Détection de dysfonctionements et d'actes malveillants basée sur des modèles de qualité de données multi-capteurs

Pedro Merino Laso

▶ **To cite this version:**

Pedro Merino Laso. Détection de dysfonctionements et d'actes malveillants basée sur des modèles de qualité de données multi-capteurs. Performance et fiabilité [cs.PF]. Ecole nationale supérieure Mines-Télécom Atlantique, 2017. Français. NNT : 2017IMTA0056 . tel-01813616

HAL Id: tel-01813616
https://theses.hal.science/tel-01813616

Submitted on 12 Jun 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

UNIVERSITE BRETAGNE LOIRE

THÈSE / IMT Atlantique

sous le sceau de l'Université Bretagne Loire

pour obtenir le grade de

DOCTEUR D'IMT Atlantique

Spécialité : Informatique

École Doctorale Mathématiques et STIC

Présentée par

Pedro Merino Laso

Préparée dans le département Image & traitement de l'information

Laboratoire Labsticc

Détection de dysfonctionements et d'actes malveillants basée sur des modèles de qualité de données multi-capteurs

Thèse soutenue le 07 décembre 2017

devant le jury composé de :

Martine Collard
Professeure des universités, Université des Antilles / présidente

Eloi Bosse
Chercheur, McMaster University - Canada / rapporteur

Joaquin Garcia-Alfaro
Professeur, Télécom SudParis / rapporteur

Michaël Hauspie
Maître de conférences (HDR), Université de Lille 1 / examinateur

David Brosset
Maître de conférences, Ecole Navale – Brest / examinateur

John Puentes
Professeur, IMT Atlantique / directeur de thèse

Philippe Leroy
Spécialiste Senior Cyber, Thales Communication & Security – Gennevilliers / invité

Patrick Hebrard
Responsable Recherche et Innovation, Naval Group - Ollioules / invité

N° d'ordre : 2017IMTA0056

Sous le sceau de l'Université Bretagne Loire

IMT Atlantique
Bretagne-Pays de la Loire

En accréditation conjointe avec l'Ecole Doctorale Sicma

Détection de dysfonctionnements et d'actes malveillants basée sur des modèles de qualité de données multi-capteurs

---

Thèse de Doctorat

Mention : Informatique

Présentée par **Pedro Merino Laso**

Département : Image et traitement de l'information

Laboratoire : Lab-STICC, UMR CNRS 6285, Équipe DECIDE Pôle : CID

Thèse réalisée au sein de la Chaire de cyberdéfense des systèmes navals

Directeur de thèse : John Puentes

Soutenue le 7 Décembre 2017

Jury :

Éloi BOSSÉ, Chercheur, McMaster University (Rapporteur)
Joaquin GARCIA-ALFARO, Professeur, Télécom SudParis (Rapporteur)
John PUENTES, Maître de Conférences, HDR, IMT Atlantique (Directeur de thèse)
David BROSSET, Maître de Conférences, École Navale (Encadrant)
Martine COLLARD, Professeur des universités, Université des Antilles (Examinatrice)
Michaël HAUSPIE, Maître de Conférences, HDR, Université de Lille 1 (Examinateur)
Philippe LEROY, Spécialiste Senior Cyber, Thales Communications & Security (Invité)
Patrick HÉBRARD, Responsable Recherche et Innovation, NAVAL Group (Invité)

Table des matières

Table des matières **i**

Table des figures **v**

Liste des tableaux **vii**

Introduction **1**

Contexte de la thèse 1

Contexte de la problématique 2

Objectifs 3

Contributions 4

Plan du manuscrit 6

I État de l'art **9**

I.1 Qualité des données et qualité de l'information 10

I.1.1 Pyramide DIKW 10

I.1.2 L'étude de la qualité 12

I.1.3 Processus d'amélioration de la qualité 14

I.2 Détection d'anomalies 15

I.2.1 Classification 16

I.2.2 Modèles statistiques 18

I.2.3 *Clustering* 21

I.2.4 Techniques spectrales . . . 22
I.2.5 Théorie de l'information . . . 22
I.2.6 Apprentissage automatique . . . 23
I.3 Cyberattaques . . . 24
I.3.1 Définition de cyberattaques . . . 24
I.3.2 Évolution des cyberattaques . . . 26
I.3.3 Cibles pour les cyberattaques . . . 28
I.3.4 Attaques connues . . . 29
I.4 Systèmes cyber-physiques . . . 31
I.4.1 Définition . . . 31
I.4.2 Types de CPS . . . 32
I.4.3 Description d'un CPS . . . 33
I.4.4 Différences entre un CPS et un Système d'Information classique . . . 34
I.4.5 Systèmes navals . . . 36
I.4.6 Sécurité des CPS . . . 41
I.5 Conclusion . . . 44

II Méthodologie pour la mesure de qualité 47
II.1 Définitions contextuelles . . . 48
II.2 Définition des entités . . . 49
II.2.1 Données . . . 49
II.2.2 Information . . . 50
II.2.3 Connaissance . . . 51
II.2.4 Intelligence . . . 51
II.3 Mesure de la qualité à différents niveaux dans les CPS . . . 52
II.3.1 Qualité des données . . . 54
II.3.2 Qualité de l'information . . . 55

II.3.3 Qualité de la connaissance 58
II.3.4 Qualité de l'intelligence 59
II.3.5 Méthodologie d'évaluation de la qualité 60
II.4 Détection des anomalies 62
II.4.1 Hypothèses 62
II.4.2 Approche de détection et catégorisation 63
II.4.3 Implémentation de l'approche 65
II.5 Conclusion 69

III Application de la méthode **71**
III.1 Introduction 72
III.2 Système cyber-physique de deux cuves 72
III.2.1 Description 73
III.2.2 Mesure de la qualité 79
III.2.3 Détection d'anomalies (actes malveillants) 83
III.2.4 Catégorisation d'anomalies (actes malveillants) 88
III.3 Drones aériens 89
III.3.1 Description 92
III.3.2 Mesure de la qualité 95
III.3.3 Détection d'anomalies (actes malveillants) 101
III.4 Comparaison et positionnement par rapport à d'autres méthodes 105
III.4.1 Comparaison avec d'autres méthodes de mesure de la qualité 105
III.4.2 Positionnement de l'approche par rapport aux autres méthodes de détection d'anomalies 107
III.5 Conclusion 108

IV Conclusion générale et perspectives **109**
IV.1 Problématique 109

IV.2 Travaux réalisés 110
IV.3 Discussion 112
IV.4 Travaux futurs 114

Liste de publications **117**

IV.5 Article de revue 117
IV.6 Articles de conférence 117

Bibliographie **119**

Annexes **133**

A *Dataset* pour le cas d'étude du système des cuves **135**

A.1 Détails de la plate-forme 135
A.1.1 Composants 135
A.2 Dataset 136
A.2.1 Protocole 136
A.2.2 Structure 138

B Acronymes **139**

Table des figures

1 Diagramme représentant les principaux domaines utilisant les systèmes cyber-physiques (CPS). 2

2 Représentation schématique des différents domaines étudiés et leurs axes d'application. 4

3 Description des relations entre les chapitres du manuscrit de thèse. 6

I.1 Pyramide de relation entre données, informations et connaissances. 11

I.2 Apparition de la qualité dans chaque domaine par ordre chronologique. 12

I.3 Le processus d'amélioration continue de la qualité. 14

I.4 Techniques pour la détection d'anomalies et leur catégorisation. 16

I.5 Fonction de distribution gaussienne avec les probabilités correspondantes en choisissant d'intervalles de $\pm\sigma$, $\pm 2\sigma$ et $\pm 3\sigma$. 19

I.6 Évolution des compétences de l'attaquant versus la sophistication de l'attaque [LC09]. 27

I.7 Évolution des cyberattaques [PWC15, PWC16, PWC17]. 27

I.8 Commerce des 28 pays européens par moyen de transport en 2014[Com16]. . . 37

I.9 Message de Maersk sur Twitter en reconnaissant une cyberattaque affectant ses systèmes informatiques. 41

II.1 Schéma des catégories des dimensions de la qualité de l'information. 57

II.2 Méthodologie de mesure de la qualité dans le processus de définition d'une connaissance. 61

II.3 Schéma explicatif de l'approche proposée pour la détection et catégorisation d'anomalies. 64

II.4 Exemple de connexion de modules CPS d'un bateau [Nat04]. 66

II.5 Processus d'amélioration de la qualité de la méthodologie proposée. 68

III.1 Schéma de la plate-forme des cuves. 74

III.2 Diagramme du réseau de la plate-forme. 74

III.3 Représentation des données pendant un fonctionnement normal. 76

III.4 Photo du scénario de la mesure du capteur à ultrason bloquée et son impact sur le signal. 78

III.5 Photo du scénario de 7 objets flottants introduits dans la cuve principale et son impact sur le signal. 78

III.6 Évolution dans le temps de la mesure réalisée par le capteur à ultrason et de ses mesures de qualité respectives (précision réelle, précision de la source, opportunisme, confiance et cohérence). 84

III.7 Effet des objets flottants sur la précision réelle du capteur à ultrason et représentation de l'AL. 85

III.8 Histogramme de la précision réelle lorsque la valeur mesurée est plus grande que $3000u$. 86

III.9 Effet du capteur à ultrason bloqué sur la précision réelle et représentation de l'AL. 87

III.10 Effet d'une attaque DoS sur l'opportunisme et représentation des ALs. 88

III.11 « *Parrot rolling spider* » (à gauche) et « Crazyflie 2.0 » (à droite) 92

III.12 Exemple de séries temporelles des mesures produites par les capteurs du drone Parrot. 93

III.13 Mesure de l'opportunisme d'un vol du parrot. 100

III.14 Mesure de l'opportunisme pendant un vol du crazyflie (en haut). Histogramme de l'opportunisme mesuré pendant le vol (en bas). 100

III.15 Mesures du baromètre pour deux vols (un vol par colonne) et leur évaluation de la précision réelle (dans le temps et histogramme). 101

Liste des tableaux

I.1 Les huit méthodologies les plus citées pour la mesure de la qualité de l'information par ordre chronologique. 13

I.2 Niveaux d'un CPS selon le standard ANSI/ISA-95. 34

I.3 Différences entre un SI traditionnel et un CPS. 35

I.4 Priorités dans la protection des systèmes. 35

I.5 Systèmes navals et risques associés 40

I.6 Couches de sécurité d'un CPS. 42

II.1 Dimensions Intrinsèques de la Qualité de l'Information dans les CPS 56

II.2 Dimensions Contextuelles de la Qualité de l'Information dans les CPS 57

II.3 Dimensions Extrinsèques de la Qualité de l'Information dans les CPS 58

II.4 Aspects de la Qualité de l'intelligence dans les CPS 60

III.1 Catégorisation des anomalies identifiées pour le capteur à ultra-son à partir des éléments impactés. 90

III.2 Catégorisation des anomalies identifiées pour les flotteurs à partir des éléments impactés. 91

III.3 Catégorisation des anomalies réseau identifiées à partir des éléments impactés. 91

III.4 Catégorisation de types d'anomalies en fonction du facteur d'origine. 92

III.5 Catégorisation des anomalies identifiées pour les drones à partir des éléments impactés. 102

III.6 Comparaison avec d'autres méthodes pour la mesure de la qualité 105

A.1 Composants installés sur le sous-système. 135

A.2 Registres du PLC. 136

A.3 Fichiers qui composent le *dataset* des cuves. 137

Introduction

Sommaire

Contexte de la thèse **1**
Contexte de la problématique **2**
Objectifs **3**
Contributions **4**
Plan du manuscrit **6**

Cette introduction a pour but d'expliquer le sujet de la thèse tout en le positionnant par rapport aux domaines scientifiques connexes. Pour cela, le contexte, la problématique et les objectifs attendus sont présentés. Les contributions scientifiques apportées sont résumées dans la dernière partie.

