergy, Industroyer (Hemsley & Fisher, 2018).

« The geopolitic connection » : 2022 – 2026

Russie · Ukraine · Europe

Avr. 2022 — Industroyer2

Nouvelle tentative contre le réseau électrique ukrainien, déjouée

Janv. 2024 — FrostyGoop

Lviv : commandes Modbus TCP, 600 immeubles sans chauffage en plein hiver

Avr. 2025 — Barrage (Norvège)

Vanne ouverte à distance, 500 L/s pendant 4 h ; acteurs pro-russes suspectés

Iran · Israël

Nov. 2023 — CyberAv3ngers

Groupe lié aux Gardiens de la révolution : automates Unitronics compromis, dont l'eau potable aux États-Unis

Mars & août 2025 — Lab Dookhtegan

Communications satellite de 116 puis 60 navires iraniens (NITC, IRISL) mises hors service

Prépositionnement

2022 — Pipedream / Incontroller

Boîte à outils modulaire visant plusieurs marques d'automates (Schneider, Omron, OPC UA)

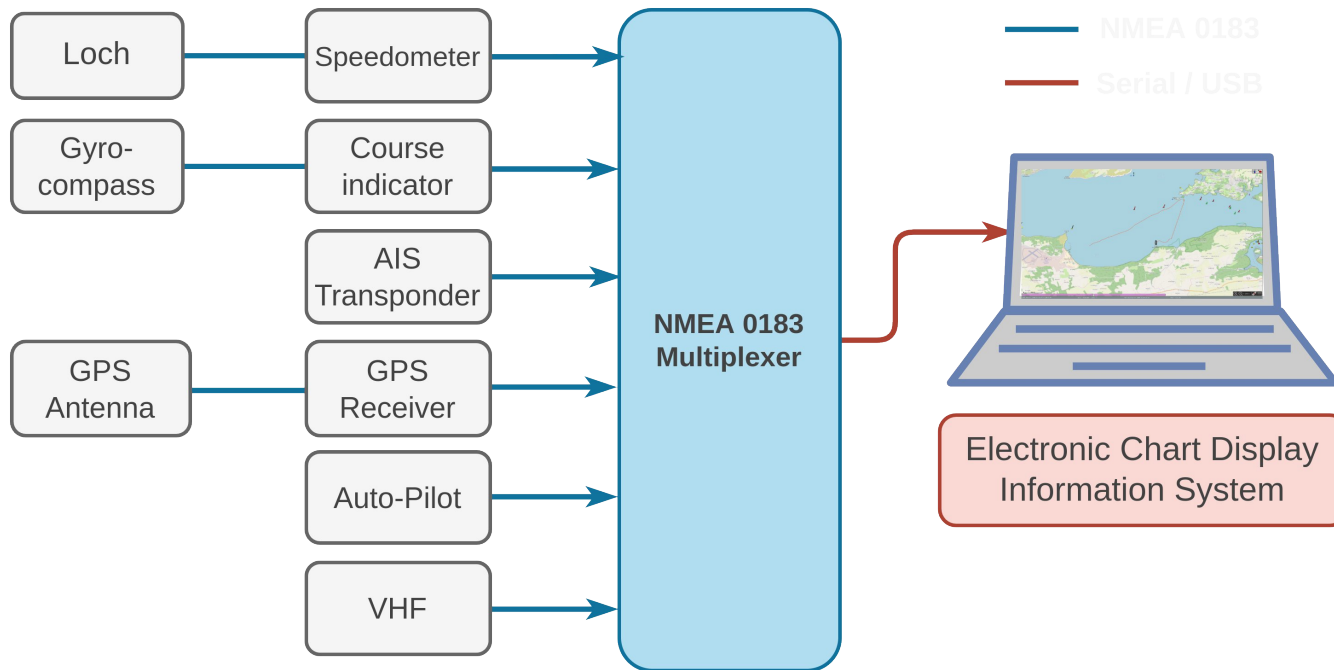
2023 → 2026 — Volt Typhoon

Présence discrète et durable dans les infrastructures critiques américaines, dont les transports : se préparer plutôt qu'attaquer