Contexte de la thèse

De nombreux domaines comme le transport, l'industrie, les villes intelligentes et l'internet des objets utilisent des systèmes de plus en plus informatisés. Ces systèmes apportent un contrôle à distance et la réalisation d'actions complètement ou partiellement automatisées selon différentes situations définies. Les solutions utilisées dans ces domaines sont composées de plusieurs sous-systèmes qui interagissent avec l'environnement en prenant des mesures à l'aide de capteurs et en exécutant des réponses avec des automates et des actionneurs. Ces systèmes sont appelés systèmes cyber physiques (CPS) en raison du lien créé entre le mode physique et l'informatique. C'est pour cela que ces systèmes représentent aujourd'hui un atout majeur dans ces nouveaux domaines comme l'indique la Figure 1 mais aussi de multiples défis.

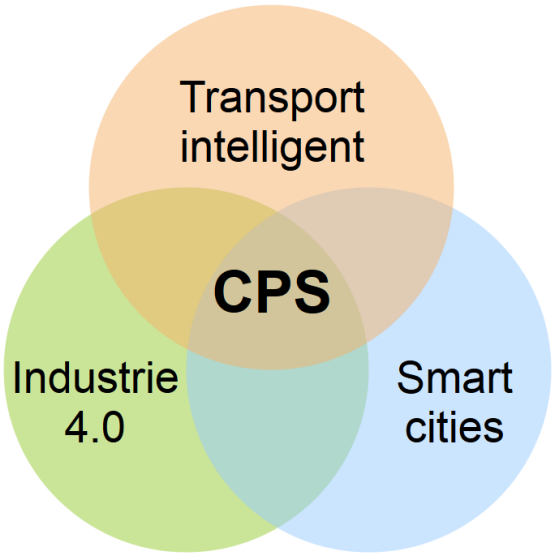

FIGURE 1: Diagramme représentant les principaux domaines utilisant les systèmes cyber-physiques (CPS).

Les CPS sont composés de sous-systèmes de différentes natures comme des capteurs, des systèmes de contrôle et des modules de communication. Chacun de ces composants va générer des flux de données dont l'objectif principal est de contrôler et surveiller un processus en gardant la cohérence des valeurs mesurées. Des objectifs complémentaires très variés sont également visés comme par exemple la sécurité ou l'aide à la décision. Ces systèmes en constante évolution intègrent souvent des connexions vers l'extérieur permettant la surveillance et le contrôle à distance. Quand ils présentent ces connexions, ils deviennent un cas particulier de CPS appelé SCADA (*Supervisory Control And Data Acquisition*). Ces systèmes sont très utilisés dans l'industrie, et c'est pour cela que le terme SCADA est fréquemment utilisé comme synonyme de CPS voire d'automate industriel.

De par leur importance et leur criticité, la sécurité de ces systèmes devient un enjeu majeur. Il faut assurer, à tout instant, leur disponibilité, leur intégrité, leur confidentialité et leur traçabilité. Ces critères de classification de l'information sont toujours utilisés pour la sécurité des systèmes informatiques et sont définis dans la norme ISO/CEI 27001 [ISO05].

Contexte de la problématique

Lors de leur phase de conception, les CPS ont cependant été faiblement sécurisés du fait de la grande importance donnée au fonctionnement optimal et performant. Par conséquent,

la rapidité de traitement fut au cœur de la conception au détriment de la sécurité. La sûreté de fonctionnement prenant le pas sur la sécurité. Actuellement, les systèmes traditionnels de détection et de réaction contre les cyber-menaces ne sont pas optimaux, notamment pour les CPS. Les raisons sont, par exemple, la variabilité des flux de données, la limitation des réponses en temps réel et la vulnérabilité des systèmes en réseau. Ainsi, l'intégration de systèmes pour le contrôle de la sécurité présente une complexité particulière en raison de la singularité des sous-systèmes concernés.

Conventionnellement, l'ensemble de la sécurité des CPS était basée sur l'isolation des systèmes, sur l'aspect personnalisé de ces systèmes en fonction des besoins et sur le fait que leurs spécifications étaient propriétaires et protégées par des contrats de non-diffusion. Depuis une décennie, les CPS sont de plus en plus interconnectés, accessibles de l'extérieur et utilisent des standards de l'industrie ou des réseaux informatiques. Leur protection devient ainsi plus difficile.

Dans la protection des CPS, un prérequis indispensable est la capacité d'identifier des situations inconnues des systèmes. C'est pour cela que les systèmes de détection basés sur des signatures ne sont pas suffisants. Ces systèmes évaluent le comportement ou la situation des systèmes par comparaison en utilisant une base de signatures. Ainsi, un comportement anormal ou un virus inconnu ne pourront pas être détectés. Des alternatives à l'analyse de trafic réseau doivent être proposées afin de sécuriser au mieux ce type de système. La catégorisation des détections est souhaitable pour pouvoir différencier une anomalie environnementale d'une attaque et pouvoir donner une réponse adéquate. L'interdépendance des sous-systèmes des CPS ajoute une complexité à la problématique.

Objectifs

L'objectif de cette thèse est de définir une méthodologie de détection d'anomalies et de cyberattaques comme cas particulier de celles-ci sur les flux des capteurs, en mesurant la qualité des données et des informations qui en découlent. Grâce à cette détection, nous cherchons à éviter ou modifier les décisions basées sur des informations anormales selon le problème subi. À cet égard, une catégorisation des détections est fortement souhaitée.

La qualité des données et de l'information est de plus en plus étudiée, mais à notre connaissance actuellement aucune méthodologie n'est adaptée aux flux de données et d'information des CPS. Ainsi, une nouvelle méthodologie adaptée doit être créée pour ces systèmes.

Afin d'atteindre les objectifs de détection, nous partons de l'hypothèse que les anomalies et les cyberattaques, considérées comme anomalies provoquées, peuvent avoir un impact sur les mesures de qualité. De cette façon les méthodes d'évaluation de la qualité se portent comme des candidats pour la résolution de cette problématique.

Par conséquent, une méthode adaptée de mesure de qualité permettant la détection d'anomalies est l'objectif principal de notre recherche. Ainsi, les éléments de cette méthodologie doivent permettre d'apporter une information additionnelle sur les détections permettant d'enrichir les systèmes d'aide à la décision.

Contributions

Les contributions de cette thèse sont le résultat de l'union de trois domaines de recherche : l'étude des anomalies et des actes malveillants, l'étude de la qualité et l'étude des CPS. Les axes d'intersection, représentés dans la Figure 2, définissent les applications étudiées.

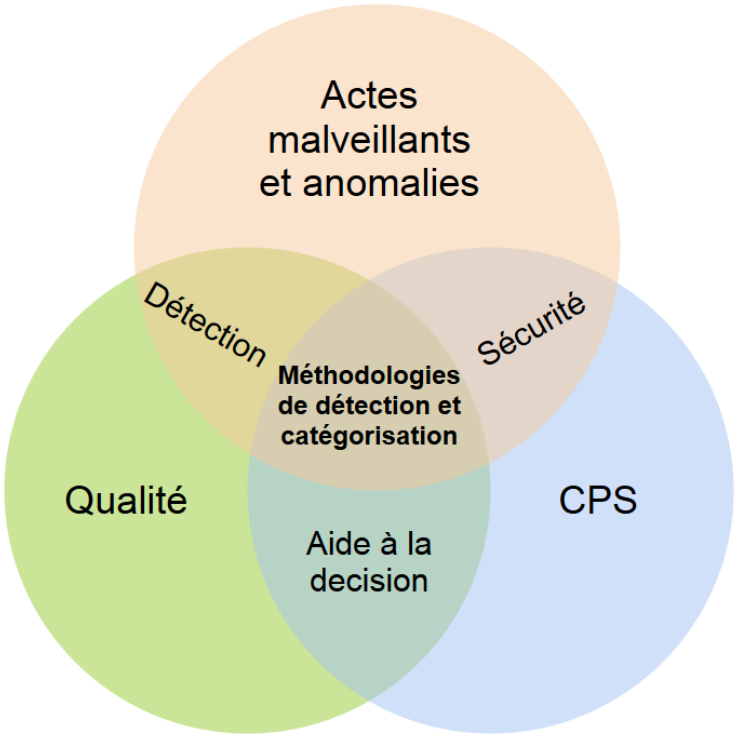

FIGURE 2: Représentation schématique des différents domaines étudiés et leurs axes d'application.

La première contribution est l'identification des éléments pertinents pour l'évaluation de la qualité dans les CPS. Cette identification permet d'étendre les approches de mesure de la qualité existantes vers d'autres domaines : les CPS, la détection d'anomalies et la cybersécurité. Ces éléments suggèrent une possibilité d'application générique dans les CPS des cas d'étude, nécessitant une adaptation à chaque système particulier.

La deuxième contribution est une méthodologie adaptée, potentiellement générique et originale pour la mesure de la qualité des flux créés par les sous-systèmes des CPS. Grâce aux éléments identifiés, cette approche est capable de mesurer la qualité de chaque flux à différents niveaux préalablement définis par la méthodologie. Cela permet d'alimenter les systèmes d'aide à la décision, ce qui pourra mener à une amélioration en continu de la qualité.