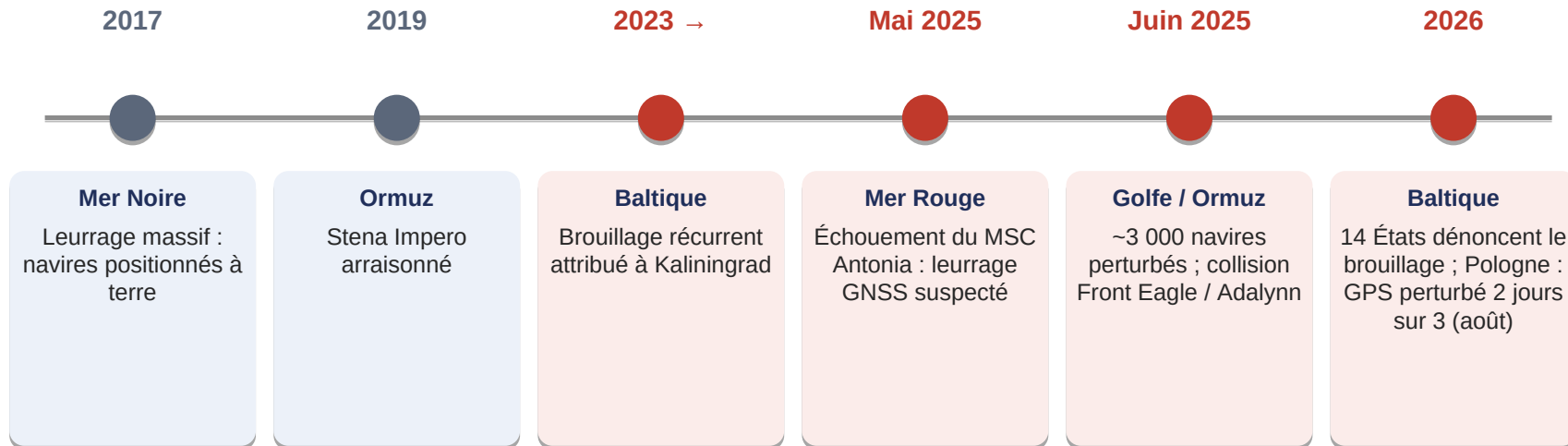
l'OT est devenu, comme présagé, un terrain de conflit à part entière — sabotage, perturbation, déstabilisation, pré-positionnement... le navire et son écosystème en fait partie.

ESET (2022) ; Dragos, FrostyGoop (2024) ; The Record (2025) ; CISA AA23-335A & AA24-038A ; Security Affairs (2025).

Les systèmes de navigation



Quelques attaques sur les systèmes de navigation



Les attaques étudiés et analysées, sont devenues presque quotidiennes dans certaines zones.

Inside GNSS (mai 2025 ; janv. 2026) ; GPSPatron (2025) ; Windward / gCaptain (juin 2025) ; Euronews (août 2026). Rôle du GNSS dans la collision Front Eagle encore débattu.

Le point d'entrée : le standard NMEA-0183

Standard historique (IEC 61162-1) : trames ASCII, liaison série, un émetteur vers plusieurs récepteurs

```
$GPGGA,052001.000,4822.8632,N,00429.8424,W,1,08,1.3,3.4,M,52.1,M,,0000*45
```

GP

Talker : GPS

GGA

Type : fix data

052001

Heure UTC

4822.8632 N

Latitude

00429.8424
W

Longitude

1

Qualité : GPS

08

Satellites

1.3

HDOP

*45

Checksum

Aucune authentification, aucun chiffrement

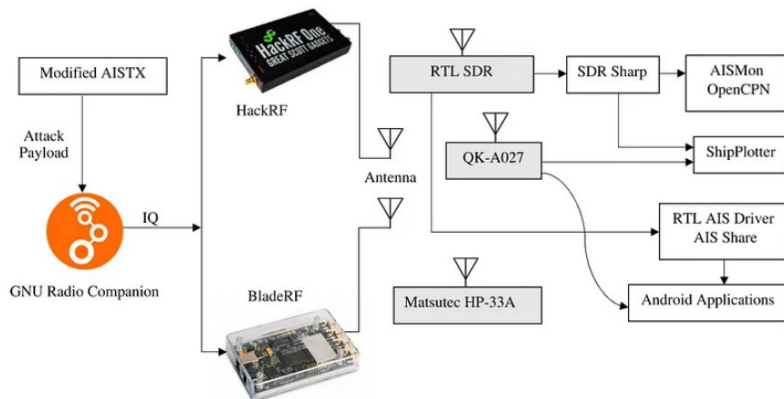
Un simple checksum XOR. Qui contrôle le flux contrôle la position affichée sur l'ECDIS.

Les successeurs ne l'ont pas remplacé

NMEA 2000 (bus CAN, IEC 61162-3), puis OneNet / IEC 61162-450 (Ethernet) : NMEA-0183 reste omniprésent dans le parc installé.

BELAMY : l'expérience de leurrage

Expérience — Basic Experimentation of Light spoofing Attack on Maritime SYstem



Ce que l'attaque laisse dans les trames

Trames GPRMC (comportement normal)									
\$	GPRMC	,125630,A,4817.0063,N,00423.8260,W	22.9	105.8	,010920,1.3,W,A*23				
\$	GPRMC	,125631,A,4817.0062,N,00423.8250,W	22.9	105.8	,010920,1.3,W,A*20				
\$	GPRMC	,125632,A,4817.0062,N,00423.8243,W	22.9	105.8	,010920,1.3,W,A*21				
Trames GPRMC (comportement anormal)									
\$	GPRMC	,125633,V,4817.0061,N,00423.8238,W	0,0	10920	1.3,W,N*02				
\$	GPRMC	,125634,V,4817.0061,N,00423.8234,W	0,0	10920	1.3,W,N*09				
\$	GPRMC	,125635,V,4817.0061,N,00423.8231,W	0,0	10920	1.3,W,N*0D				