La troisième contribution est la définition d'états de fonctionnement normal dans les CPS. Grâce à eux, la qualité devient un indice pour la détection d'anomalies. Ainsi, les différents éléments de la qualité affectés par les anomalies peuvent définir un système de catégorisation des détections. La définition des valeurs normales des éléments de la qualité peut être partiellement automatisée grâce aux techniques existantes pour la détection d'anomalies. C'est pour cela que cette méthodologie de mesure de qualité a abouti à l'implémentation de systèmes semi-automatiques de détection d'anomalies dans les CPS. Les cyberattaques considérées comme des comportements anormaux d'un système pourront ainsi être détectées.

Pour la validation des approches de mesure de qualité et de détection, nous avons défini deux protocoles expérimentaux. Les deux systèmes étudiés concernent deux systèmes critiques qui présentent des risques importants régulièrement mis en lumière dans l'actualité. Le premier cas d'étude est constitué d'un système SCADA de deux cuves qui peuvent représenter aussi bien un système d'alimentation d'eau comme un château d'eau ou bien un ensemble de réservoirs de carburant sur un navire. La quatrième contribution est le partage du *dataset* avec les fichiers de logs utilisés pour son étude à fin que d'autres chercheurs puissent les réutiliser. Le deuxième protocole expérimental est réalisé à l'aide de deux drones aériens possédant des caractéristiques différentes. L'étude de ces deux systèmes différents montre comment cette approche originale peut être généralisée. Le partage des données utilisées pour ce cas d'étude permet de les réutiliser dans la validation d'autres approches proposées par d'autres chercheurs de différents domaines par exemple la détection de pannes et d'anomalies et le fait de pouvoir les comparer.

Plan du manuscrit

Le manuscrit est composé de quatre chapitres (Figure 3).

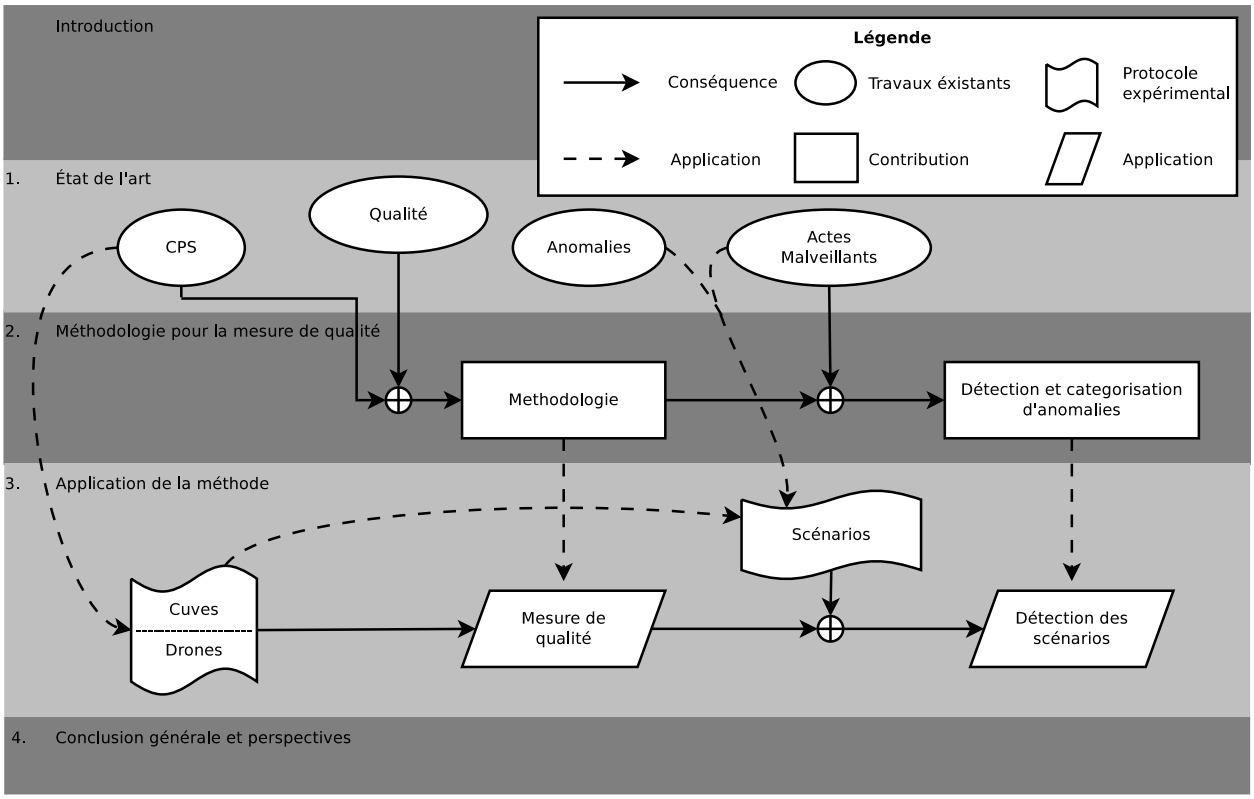

FIGURE 3: Description des relations entre les chapitres du manuscrit de thèse.

Le premier chapitre présente l'état de l'art des différents domaines qui sont explorés : la qualité et la détection d'actes malveillants à partir d'anomalies. Les caractéristiques et particularités des CPS sont également exposées.

Dans le deuxième chapitre, une méthodologie d'évaluation de la qualité d'application générale est proposée pour les sous-systèmes des CPS. L'objectif de cette nouvelle méthodologie est d'être utilisé pour la détection d'anomalies et de pouvoir détecter des actes malveillants et des dysfonctionnements.

Dans le troisième chapitre, deux cas d'étude sont présentés : un système SCADA composé de deux cuves et deux drones aériens. Grâce aux données créées en simulant différents scénarios, la méthodologie proposée est appliquée, d'abord la partie évaluation et ensuite la partie détection. Une détection à partir d'une évaluation multidimensionnelle de la qualité permettra d'introduire la catégorisation des détections.

Le quatrième chapitre conclut le manuscrit en discutant des résultats obtenus, leur généralisation et en présentant les perspectives envisagées du travail réalisé.

Chapitre I

État de l'art

Sommaire

I.1 Qualité des données et qualité de l'information 10
I.1.1 Pyramide DIKW 10
I.1.2 L'étude de la qualité 12
I.1.3 Processus d'amélioration de la qualité 14
I.2 Détection d'anomalies 15
I.2.1 Classification 16
I.2.2 Modèles statistiques 18
I.2.3 *Clustering* 21
I.2.4 Techniques spectrales 22
I.2.5 Théorie de l'information 22
I.2.6 Apprentissage automatique 23
I.3 Cyberattaques 24
I.3.1 Définition de cyberattaques 24
I.3.2 Évolution des cyberattaques 26
I.3.3 Cibles pour les cyberattaques 28
I.3.4 Attaques connues 29
I.4 Systèmes cyber-physiques 31
I.4.1 Définition 31
I.4.2 Types de CPS 32
I.4.3 Description d'un CPS 33
I.4.4 Différences entre un CPS et un Système d'Information classique . 34
I.4.5 Systèmes navals 36
I.4.6 Sécurité des CPS 41
I.5 Conclusion 44

Ce chapitre décrit les travaux de recherche antérieurs sur la qualité des données, la

qualité de l'information et la détection d'anomalies. Dans cette thèse, les cyberattaques qui constituent des anomalies particulières sont étudiées. La dernière partie de ce chapitre leur est consacrée.

I.1 Qualité des données et qualité de l'information

Quand les systèmes informatiques sont apparus, les logiciels n'étaient pas considérés comme une propriété avec une valeur concrète au sens économique, même s'ils étaient considérés comme un objet d'importance. Avec leur développement et leur complexité, les logiciels ont atteint leur statut actuel. Les données ont généralement été sous-estimées bien que les connaissances étaient considérées comme un point stratégique. Aujourd'hui les données ont gagné en importance et quelques entreprises font d'elles leur marché. Google est un exemple bien connu de l'importance actuelle des données. Ainsi, l'automatisation des métiers a fait que les systèmes industriels ont profité des applications des CPS pour évoluer vers l'industrie 4.0, domaine dans lequel les décisions se basent sur les données de capteurs. Cela a donné lieu à l'apparition de nouvelles problématiques et donc de nouveaux défis pour la recherche.

I.1.1 Pyramide DIKW

Les définitions des données, information et connaissances ne sont pas généralement claires et uniques. Une des façons les plus générales de représenter cette hiérarchie est avec une pyramide ou à différents niveaux de cognitive avec augmentation de sémantique. Une des plus réputées est la pyramide DIKW (*Data-Information-Knowledge-Wisdom*) qui représente la relation entre les données, les informations, les connaissances et l'intelligence [1] (Figure I.1) [Zel05]. Bien qu'il existe des travaux rigoureux sur la problématique [Zel05], il existe une étude qui fait une comparaison des différentes définitions proposées pour conclure qu'il n'y a pas une signification commune pour chaque terme [Zin07]. Cependant, la pyramide DIKW s'adapte parfaitement aux éléments des CPS (données binaires, informations, résultats de traitements et algorithmes) et donc, elle sera appropriée pour les travaux réalisés dans cette thèse.

1. La traduction du terme anglais « *wisdom* » n'est pas évidente, car le terme est également sujet à polémiques en anglais. Un choix a été réalisé entre plusieurs termes selon le contexte comme « compréhension » et « sagesse ».

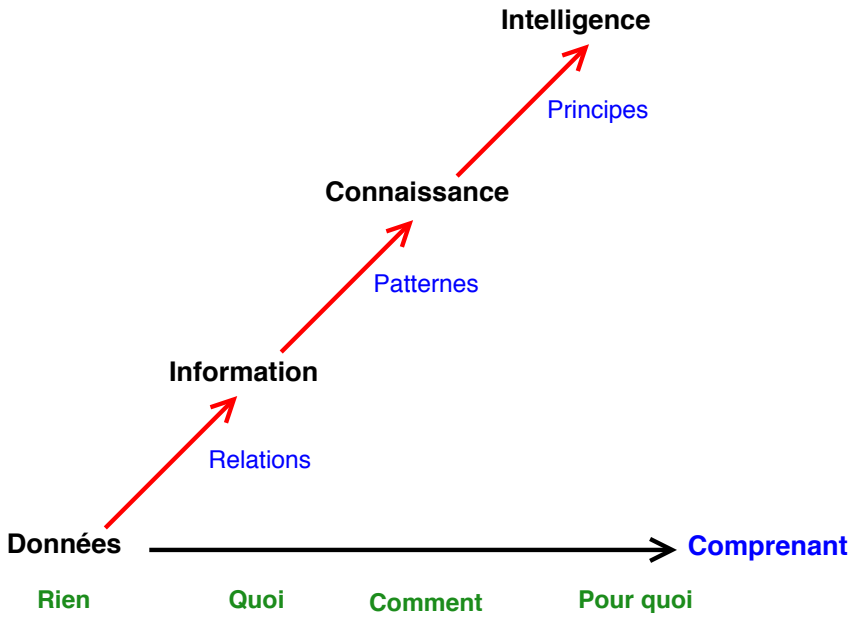

FIGURE I.1: Pyramide de relation entre données, informations et connaissances.

Sur la figure I.1 les définitions des différents niveaux peuvent être résumées à partir de leurs niveaux de compréhension. D'abord, les données sont des entités où **rien** n'est compris (*know-nothing*). Une fois qu'une série de relations fait une mise en contexte des données, l'information peut être extraite et nous pouvons savoir de **quoi** il s'agit (*know-what*). Une connaissance est créée quand un motif a été défini pour un système et par conséquent nous savons **comment** l'information de l'étape précédente peut être utilisée (*know-how*). Quand le principe de fonctionnement des connaissances est extrait, cela veut dire, l'explication de leur **pour quoi**, une intelligence est définie (*know-why*).

Pour illustrer ces concepts, imaginons qu'il existe une connaissance des températures d'un type de moteur en fonction du contexte de fonctionnement. Un capteur de température enregistre une donnée avec une valeur binaire qui représente la température et la transmet via un réseau. Une fois que cette donnée est récupérée, on la met en contexte avec les spécifications du capteur, de la codification et du protocole utilisés. Le résultat de ce processus est l'information que le moteur est à 20^{o}C. Grâce à l'intelligence, nous pouvons extraire la connaissance indiquant que le moteur n'est pas en fonctionnement parce qu'il est « froid ». Dans ce cas, l'intelligence est composée par deux intervalles qui indiquent les températures du moteur en fonction de son état.

I.1.2 L'étude de la qualité

La qualité a été étudiée depuis longtemps dans différents domaines. Les trois domaines qui s'y sont intéressés le plus sont les systèmes d'information du management (MIS), les systèmes d'information web (WIS) et les systèmes de fusion d'informations (IFS) [TLKLC13]. Sur la Figure I.2, l'apparition de différents concepts sur la qualité au cours des années est représentée pour chacun des domaines qui l'a introduit. Une des premières fois que la qualité a été citée fut en 1985 sur les MIS. Cette recherche définissait la qualité des données pour les systèmes de décision basés sur des systèmes d'information avec plusieurs données comme entrée [BP85]. Une décennie plus tard, l'imperfection de l'information a été étudiée pour les IFS [Sme97]. Ainsi, les chercheurs qui étudiaient les MIS commencèrent à se demander directement « qu'est-ce que la qualité ? » dans des contextes précis [WS96] et ainsi plus tard, ils ont commencé à chercher des méthodologies pour sa mesure [MWLZ09]. Cette première définition de la qualité pour les MIS a fait émerger la discipline. Les WIS, conçus pour l'utilisation de grandes bases de données et l'exploration de celles-ci, étudient la même problématique depuis les années 1990 [Che01, Nau01]. Nous pouvons également apprécier que la qualité a une importance majeure pour les systèmes multi sources [RB10] et plus précisément pour les systèmes multi capteurs [KL09b] et leur fiabilité [RN04]. Actuellement, le Big Data se positionne comme un défi interdisciplinaire pour la qualité des données [CZ15].

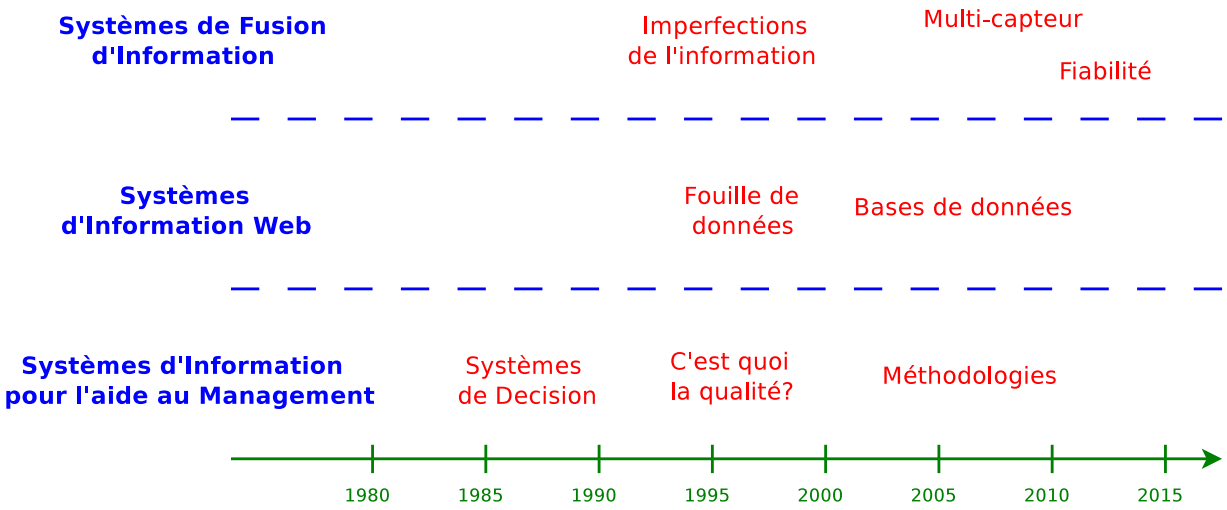

FIGURE I.2: Apparition de la qualité dans chaque domaine par ordre chronologique.

Il existe plusieurs méthodologies différentes pour la mesure de la qualité de l'information. Sur le tableau I.1 les plus citées sont listées. Il existe plusieurs études qui les examinent avec le but de les comparer et les mettre en perspective [WP10].

Acronyme	**Nom**	**Référence**
TDQM	Total Data Quality Management	[WS96]
TIQM	Total Information Quality Management	[Eng99]
COLDQ	Cost-effect Of Low Data Quality	[Los01]
AIMQ	A methodology for Information Quality Assessment	[LSKW02]
DQA	Data Quality Assessment	[PLW02]
DaQuinCIS	Data Quality in Cooperative IS	[SVM+04]
QAFD	Quality Assessment on Financial Data	[DA04]
CDQ	Comprehensive methodology for Data Quality Management	[BS06]

TABLE I.1: Les huit méthodologies les plus citées pour la mesure de la qualité de l'information par ordre chronologique.

Les différences principales entre les méthodologies citées se focalisent sur les domaines d'application. TDQM [WS96] utilise une catégorisation des dimensions de la qualité identifiées par un groupe de consommateurs pour définir la qualité des données. TIQM [Eng99] présente la qualité de l'information comme un point important pour réduire les coûts et incrémenter les bénéfices dans les modèles de management. AIMQ [LSKW02] est une généralisation de méthodologies existantes au moment de sa publication. DQA [PLW02] propose différentes métriques qui peuvent servir à la mesure de la qualité. COLDQ [Los01] examine la qualité avec une méthodologie pour répondre aux besoins de certaines entreprises. Aussi, la problématique de la mesure de la qualité par techniques collaboratives a été abordée par DaQuinCIS [SVM+04]. Pour les systèmes d'information de finances, la méthodologie QAFD [DA04] étudie comment améliorer la qualité des données les plus pertinentes. Un autre objectif de la qualité est l'amélioration de l'efficience et l'efficacité des modèles organisationnels et de management et elle est traitée par la méthodologie CDQ [BS06].

La qualité des données et de l'information dans un système a été rendue nécessaire au moment même que les systèmes sont apparus. Par exemple, l'étude de cette qualité peut être réalisée sur l'information d'une organisation [Wan98]. L'étude de la qualité pour les CPS est encore un sujet à l'étude. Les publications existantes n'abordent que des problématiques très concrètes par exemple la surveillance d'un barrage d'eau avec un réseau de capteurs sans fil [GL15]. Aussi, la qualité des sources d'information des capteurs a été étudiée comme source possible d'information dans les systèmes de fusion d'information [RB10]. D'autres recherches ont analysé la qualité des informations obtenues par un réseau de capteurs pour les

intégrer dans un modèle de *business* plan [KL09b]. Quelques travaux abordent des éléments générales de la qualité qui peuvent être utilisés par un ensemble de capteurs du même type, par exemple par des capteurs médicaux [PMLL13]. La propagation de qualité a été aussi une problématique traitée sur les CPS avec une application concernant un système de détection radar [Tod14]. La mesure de la qualité dans les CPS est encore un défi qui doit être exploré à cause entre autres, du manque de méthodologies générales pour son évaluation [SZ15].

I.1.3 Processus d'amélioration de la qualité

Parmi ces méthodologies, TDQM est prise généralement comme la référence. Cette méthodologie montre le processus d'amélioration continue de la qualité qui est un des défis les plus présents en ce qui a trait à la qualité. Ce processus est représenté par la Figure I.3. Il consiste en quatre étapes qui se répètent en boucle :

1. La première étape est de **définir** la qualité pour le cas d'un système particulier et comment la mesurer.
2. La deuxième étape consiste à **mesurer** tous les attributs identifiés de la qualité.
3. Ensuite, une **analyse** est réalisée à partir des mesures pour identifier les dimensions déficitaires.
4. Finalement, une fois identifiée la source du problème, elle sera corrigée dans le but d'**améliorer** la qualité du système.

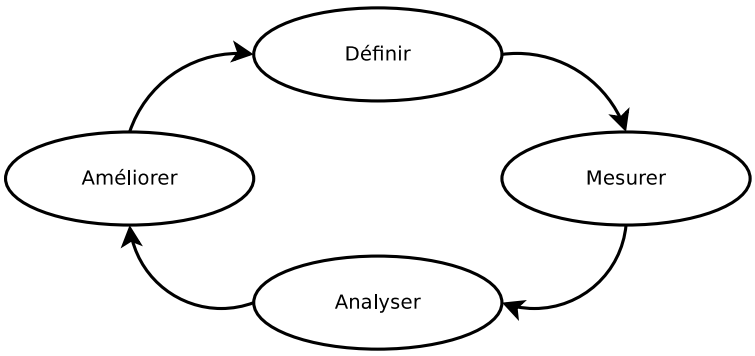

FIGURE I.3: Le processus d'amélioration continue de la qualité.