Phase de brouillage

Transition brouillage - leurrage,
champs NMEA impactés

Trames GPGBA (comportement normal)									
\$	GPGBA	,125630,4817.0063,N,00423.8260,W	1	00	2.8	3.5	M,,,*47		
\$	GPGBA	,125631,4817.0062,N,00423.8250,W	1	00	2.8	3.5	M,,,*44		
\$	GPGBA	,125632,4817.0062,N,00423.8243,W	1	00	2.8	3.5	M,,,*45		
Trames GPGBA (comportement anormal)									
\$	GPGBA	,125633,4817.0061,N,00423.8238,W	0	00	0,0	M,,,*46			
\$	GPGBA	,125634,4817.0061,N,00423.8234,W	0	00	0,0	M,,,*4D			
\$	GPGBA	,125635,4817.0061,N,00423.8231,W	0	00	0,0	M,,,*49			

Marqueurs observés (QR2)

- **RMC** : statut A → V, vitesse et cap à 0
- **GGA** : qualité → 0, satellites → 00, HDOP et altitude vides
- **Puis** : retour à des valeurs cohérentes... mais fausses

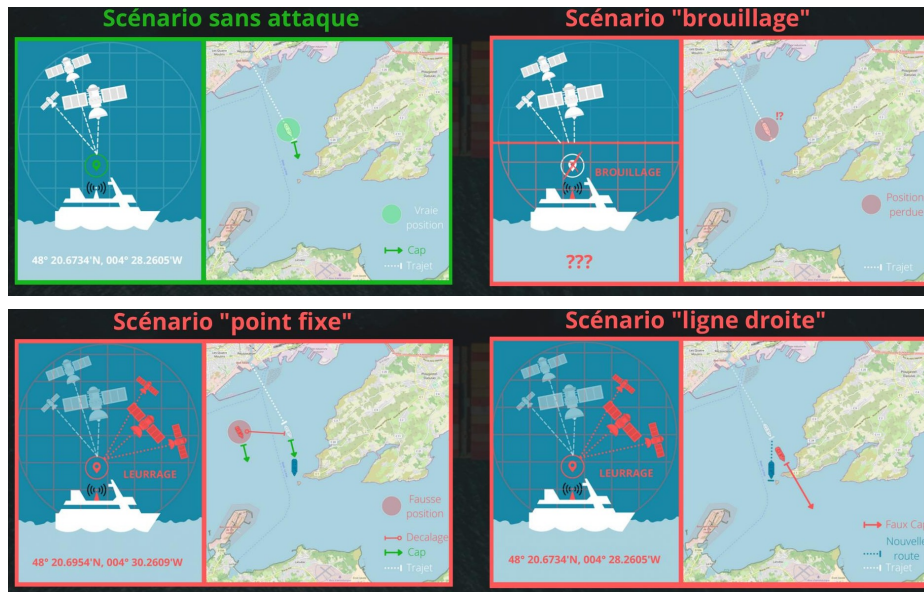
→ Le brouillage est visible ; le leurrage bien fait ne l'est presque plus.

[1] Amro, A., Oruc, A., Gkioulos, V. & Katsikas, S. (2022). Navigation Data Anomaly

Amro, Oruc, Gkioulos & Katsikas (2022), Navigation Data Anomaly Analysis and Detection, Information 13(3).

Démonstrateur NAGE :

Démonstrateur de génération d'attaques sur la navigation — NMEA Attack Generation Engine



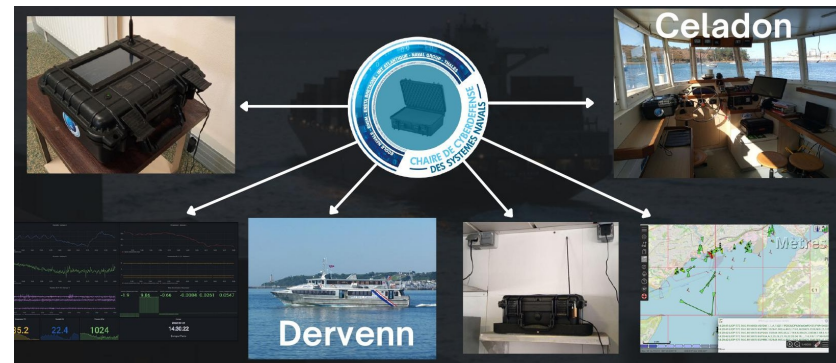
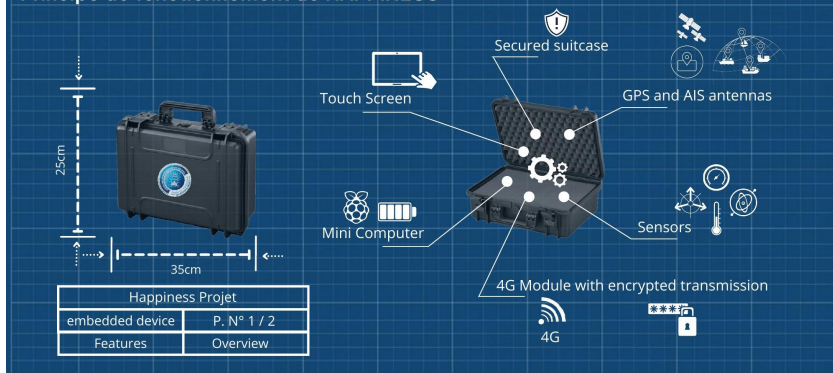
Le pendant « navigation » d'AEGIS :

- **Entrée** : flux NMEA réels (récepteur) ou simulés (Bridge Command)
- **Scénarios** : sans attaque, brouillage, point fixe, ligne droite, décalage
- **Sortie** : trames modifiées et labellisées, directement exploitables par l'ECDIS
- **Usage** : constituer des jeux de données pour entraîner et évaluer les détecteurs

Mallette d'acquisition de données HAPPINESS :

Projet — Holistic APProach of Integrated Navigation Equipment for Cybersecurity at Sea

Principe de fonctionnement de HAPPINESS



Mallette autonome : système embarqué indépendant du navire : mini-PC, antennes GPS/AIS, capteurs, 4G chiffrée

Données ouvertes : collecte NMEA en temps réel, sans protocole propriétaire, à bord du Celadon et du Dervenn

Moins de biais : des scénarios réalistes rejoués dans le Naval Cyber Range, au lieu du seul simulé

Générer des données : hier et aujourd'hui

	Au moment de la thèse (2019-2022)	Aujourd'hui (2026)
Sources	Naval Cyber Range + données réalistes de navires (HAPPINESS)	AIS terrestre et satellitaire massif ; détection d'interférences à grande échelle (ex. Spire, Windward)
Jeux publics	Quasi inexistants pour la cyber navigation	MARSIM (Fraunhofer FKIE, 2023) : 43 320 scénarios de leurrage simulés, en accès ouvert
Attaques	Générées et maîtrisées (démonstrateurs AEGIS, NAGE ; expérience BELAMY)	Attaques réelles observées et documentées : une vérité terrain enfin disponible
Limite	Environnement limité, diversification limitée	Labellisation, confidentialité et partage des données sensibles

Partir de données terrain plutôt que des données simulés est devenue la norme ; (le défi des chercheurs s'est déplacé vers la qualité des labels et la représentativité des données partagées).

Spravil et al. (2023), Detecting Maritime GPS Spoofing Attacks Based on NMEA Sentence Integrity Monitoring, JMSE 11(5):928 (jeu MARSIM).

PARTIE 4

DÉTECTER LES ANOMALIES

Des marqueurs aux détecteurs



Pourquoi utiliser la détection d'anomalies ?

Détection par signatures

- Reconnaît les attaques déjà connues, avec peu de faux positifs
- **Aveugle** face à l'inédit (0-day) et aux protocoles métiers (NMEA, Modbus, S7)

Détection par analyse d comportement

- Modélise le **comportement normal** et signale tout écart
- Détecte l'**inconnu** : attaques inédites... et pannes
- Contrepartie : faux positifs et besoin de **données représentatives**

15-40

Durée de vie

Équipements certifiés, durée de vie de 15 à 40 ans : on ne pourra pas tout corriger, il faudra surveiller

SAT

Isolé de la terre (normalement)

Connectivité limitée : la détection doit être au maximum autonome, à bord.

CPS

Systèmes Cyber Physique

Une anomalie peut être une attaque ou une panne : mélange entre sûreté et sécurité.