I.2 Détection d'anomalies

La détection d'anomalies est un domaine interdisciplinaire qui a été étudié depuis longtemps. Les premiers travaux qui cherchaient des valeurs aberrantes (*outliers* en anglais) ou anomalies dans des données avec des méthodes statistiques datent du 19^{e} siècle [Edg87]. Les objectifs poursuivis varient en fonction du domaine de recherche. Par exemple, la détection d'une anomalie dans le domaine de la santé peut être un indice pour le diagnostic d'une maladie, par contre dans le secteur bancaire, une anomalie dans les transactions peut signaler une fraude.

Plusieurs définitions de détection d'anomalies existent en fonction du domaine. Une définition générale qui peut être utilisée dans le cas des systèmes cyber-physiques est :

> La **détection d'anomalies** consiste à identifier un schéma dans les données qui représente un comportement inattendu. [2] [CBK09]

Avec cet objectif, plusieurs dimensions de ces données peuvent être étudiées. Chacune de ces dimensions représente une caractéristique de la donnée. C'est pour cela que souvent les données sont représentées comme des points sur un hyperespace de n dimensions. Cette représentation sera utilisée dans les sections suivantes pour l'introduction des méthodes utilisées.

Les systèmes critiques [3] s'intéressent spécialement à cette problématique. Par exemple, les systèmes navals doivent détecter les anomalies dans leurs systèmes parce qu'ils travaillent dans un environnement rude et imprévisible. D'autres systèmes avec des caractéristiques similaires ont été étudiés pour la détection d'anomalies par exemple les véhicules spatiaux [HB95], les centrales nucléaires [RC08] et les voitures autonomes [PS14]. Ces domaines d'application présentent plusieurs méthodologies de détection d'anomalies avec des caractéristiques et des performances différentes. Dans les sections suivantes, nous allons parcourir les différentes méthodes de détection d'anomalies existantes dans la littérature. Un aperçu de ces méthodes et leurs catégories est présenté dans la Figure I.4.

2. Anomaly detection refers to the problem of finding patterns in data that do not conform to expected behavior.

3. Un système est considéré critique si son mauvais fonctionnement peut avoir un impact important sur la sécurité ou la vie des personnes, des entreprises ou des biens

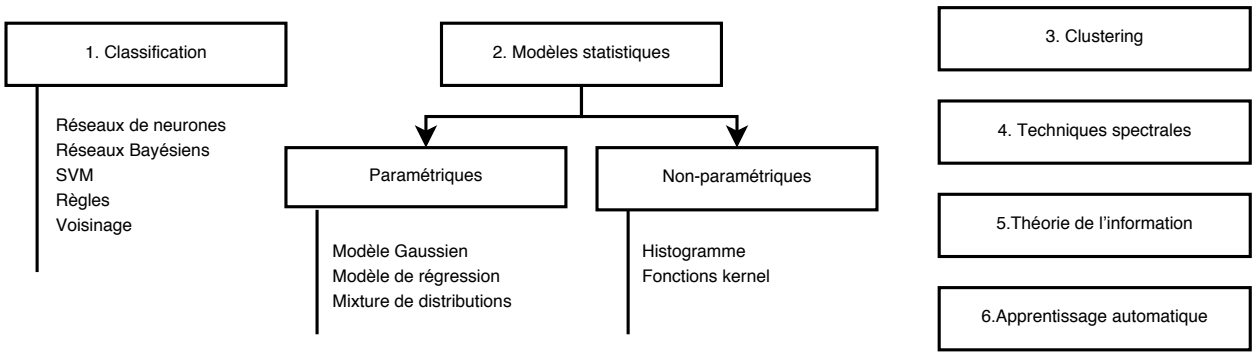

FIGURE I.4: Techniques pour la détection d'anomalies et leur catégorisation.

I.2.1 Classification

Les techniques de classification consistent à assigner une étiquette à une donnée pour la trier dans une des catégories prédéfinies. Ces méthodes sont adaptées pour la détection d'anomalies en créant une catégorie qui prend en considération les caractéristiques des anomalies ou en définissant comme anormales les données qui ne correspondent avec aucune classe. Pour atteindre cet objectif, plusieurs méthodes ont été étudiées.

Les **réseaux de neurones** sont un ensemble d'algorithmes qui essayent de répliquer le fonctionnement des neurones présents dans un cerveau. Il existe plusieurs méthodes qui peuvent être séparées en deux familles : les applications statistiques et l'intelligence artificielle ou AI (*Artificial Intelligence*). Dans la première famille, on peut trouver des méthodes de classification topologique par exemple les réseaux de Kohonen. Dans la famille de l'intelligence artificielle, où on trouve le *deep learning* dont les solutions ont pur but de donner un aspect rationnel aux réseaux. Les méthodes de détection utilisant les réseaux neurones comportent deux phases : une phase d'apprentissage et une phase de fonctionnement. Dans la phase d'apprentissage, les neurones vont « apprendre » les différentes classes normales qui existent dans le *dataset*[4] d'entraînement. Plus tard, dans la phase de fonctionnement le réseau sera capable de classifier les données d'entrée comme normales ou anormales. Cette technique est utilisée, entre autres, pour la détection d'intrusion réseau et la détection de fraudes [LG12]. Il existe plusieurs types de réseaux de neurones qui peuvent être utilisés dans ce but [ST12].

Les **réseaux bayésiens** sont des méthodes qui se servent d'un graphe modélisant une structure en calculant la probabilité d'être dans une situation particulière. Pour le cas d'étude, cette structure peut représenter un sous-système ou un système complet. Ces proba-

4. Anglicisme pour jeu de données.

bilités sont stockées dans des tables de probabilité appelées aussi paramètres, qui prennent en compte chaque variable du système étudié. Un des domaines d'application est la détection d'anomalies dans les réseaux de capteurs parce qu'ils peuvent réaliser une classification en temps réel [HMA07]. Aussi ils sont employés pour la détection d'anomalies dans la surveillance maritime [GWCC08].

Les **machines à vecteurs de support** ou SVM (*Support Vector Machine*) sont un ensemble de techniques utilisées par le *machine learning*[5] pour la classification de données. Ce sont des méthodes qui ont besoin d'un apprentissage supervisé. Le principe est de trouver un hyperplan qui soit capable de séparer un ensemble de données en deux groupes. Parfois, il n'existe aucune solution d'hyperplan avec cette condition, donc une transformation avec une fonction *kernel* est réalisée avec le but de définir un nouvel espace où il en existe au moins une. Quand il y a plusieurs solutions possibles, on privilégiera celle qui optimise la séparation entre les deux catégories en laissant une marge parce qu'elle minimisera les fausses détections. Ces méthodes peuvent être appliquées à la détection d'intrusions machine [HLV03] et d'intrusions réseau [FZHH14].

La détection d'anomalies par des techniques basées sur **règles** consiste à décrire le fonctionnement normal d'un système par un ensemble de règles. Quand un comportement n'est pas décrit par une des règles, il sera considéré comme un comportement anormal. Les différentes techniques qui ont été proposées diffèrent par l'algorithme qui crée ces règles, par exemple en utilisant des arbres de décision ou RIPPER (*Repeated Incremental Pruning to Produce Error Reduction*) [Coh95]. Ces systèmes de détection sont utilisés dans les réseaux informatiques [IKP95], pour la découverte de fraudes bancaires [BLH99], dans les réseaux de capteurs [BGS+13] et les systèmes cyber-physiques [CCBGA05].

D'autres solutions sont les **techniques de voisinage**. La méthode des k plus proches voisins ou "knn" (*k-nearest neighbor*) est une méthode supervisée de classification basée sur la probabilité qu'un point appartient à une catégorie. Cette probabilité est calculée en fonction de la catégorie de ses k voisins plus proches. Les voisins sont établis pendant une phase d'apprentissage. La performance de cet algorithme est très liée à la qualité du *dataset* utilisé dans cette phase. Cette qualité varie en fonction de la densité des points et la qualité de leur classification. Une variante de ces méthodes prend en compte la densité des points dans les alentours du point étudié au lieu d'analyser une quantité k de voisins. Ces techniques ont démontré leur intérêt dans la détection d'intrusions réseau [LV02] et le domaine des réseaux capteurs sans fil [RLPB06]. Parfois, cette technique est enrichie avec un autre type

5. Apprentissage automatique.

de méthode comme les modèles gaussiens [KNN16].

Parfois, les systèmes de détection d'anomalies implémentent plusieurs méthodologies pour atteindre un objectif. Par exemple, les techniques de voisinage, les réseaux de neurones et les SVM peuvent être enchaînées pour faire une meilleure détection d'anomalies réseaux [CR13]. Cette fusion permet de profiter des meilleures caractéristiques de chaque technique, sans obtenir toutefois une performance très élevée.

I.2.2 Modèles statistiques

Les modèles statistiques vont considérer un système producteur de données comme un modèle stochastique. C'est pour cela, qu'une méthode de détection basée sur ces techniques va définir les anomalies comme des données qui n'ont pas été produites par un modèle stochastique particulier. En conséquence, les données normales vont se concentrer dans les régions de haute probabilité tandis que les anomalies vont se produire dans les régions de faible probabilité des modèles.

Les techniques de détections d'anomalies basées sur des modèles statistiques peuvent être paramétriques ou non paramétriques. La différence entre ces deux familles est que pour l'une le point de départ est la connaissance de la distribution, ayant besoin de chercher les paramètres pour un cas particulier, et que pour l'autre la distribution est inconnue.

Paramétriques

Dans ces techniques, une fonction de densité $f(x, \theta)$, où x sont les données observées et θ les paramètres de la distribution, est prédéfinie pour un *dataset* particulier. Afin d'obtenir la fonction de distribution des anomalies, la fonction inverse de $f(x, \theta)$ est calculée. Ces techniques peuvent être classifiées à partir du modèle choisi en : modèle gaussien, modèle de régression et mixture de distributions paramétriques.

Étant donné que souvent les données présentent des distributions gaussiennes, plusieurs méthodes assument par défaut que les données étudiées vont être résultat d'un **modèle gaussien**. Une étape avec des données d'entraînement est nécessaire pour calculer les paramètres de la fonction. Une fois que les paramètres ont été choisis, l'intervalle le plus probable définit les données normales et donc le reste des valeurs non compris dans cet intervalle est désigné comme anormal. Pour fixer les limites de l'intervalle la règle 68-95-99,7

est souvent utilisée. Cette règle prend comme référence pour l'intervalle la moyenne de la distribution et étends l'intervalle une, deux et trois fois la variance (sigma) vers les deux sens ($\mathbb{P}(\mu - n\sigma \leq x \leq \mu + n\sigma)$ pour n fois) ce qui englobe respectivement 68%, 95% et 99,7% des valeurs normales comme il est représenté dans la Figure I.5. Par exemple, dans l'application sur des réseaux de capteurs sans fil, une distribution gaussienne s'adapte mieux que d'autres distributions par exemple qu'une distribution uniforme [WFA13]. Par ailleurs, les modèles gaussiens complètent parfois d'autres techniques de détection comme k-moyennes [KNN16].