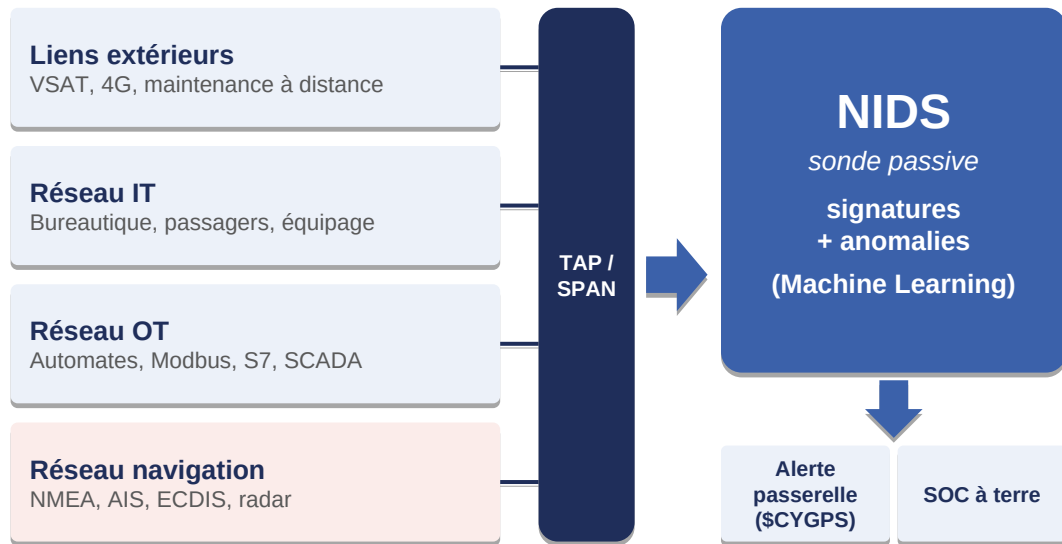
CSA

Vue d'ensemble situationnelle

Maritime Cyber Situational Awareness : donner suffisamment d'information à l'équipage pour aider à la décision.

Chandola et al. (2009), *Anomaly detection: a survey* ; Vaarandi et al. (2025), *Literature Review: Cyber Security Monitoring in Maritime*, arXiv:2503.18173.

Les NIDS :



Dans le Naval Cyber Range : sondes réseau de détection d'intrusion + sondes capteurs, corrélées dans un même tableau de bord.

Pourquoi ils sont primordiaux

- **Passifs** : écoute via TAP ou port miroir, sans latence ni impact sur des équipements certifiés
- **Vision globale** : flux IT, OT et navigation, protocoles métiers compris
- **Filet de sécurité** : la segmentation réduit le risque, le NIDS voit ce qui passe quand même
- **Exigés** : fonction « Détecter » de l'IACS UR E26 ; pas de notification d'incident (NIS2, USCG) sans détection

Détection d'une attaque sur les automates

(Comportement normal)



(Comportement anormal / attaque "PLC-stop")

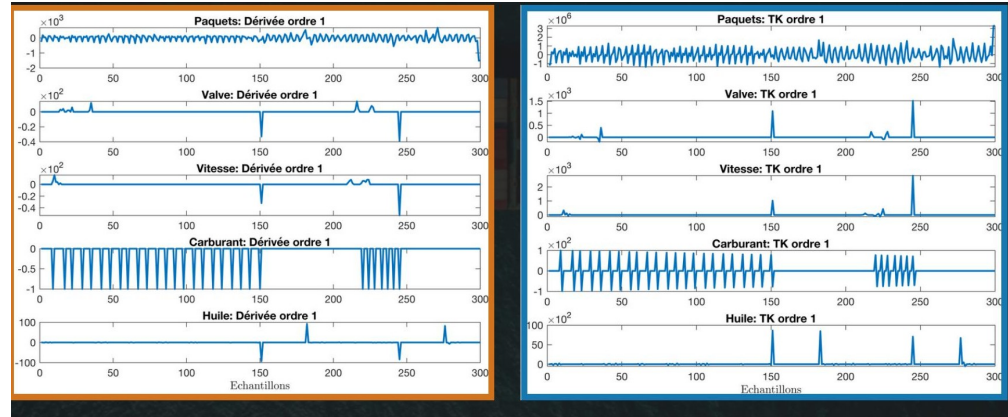


- **Corrélation** capteurs / actionneurs / trafic réseau
- **Attaque « PLC-stop »** : chute de la vitesse d'hélice, de la consommation, rupture du trafic
- **Stratégie d'extraction** de données pour des alertes cyber-physiques (CyberSA 2020)

L'opérateur d'énergie de Teager-Kaiser

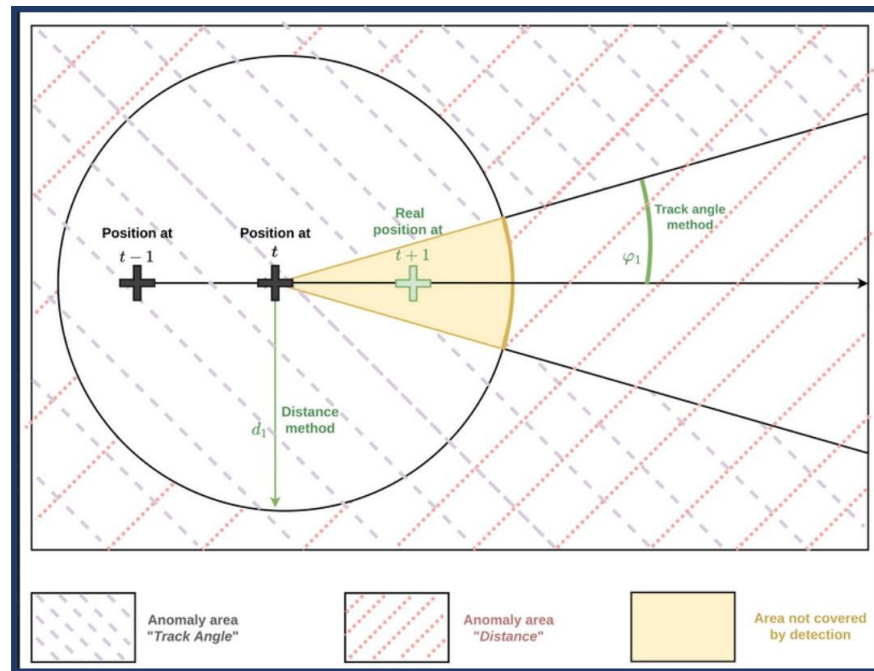
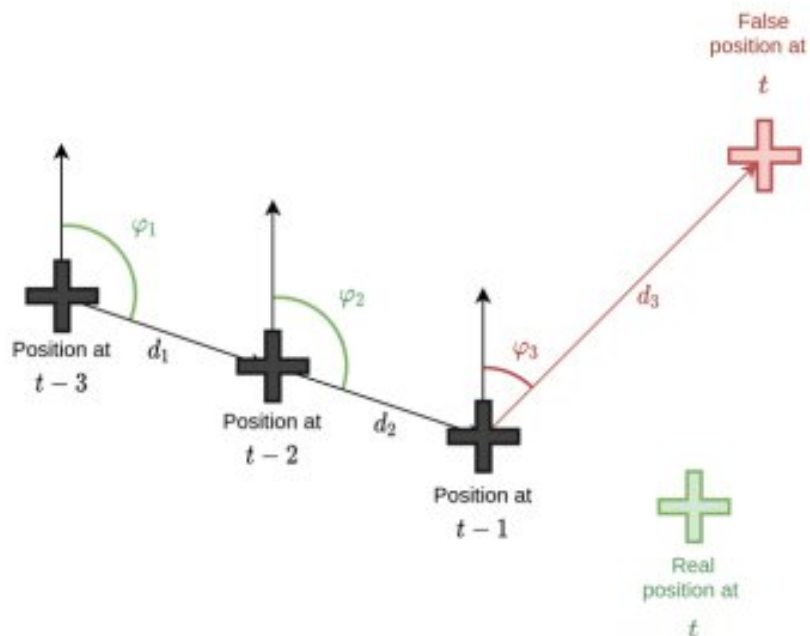
$$\begin{aligned}
 (1) \quad \Psi_c[x(t)] &= [\dot{x}^2(t)] - x(t)\ddot{x}(t) \\
 (2) \quad \Psi_d[x(n)] &= x^2(n) - x(n-1)x(n+1) \\
 (3) \quad \Psi_{p,q,m,l}[x(t)] &= x^{(p)}(t)x^{(q)}(t) - x^{(m)}(t)x^{(l)}(t) \\
 (4) \quad \Psi_{d_m}[x(n)] &= x^2(n) - x(n-m)x(n+m)
 \end{aligned}$$

- **Idée** : suivre les changements locaux et discontinuités du signal
- **Plus sensible que la dérivée** aux ruptures (valve, vitesse, huile)
- **Échelles de résolution** : généralisation de l'ordre 1 à 4



Kaiser (1990), ICASSP ; Boudraa & Salzenstein (2018), Teager-Kaiser energy methods for signal and image analysis: a review, DSP 78.

Détection d'anomalie par cinématique



Machine learning : l'arbre généalogique



En 2020, placer « SVM » et « Isolation Forest » dans une réunion suffisait à passer pour un expert en IA. Spoiler : ça n'a pas duré.

Thapa et al. (2020), Comparison of ML and DL models for network intrusion detection systems, Future Internet ; Perera et al. (2021), One-class classification: a survey.

Navigation : apprentissage automatique

Algorithme de détection de nouveauté	Attaque de leurrage GNSS	Précision	Score F1
OCSVM	Static	97.9%	99.0%
LOF	Static	98.4%	99.2%
IF	Static	99.0%	99.5%
RC	Static	89.1%	94.8%
OCSVM	Decoy	95.0%	97.5%
LOF	Decoy	94.0%	96.9%
IF	Decoy	96.9%	98.4%
RC	Decoy	80.1%	88.9%
OCSVM	Straight	98.9%	99.5%
LOF	Straight	97.8%	98.9%
IF	Straight	98.9%	99.5%
RC	Straight	83.0%	90.7%
OCSVM	Offset	97.4%	98.7%
LOF	Offset	96.6%	98.3%
IF	Offset	97.8%	98.9%
RC	Offset	93.1%	96.4%

- **Supervisé** : très bons scores, mais dépend des attaques déjà connues
- **Non supervisé** (détection de nouveauté) : OCSVM, LOF, Isolation Forest, Robust Covariance
- **4 scénarios de leurrage** : statique, leurre, ligne droite, décalage

Isolation Forest

meilleur compromis : précision 96,9 – 99 %, score F1 ≥ 98,4 % sur les 4 scénarios