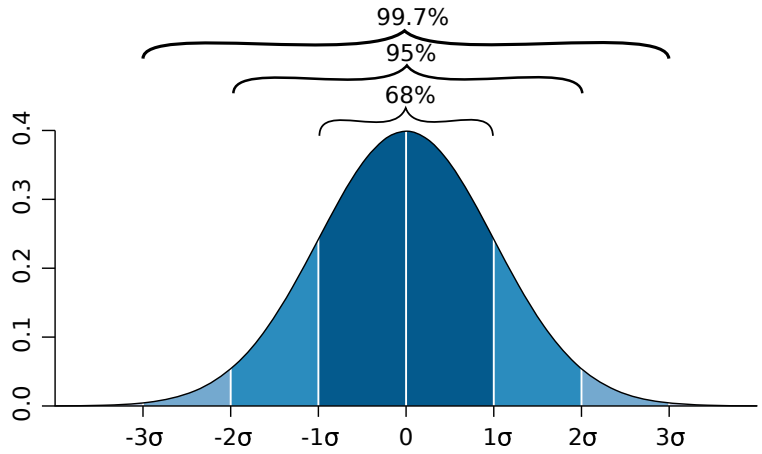

FIGURE I.5: Fonction de distribution gaussienne avec les probabilités correspondantes en choisissant d'intervalles de $\pm\sigma$, $\pm 2\sigma$ et $\pm 3\sigma$.

Les **modèles de régression** sont fréquemment utilisés pour faire une détection d'anomalies. Une fois que le modèle a été adapté à un système, l'évaluation des données fournira comme résultat un résidu qui représente le score de l'anomalie. Souvent les *datasets* qui sont utilisés pour calculer les paramètres du modèle contiennent des données anormales. Pour réduire leur impact pendant le choix de paramètres, un type de régression appelée robuste peut être utilisé. Ces modèles ont été utilisés depuis longtemps pour la détection d'*outliers* sur des séries temporelles [AC89]. Par la suite, ils se sont positionné comme une alternative appropriée pour la détection d'intrusions réseau [Wan05, MSJ10].

Ainsi, une **mixture de distributions** paramétriques présentées peut être plus performante dans quelques cas particuliers. Par exemple, deux distributions peuvent être utilisées pour étudier séparément la distribution de données normales et anormales. À partir de ces deux distributions un score d'anomalie peut être défini [Esk00, Agg05]. Aussi, il y a des solutions qui proposent une addition de plusieurs distributions connues pour représenter la distribution des données étudiées. Ce type de méthodologies a servi à implémenter des systèmes pour la détection d'intrusions dans les réseaux informatiques [YTWM00]. Cette

implémentation permet d'avoir une méthodologie qui s'adapte aux données non stationnaires avec un coût de computation faible, ce qui est communément un prérequis indispensable pour les systèmes de détection.

Non-paramétriques

Les systèmes de détection qui utilisent des modèles statistiques non paramétriques n'ont pas de paramètres définis *a priori*. Leur fonctionnement s'adapte à fur et à mesure aux données évaluées.

Les méthodes de détection par **histogramme**, aussi appelées de fréquence ou méthodes de comptage, sont parmi les méthodes les plus utilisées dues à leurs caractéristiques. Ces méthodes consistent à créer un histogramme à partir d'un *dataset* d'entraînement. Cela permet de vérifier si les valeurs à tester font partie ou non d'une des barres de l'histogramme créé et de déterminer si elles sont respectivement normales ou anormales. Leurs caractéristiques les plus importantes sont leur simplicité et leurs réponses en temps réel avec un bas coût de computation. Ces méthodes sont amplement utilisées dans plusieurs domaines de détection comme les fraudes téléphoniques [CEWB97], les intrusions machine [WFP99] et les intrusions réseau [CHMS09].

Ainsi que pour les techniques basées sur des SVM, des **fonctions *kernel*** peuvent être utilisées pour transformer les données, de telle façon que les fonctions de densité puissent faire une détection optimale. Par exemple, pour faire une détection d'intrusions réseau, une transformation avec des *kernels* gaussiens est réalisée avant d'utiliser une estimation par noyau [YC02].

Mixtes

Il existe des techniques qui, en fonction de leur configuration, peuvent être paramétriques ou non-paramétriques. Les **filtres de Kalman** en sont un exemple [Ard83]. Même si ces filtres sont généralement paramétriques, grâce à l'utilisation en parallèle de certaines techniques ils peuvent devenir non-paramétriques en adaptant automatiquement ces paramètres au contexte [HBS16].

Le filtre de Kalman est une technique utilisée avec différents objectifs par exemple la prédiction d'une valeur et la fusion de sources d'information. Ce type de filtre a eu une

importance majeure dans les cinquante dernières années. Une croyance générale affirme qu'il est l'une des raisons pour lesquelles l'arrivée à la lune a été possible [GA10]. Ce sont des filtres récursifs qui donnent une réponse en temps réel et c'est pour cela qu'ils se positionnent comme une option pour les systèmes critiques.

En plus des applications citées, les filtres de Kalman sont utilisés pour la détection d'anomalies et particulièrement pour la détection des cyberattaques [RHDCGA16]. Un exemple est la détection de pannes [KS07]. D'autre part, l'application d'un filtre non paramétrique est utilisée pour la détection de failles dans le software qui produisent par exemple un *memory overflow* [KL08].

I.2.3 *Clustering*

Le *clustering* est une technique de partitionnement de données qui consiste à diviser un ensemble de données en plusieurs sous-ensembles. Chacun de ces sous-ensembles contiendra des données avec des caractéristiques similaires. Ces techniques sont communément utilisées dans de nombreux domaines comme la médecine ou l'image satellitaire avec différents objectifs par exemple la segmentation, la classification et l'extraction de connaissances.

Les techniques existantes pour la détection d'anomalies en utilisant les méthodes de *clustering* partent de trois hypothèses. La première considère l'idée que les données normales vont appartenir à un *cluster* tandis que les anormales seront inclassables. La deuxième hypothèse assume que les données normales vont se placer près du centre d'un *cluster* tandis que les données anormales vont apparaître plus loin. La dernière hypothèse prend comme point de départ qu'on aura plus de données normales qu'anormales. C'est pour cela, que ces méthodologies vont considérer les *clusters* grands et denses comme normaux et les *clusters* petits et éparpillés comme anormaux.

Les méthodes de détection basées sur des techniques de *clustering* peuvent être utilisées avec une grande variété d'objectifs. Par exemple, elles sont utilisées pour la détection de fraudes avec cartes bancaires [BH01] d'une part, et l'implémentation de NIDS (*Network Intrusion Detection System*) d'autre part [LKT15, KBD+15].

I.2.4 Techniques spectrales

Les techniques qui cherchent des anomalies dans le spectre se basent sur l'idée que les anomalies peuvent être cachées quand on les regarde dans un espace déterminé. Pour les trouver, ces techniques réduisent l'espace dimensionnel à un sous-espace dans lequel les anomalies apparaissent d'une façon plus signifiante. Ces travaux cherchent à trouver les sous-espaces adéquats en utilisant différentes méthodes comme les projections ou l'imbrication de dimensions.

Pour réussir à trouver les dimensions les plus pertinentes, plusieurs techniques existent. Une des plus utilisées est l'analyse en composantes principales ou PCA (Principal Component Analysis) [SCSC03]. En général, elle permet de décorréler les dimensions d'un *dataset* pour donner ensuite la possibilité de réduire l'espace dimensionnel.

Une des applications de ces techniques est la détection d'intrusions dans les réseaux informatiques. Quelques travaux résument ce mécanisme en affirmant que « moins est plus » [SXZF07]. Cela est possible en faisant une décomposition de matrices permettant de faire une détection d'anomalies aussi performante que d'autres techniques, mais en utilisant moins de ressources de stockage et avec un plus faible coût de computation. D'autres applications de détection d'anomalies dans les réseaux font aussi une réduction avant d'appliquer un autre type de technique par exemple du *clustering* [LCD05].

Ainsi, les techniques spectrales ont démontré leur intérêt dans d'autres domaines. Par exemple, elles ont été utilisées pour la détection de dysfonctionnements et de compromissions dans les réseaux de capteurs sans fil [CPGM06]. La corrélation entre les mesures et la hiérarchie présente dans ce type de réseaux font que l'application de ces techniques a un intérêt particulier. Grâce à elles, il est possible d'atteindre une réduction de besoin de ressources, qui est une des plus fortes contraintes de ces systèmes.

I.2.5 Théorie de l'information

La théorie de l'information est une théorie énoncée par Claude Shannon en 1948 [Sha48]. Shannon présente avec cette théorie la problématique de mesurer la quantité d'information présente dans un ensemble de données. Cette question se pose pour multiples applications par exemple le codage de l'information et la compression de données. Différentes mesures ont été proposées comme l'entropie ou la complexité de Kolomogorov pour la quantifier.

Afin de pouvoir réaliser une détection d'anomalies à partir de cette théorie une hypothèse a été introduite : les anomalies vont introduire des irrégularités dans un *dataset*, ce qui va faire augmenter sa complexité. Ceci se traduit pour un *dataset* D avec une complexité $C(D)$ comme le problème de trouver des sous-ensembles I qui maximisent $C(D) - C(D - I)$. Ces sous-ensembles I seront considérés comme anormaux puisqu'en les enlevant du *dataset* la complexité de celui-ci diminue.

Quelques travaux vont profiter de la redondance implicite qui existe dans les *datasets* pour détecter des déviations [AAR96]. Cette théorie a ainsi été appliquée à la détection de fraudes et à la détection d'intrusions dans les réseaux informatiques [NC03] grâce à la réutilisation et à l'amélioration des résultats des travaux passés [LX01]. Dans ce but, elle inclut la théorie de graphes pour avoir une meilleure vision de la prédictibilité des données dans un *dataset*.

I.2.6 Apprentissage automatique

L'apprentissage automatique (« *machine learning* » en anglais) est un ensemble de techniques de l'intelligence artificielle qui permet à une machine de résoudre des problèmes complexes. Un processus systématique permet à une machine d'évoluer vers la résolution d'une tâche. Ce domaine permet de déchiffrer des problématiques qui ne peuvent pas être traitées ou résolues facilement par des méthodes classiques. Ces techniques servent à étudier par exemple des courbes et plusieurs structures de données comme les graphes ou les arbres.