Remonter l'alerte : la trame \$CYGPS

Field	Comments
\$CYGPS	Sentence ID
CY	Talker ID, "cyber" component
120004	time of the fix, here 12 :00 :04 UTC
1-0	Anomaly indicator (1 : an anomaly is detected 0 : NO anomaly detected).
*64D	the checksum of the data

→ \$CYGPS, 120004, 1 *.64D

Exemples d'utilisation dans le cas d'un leurrage

Conception d'une nouvelle trame « CYBER »

- **Talker CY**, heure, indicateur d'anomalie 0/1, checksum
- **Compatible** avec l'existant : l'ECDIS la lit comme n'importe quelle trame
- **Indicateur temps réel** : vert (OK), gris (pas de données), rouge (leurrage détecté)

Valorisation des travaux de thèse

2020

*A data extraction method for
anomaly detection in naval systems*

Boudéhenn, Cexus, Boudraa — CyberSA
2020, IEEE

2021

*Navigation anomaly detection: an
added value for maritime cyber
situational awareness*

Boudéhenn, Jacq, Lannuzel, Cexus, Boudraa
— CyberSA 2021, IEEE

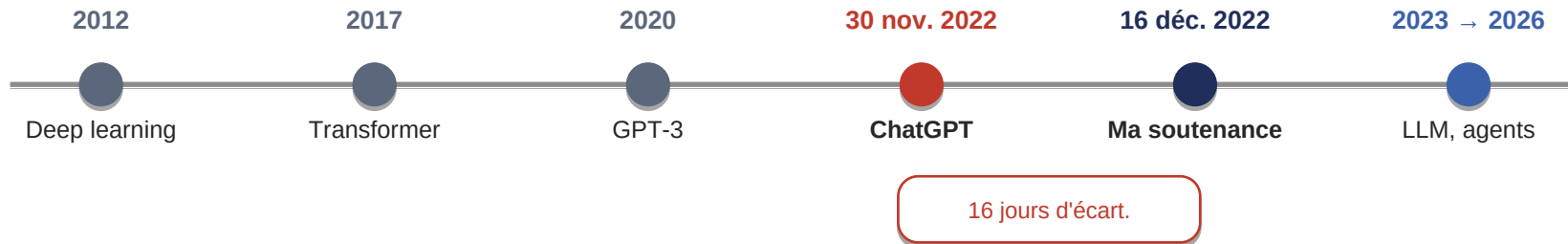
2022

*Holistic Approach of Integrated
Navigation Equipment for
Cybersecurity at Sea*

Boudéhenn, Lannuzel, Cexus, Jacq, Brosset,
Boudraa — Cyber Science 2022, Springer

Démonstrateurs AEGIS & NAGE · Expérience BELAMY · Mallette HAPPINESS

Évolutions des modèles d'apprentissage



Ce que les LLM apportent

- **Trajectoires AIS** : prédiction, anomalies et risque de collision expliqués (AIS-LLM)
- **SOC augmenté** : tri des alertes, synthèse d'incidents, aide à l'écriture de règles
- **CPS** : extraction d'invariants physiques pour la détection d'anomalies

Leurs limites

- **Hallucinations** et explicabilité : difficile à certifier
- **Coût de calcul** face aux contraintes embarquées
- **Nouvelle surface d'attaque** : injection de prompt, empoisonnement

Vaswani et al. (2017) ; Park et al., AIS-LLM (arXiv:2508.07668) ; LLM-assisted Physical Invariant Extraction for CPS Anomaly Detection (arXiv:2411.10918, 2024).

Les nouvelles méthodes en 2026

1 Contrôle d'intégrité NMEA

MANA (Fraunhofer FKIE, 2023) : 5 méthodes combinées (distance entre 2 récepteurs, éphémérides, dérive d'horloge, C/N0, cohérence physique), précision moyenne 0,92. Très proche de notre logique cinématique.

2 Transformers & LLM

TrAISformer (trajectoires AIS génératives) puis AIS-LLM (2025) : prédiction, anomalies et risque de collision, avec explications en langage naturel.

3 Apprentissage fédéré

IDS fédérés pour radars et réseaux de navires (ex. MRS-PFIDS, 2025) : apprendre sur plusieurs navires sans centraliser les données — une piste que nous citions en 2022.

4 Authentification à la source

Galileo OSNMA opérationnel depuis le 24 juillet 2025 : message de navigation signé. Efficace contre le leurrage, pas contre le brouillage. (authentification du message GNSS par hash retardés)

Mais la détection d'anomalie ne remplacera pas l'authentification (défense en profondeur).

Spravil et al. (2023) JMSE ; Nguyen & Fablet, TrAISformer (arXiv:2109.03958) ; Park et al., AIS-LLM (arXiv:2508.07668) ; Int. J. Inf. Security (2025) ; EUSPA (2025).

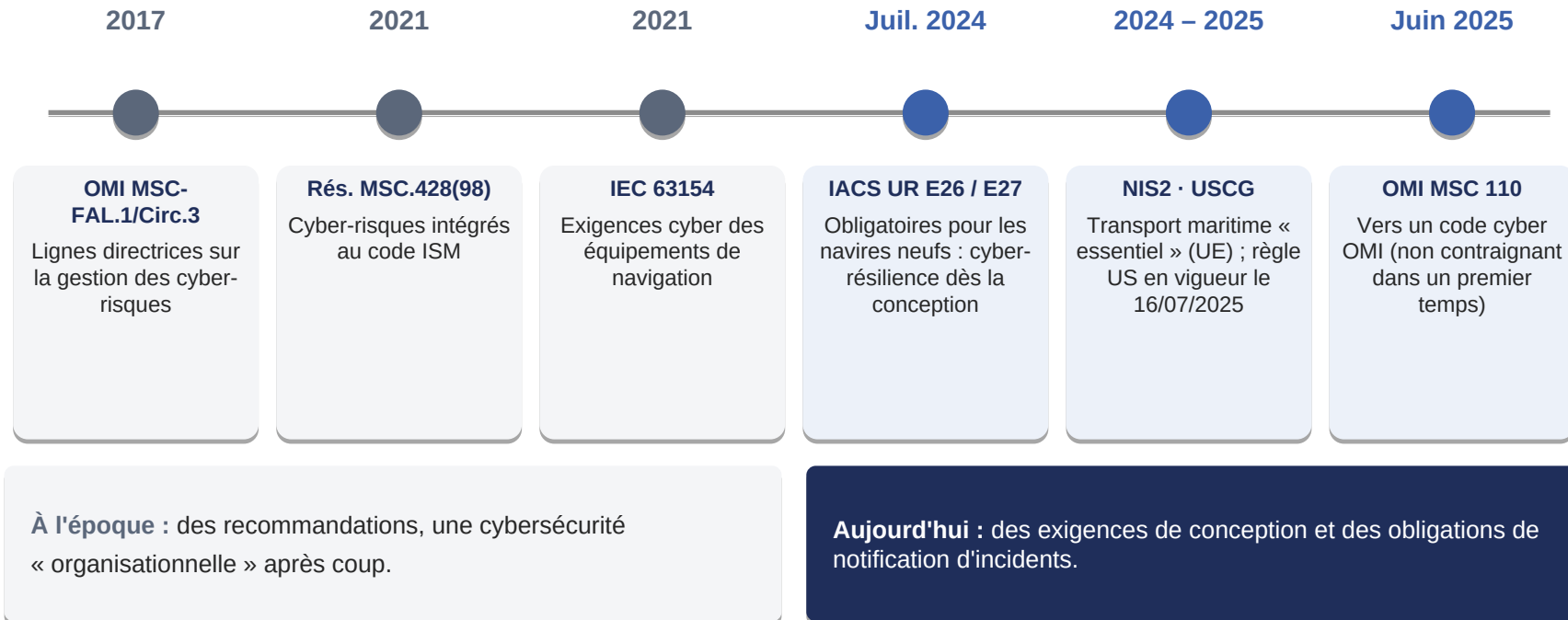
PARTIE 5

MISE EN PERSPECTIVE

De 2022 à 2026 : ce qui a changé



Réglementation : du conseil à l'obligation



IMO ; IACS (UR E26/E27) ; Directive (UE) 2022/2555 ; US Coast Guard, Final Rule 90 FR 6298 (17/01/2025) ; DNV, MSC 110 (2025).

Bilan comparatif

Ce qui avait été identifié pendant la thèse

- **La navigation** (GNSS, AIS, NMEA) comme maillon faible du navire
- **Les données terrain** ont permis de sortir du biais de simulation (Démonstrateur NAGE & AEGIS, mallette HAPPINESS et expérimentation BELAMY)
- **Des détecteurs légers**, explicables, compatibles NMEA (\$CYGPS) peuvent être intégrés facilement
- **L'apprentissage non supervisé** a été efficace face à des nouveaux types d'attaques

Ce qui a changé / a été sous-estimé

- **L'échelle** : plus besoin d'être proche du navire (flotte et armateur connectés)
- **L'AIS lui-même** de plus en plus manipulé : identités usurpées, flotte fantôme
- **La réglementation** devenue contraignante (E26/E27, NIS2, USCG)
- **La défense à la source** : OSNMA, PNT alternatifs (R-Mode, eLoran) (ou à la réception : CRPA)
- **L'IA générative** : transformers et LLM directement sur des données maritimes

ET MAINTENANT ALORS ?





**MINISTÈRE
DES ARMÉES**

*Liberté
Égalité
Fraternité*



Génération de jeux de données pour la Cybersécurité des systèmes navals

Rétrospective sur une thèse

DES QUESTIONS ?



Clet Boudéhenn — Docteur, Ingénieur en Cybersécurité à DGA-MI