Ces algorithmes permettent d'adapter une analyse ou un comportement à partir de la « pratique ». Ils sont souvent classés en fonction du mode apprentissage par exemple : « l'apprentissage supervisé » quand l'algorithme apprend avec des exemples accompagnés d'étiquettes et qui ont été créés par un expert et « l'apprentissage non supervisé » quand les données d'entraînement n'ont pas d'étiquettes et c'est l'algorithme qui doit les structurer et les classifier.

Dans les dernières années, l'apprentissage automatique a été appliqué à la cyber-sécurité [DD16, BG16]. Deux exemples d'application sont la détection d'intrusion réseau [HOA+15] et la détection de fraudes dans les assurances de santé [JRMB+15]. Même si on a introduit l'apprentissage automatique comme une catégorie des systèmes de détection, parfois il pourrait être considéré pour d'autres catégories présentées dans cette section. Par exemple, ces

techniques peuvent être utilisées pour la classification des messages de *phising*[6] [AA14].

I.3 Cyberattaques

Les cyberattaques représentent un risque très important pour les systèmes d'information. Dans cette section une définition de cyberattaque est donnée ainsi qu'une liste de différents types de celles-ci. Une analyse de leur évolution sera montrée. Finalement quelques exemples de cyberattaques connues seront présentés.

I.3.1 Définition de cyberattaques

En ce qui concerne les termes relatifs au cyberespace, il y a une définition différente pour chaque pays ou organisation. La CCDCOE (NATO Cooperative Cyber Defence Centre Of Excellence) a fait un grand et complexe glossaire avec différentes définitions proposées par les différents pays[7]. Une définition parmi les plus complètes pour le terme cyberattaque est celle-ci de l'Allemagne [Fed11] :[8]

Une **cyberattaque** est une attaque informatique dans le cyberspace ciblée sur un ou quelques autres SIs (Système Informatique) et avec l'objectif de nuire à la sécurité informatique. Les buts de la sécurité informatique, confidentialité, intégrité et disponibilité peuvent être tous ou individuellement compromis. Une cyberattaque ciblée contre la confidentialité d'un SI qui est lancé par un service de renseignement s'appelle **cyberespionnage**. Les cyberattaques ciblées contre l'intégrité et la disponibilité des SIs sont des **cybersabotages**.

Plusieurs ontologies ont été conçues sur la cybersécurité. Certaines ont pour objectif la définition de tous les éléments qui peuvent être impliqués dans une cyberattaque [OCM12]. D'autres traitent la problématique de la modélisation des cyberattaques [OCWM14]. Il existe également des ontologies plus spécialisées comme la cybersécurité sur les systèmes SCADA

6. Technique utilisée par des cyberattaquants pour obtenir des renseignements personnels dans le but de perpétrer une usurpation d'identité.

7. Resources, Cyber Definitions. NATO Cooperative Cyber Defence Centre of Excellence. Consulté le 13 Juillet 2017. `https://ccdcoe.org/cyber-definitions.html`

8. A **cyber attack** is an IT attack in cyberspace directed against one or several other IT systems and aimed at damaging IT security. The aims of IT security, confidentiality, integrity and availability may all or individually be compromised. Cyber attacks directed against the confidentiality of an IT system, which are launched or managed by foreign intelligence services, are called **cyber espionage**. Cyber attacks against the integrity and availability of IT systems are termed **cyber sabotage**.

[BC11].

L'ANSSI (Agence Nationale de la Sécurité des Systèmes d'Information) propose une classification des attaques en 15 catégories [Dir06]. Même si cette classification a été réalisée pour tout type de SI elle pourra être appliquée aux CPS comme cas particulier. Elle va permettre de parcourir tous les différents types d'attaques en fonction de leur but d'une façon générale. Les explications des différentes catégories sont présentées ci-dessous :

Destruction de matériels ou de supports : Ce sabotage a pour but de compromettre l'intégrité et la disponibilité d'un SI ou d'un de ses sous-systèmes.

Rayonnements électromagnétiques : Consiste à ajouter un brouillage a un système pour le rendre inutilisable.

Écoute passive : Cette interception de messages permet l'obtention d'information et comprend aussi la cryptographie.

Vol de supports ou de documents : Vol du matériel d'information. Une copie d'un document ou d'un fichier informatique est aussi considérée comme un vol.

Vol de matériels : Vol de ressources par exemple d'un ordinateur portable ou une mémoire.

Récupération de supports recyclés ou mis au rebut : Récupération d'information à partir d'une mauvaise destruction de données une fois finalisée leur utilisation.

Divulgation : Obtention d'informations confidentielles pour faire un chantage ou les utiliser avec un but non légitime.

Informations sans garantie d'origine : Information partielle, erronée ou trompeuse à partir de courriers électroniques.

Piégeage du logiciel : Toute attaque qui modifie le fonctionnement normal du système à partir de codes malicieux.

Saturation du SI : Saturation du système ou de ses voies de communication pour le bloquer ou pour fausser son comportement.

Utilisation illicite des matériels : Utilisation d'un système avec un but qui n'est pas celui pour lequel il est conçu.

Altération des données : Altération de données ayant comme but : la destruction, la modification ou l'insertion de messages pour bloquer ou modifier une communication ou pour identifier quels services sont disponibles dans un système.

Abus de droit : Un utilisateur avec privilèges réalise une action malveillante.

Usurpation de droit : Usurpation des privilèges pour avoir accès au système avec certains droits grâce à différentes méthodes.

Reniement d'actions : Ce type correspond à nier avoir participé à un échange d'information comme celle pour la perpétration d'une attaque.

Cette classification permet entre autres d'identifier les risques qui peuvent exister dans un système. La généralité des systèmes analysés permet l'application à n'importe quel système informatique.

I.3.2 Évolution des cyberattaques

Le cabinet d'audit PwC (PricewaterhouseCoopers) fait tous les ans un rapport de cybersécurité qui montre l'état et l'évolution de la sécurité dans les systèmes informatiques [PWC17] ainsi que Symantec [Sym17]. Ces rapports ne montrent pas seulement la quantité d'attaques qui se produisent dans le monde, mais ils font également une analyse de leurs bases de données en faisant une classification en fonction des différents critères. Aussi, il y a une étude sur l'évolution de l'investissement des entreprises pour renforcer leur cybersécurité.

Il ne faut pas penser aux nouvelles technologies comme des systèmes complètement sécurisés, mais comme de nouvelles cibles pour les attaquants. Un exemple est l'apparition des téléphones intelligents ou *smartphones*. Même s'ils ont été conçus en faisant attention à la sécurité des données, selon PwC et Symantec les incidents sur ce type de systèmes augmentent continuellement chaque année. Par exemple, un dispositif de l'IoT (Internet des Objets) est découvert et attaqué en moyenne 2 minutes après d'être connecté sur Internet.

De plus, l'apparition de logiciels et algorithmes d'analyse de systèmes et d'automatisation d'attaques permet chaque jour que des attaquants avec moins de compétences aboutissent à des attaques plus sophistiquées. La figure I.6 représente cette évolution dans le temps.

La figure I.7 représente l'évolution de la quantité de cyberattaques reportées au cours des années. Chaque année il y a plus d'attaques et la tendance est à l'augmentation. Ces valeurs doivent être mises en perspective avec l'augmentation du nombre de systèmes connectés ainsi que le nombre d'échanges au cours des années. En 2016 une réduction d'incidents a été reportée grâce à un effort économique majeur réalisé par les entreprises pour la protection de leurs systèmes. Par contre, les incidents étudiés ont montré une sophistication en

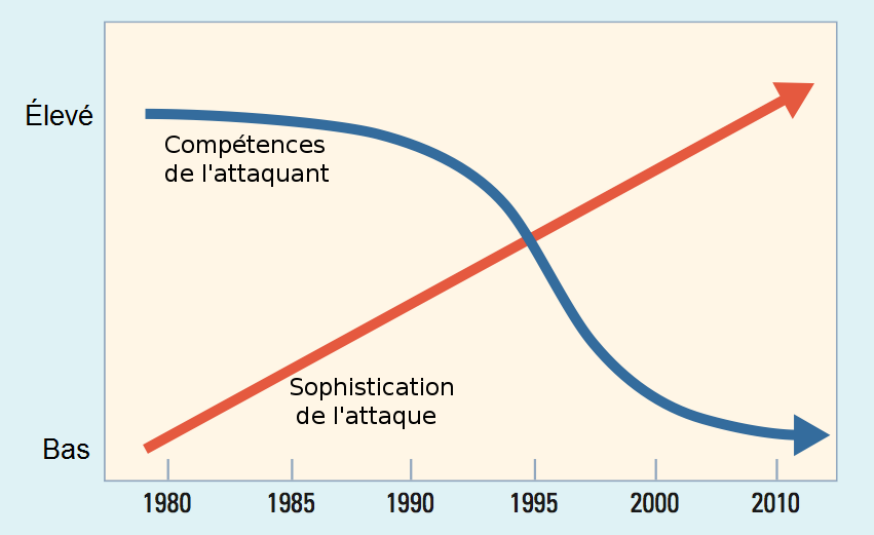

FIGURE I.6: Évolution des compétences de l'attaquant versus la sophistication de l'attaque [LC09].

hausse. L'utilisation de systèmes informatiques et leur importance sont en augmentation et en conséquence, chaque jour, ils sont plus attrayants pour les attaquants.

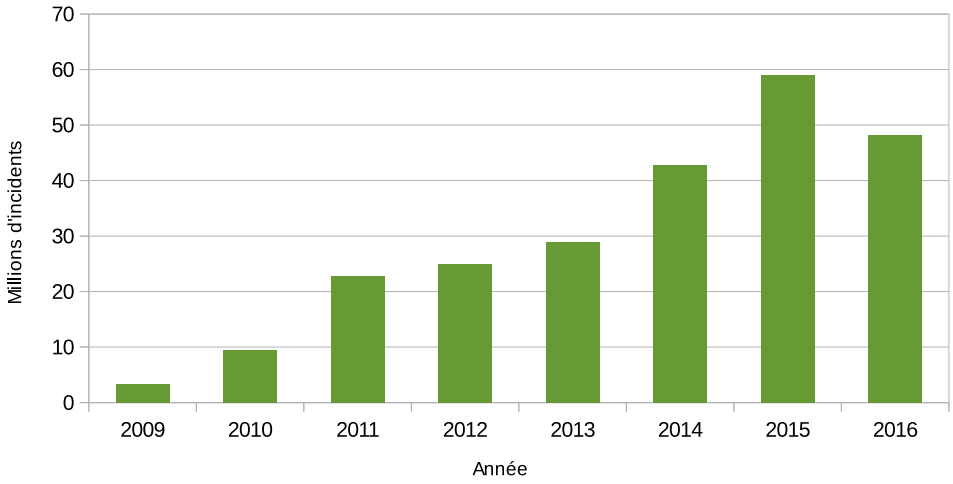

FIGURE I.7: Évolution des cyberattaques [PWC15, PWC16, PWC17].

Un défi important pour la détection de cyberattaques est que celles-ci changent leur comportement à fur et à mesure que les systèmes de détection évoluent pour devenir indétectables. De plus, parmi les méthodes de détournement, il existe des mécanismes d'autodestruction ou désactivation quand les attaques peuvent être découvertes. Un exemple

concerne les cyberattaques qui ne fonctionnent pas quand on les exécute dans une *sandbox*[9].

I.3.3 Cibles pour les cyberattaques

Les cibles d'une cyberattaque peuvent être très variables en fonction de l'objectif [LC09]. Cette étude utilise les données télémétriques de WINE (World Intelligence Network Environment), un réseau de sondes installées par Symantec qui récoltent les détails des attaques reçues par plus de dix millions d'utilisateurs répartis partout dans le monde. Dans ces documents, il y a une bonne représentation de la manière dont les cyberattaques se réalisent à distance. La plupart sont perpétrées depuis un autre pays, pour que les conditions de succès soient les plus favorables. Par exemple, une grande partie des attaques vient des pays de l'Europe de l'Est parce qu'il y a une conjonction entre une bonne infrastructure informatique et un fort taux de corruption. Un autre type d'attaque très commune relève du web et des fausses applications concentrées dans les pays riches, car le but lucratif est un des plus fréquents. Les attaques peuvent être classées en fonction de leur motivation. Par exemple, à partir des attaques connues, trois dimensions peuvent être identifiées [GSM+11] : politique, socioculturelle et économique.

Parfois la limite entre un *malware* et un logiciel est floue. Ainsi, les PUPs (*Potentially Unwanted Programs*) sont des logiciels malveillants non sollicités qui s'installent automatiquement en même temps qu'une autre application. Dans cette catégorie de PUPs il y a des logiciels bien connus par tous les utilisateurs comme ceux qui contrôlent les navigateurs pour obliger l'utilisateur à visiter les pages internet de quelques entreprises, juste pour faire de l'argent avec ces modifications. Le pire de ce cas d'attaque est que la plupart des antivirus sont accompagnés de ce type de *malware*[10] ou les ordinateurs avec un système d'exploitation installé comme le cas de *Superfish* sur les portables Lenovo.

Aujourd'hui, la cybersécurité dans les entreprises a pris une très grande importance. Actuellement, les pertes économiques dues aux cyberattaques sont chiffrées en environ 400 milliards de dollars [Cen14]. De plus, en moyenne 50 jours sont nécessaires pour résoudre les problèmes occasionnés par une intrusion et 23 jours pour combattre une *ransomware* [Acc17]. C'est à cause de ces difficultés que les efforts de l'industrie sont chaque jour plus forts.

9. Anglicisme de « bac à sable ». Environnement contrôlé qui permet l'analyse d'applications informatiques.

10. Has the antivirus industry gone mad ?. Emsisoft. Consulté le 13 Juillet 2017. `http://blog.emsisoft.com/2015/01/17/has-the-antivirus-industry-gone-mad/`

Les caractéristiques de ce type d'attaques sont qu'elles ont pu être réalisées à distance, qu'elles sont difficilement détectables et qu'il est presque impossible de reconnaître qui en est l'auteur. Ces problématiques donnent aujourd'hui de plus en plus d'importance à la cyberdéfense. Bien que l'ancien ministre français de la Défense, Jean-Yves Le Drian, a identifié la cyber comme quatrième armée [11], nous ne pouvons pas le définir encore en tant qu'armée à l'instar de celles de la terre, de la marine et de l'air, mais c'est sûr qu'elle est devenue un quatrième terrain d'affrontement.

De plus, il y a des victimes qui ne sont pas des cibles, mais qui sont toutefois affectées. C'est difficile pour un attaquant de savoir avec précision l'expansion que son *malware* aura dans le monde. Par exemple, Regin, un *malware* d'espionnage de haut niveau, a infecté selon Symantec [12] des systèmes tout genre, même s'il a probablement été conçu pour un objectif sensiblement différent.

I.3.4 Attaques connues

Il y a une grande quantité d'attaques connues à cause de leur répercussion économique, politique et médiatique. Cette section présente différents types d'attaques qui ont été lancées au long des années.

Internet est devenu un moyen de manifestation pour le grand public. Jusqu'à maintenant une manifestation consistait à bloquer les voies de communication d'une ville pour montrer des idées. Maintenant, les sites web sont les cibles de cyberattaques avec le même but [Kar14]. Mais la prise de contrôle d'appareils faiblement sécurisés a permis la création de *botnets* [13] qui sont capables d'avoir le même effet. Un exemple est l'attaque d'octobre 2016 avec le *botnet* mirai [14]. Aussi, l'accès à la culture et à l'information a été un point d'affrontement ayant comme exemple le partage décentralisé d'informations (par exemple, via Torrent) ce qui permet le partage d'informations volées. Il faut remarquer aussi le réseau Tor comme outil de navigation anonyme qui est parfois détourné pour la réalisation des activités illégales

11. Cyberguerre : au cœur de la quatrième armée. RTL. Consulté le 13 Juillet 2017. `http://www.rtl.fr/actu/societe-faits-divers/cyberguerre-au-c-ur-de-la-quatrieme-armee-7774738194`

12. Regin : Top-tier espionage tool enables stealthy surveillance. Symantec. Consulté le 13 Juillet 2017. `http://www.symantec.com/connect/blogs/regin-top-tier-espionage-tool-enables-stealthy-surveillance`

13. Réseau d'ordinateurs connectés à Internet qui travaillent en équipe pour l'exécution d'une tâche.

14. DDoS attack that disrupted internet was largest of its kind in history, experts say. The Guardian. Consulté le 13 Juillet 2017. `https://www.theguardian.com/technology/2016/oct/26/ddos-attack-dyn-mirai-botnet`

sur internet.

Les cyberattaques peuvent avoir une surface d'attaque très importante. Quelques attaques ciblent une grande quantité de victimes potentielles. Entre les attaques économiques actuelles les plus connues, l'attaque sur le site web eBay ressors [15]. eBay a été obligé de faire une communication pour demander à tous ses clients de changer leur mot de passe parce qu'ils n'ont pas pu savoir jusqu'à quel niveau la confidentialité a été compromise par les attaquants. D'autres attaques comme Slammer worm ou Sapphire sont des vers informatiques qui se propagent partout en quelques instants. 90% de machines vulnérables dans le monde ont été infectées en dix minutes [MPS+03]. Ce ver a profité d'une faille des logiciels Microsoft pour provoquer un déni de service dans les machines connectées en réseau et globalement un ralentissement d'Internet. Plus récemment en 2017, le *ransomware* [16] Wanacry [17] a démontré le risque que suppose la perte de données sur des systèmes critiques comme ceux des hôpitaux.

L'industrie est de plus un plus souvent victime des cyberattaques. Une attaque de type politique pour avoir accès aux informations industrielles stratégiques d'une centrale nucléaire en Corée du Sud a été perpétrée [18]. Dans cette attaque, les *blueprints* [19] des réacteurs nucléaires ont été volés. Même si cette attaque n'est pas considérée comme critique parce que les réacteurs n'ont pas une connexion avec l'extérieur, la filtration d'information d'une structure de ce type est une affaire très grave au niveau de cybersécurité.

Le cheval de Troie le plus cité sur les systèmes SCADA est Stuxnet, un logiciel de très haute sophistication et de complexité technique conçu probablement par la NSA (*National Security Agency*) aux États-Unis et l'unité 8200 en Israël pour attaquer les centrifugeuses d'enrichissement d'uranium iraniennes [FMC11]. Stuxnet est le premier ver informatique capable de reprogrammer les systèmes en plus de les espionner. Les dégâts de ce ver sont chiffrés à 45.000 CPS, dont la plupart étaient en Iran. L'attaque consiste en reprogrammer des automates conçus par Siemens grâce à une vulnérabilité inconnue. C'est pour cela que tout genre de système industriel utilisant un modèle d'automate vulnérable à cette faille a

15. eBay Inc. To Ask eBay Users To Change Passwords. eBay. Consulté le 13 Juillet 2017. `http://www.ebayinc.com/in_the_news/story/ebay-inc-ask-ebay-users-change-passwords`

16. Logiciel malveillant qui chiffre les données stockées sur un ordinateur pour les prendre en otage et demander de l'argent en échange de la clé qui permet de les déchiffrer.

17. What you need to know about the WannaCry Ransomware. Symantec. Consulté le 13 Juillet 2017. `https://www.symantec.com/connect/blogs/what-you-need-know-about-wannacry-ransomware`

18. South Korea nuclear plant hit by hacker. Cnet. Consulté le 13 Juillet 2017. `http://www.cnet.com/news/south-korea-nuclear-plant-hit-by-hackers/`

19. Plans techniques détaillés d'un objet qui sont très utilisés dans l'industrie.

été impacté.

Même si Stuxnet est l'attaque la plus réputée dans les systèmes SCADA, il y a d'autres exemples qui ont été détectés comme Black Energy2, Havex et Industroyer/CrashOverride. Les deux premiers sont des logiciels qui ont été utilisés pour l'espionnage industriel tandis que Industroyer/CrashOverride a les mêmes capacités que Stuxnet. Ce dernier a prouvé son pouvoir avec l'attaque perpétrée contre le réseau électrique ukrainien [20] [21].

Ces dernières années les arsenaux d'outils informatiques ont poussé en importance et en puissance. Une révélation de Wikileaks montre un échantillon des outils informatiques que potentiellement les États Unis pourraient utiliser [22].

I.4 Systèmes cyber-physiques

Les CPS sont des systèmes qui permettent le monitoring et le contrôle de processus physiques grâce aux systèmes informatiques. Ils sont utilisés dans plusieurs domaines comme le transport, la médecine et les *smart-grids* [23]. Ainsi, ils sont une pièce élémentaire dans la transition vers les usines de nouvelle génération appelées « Usine 4.0 » [LBK15] et les voitures autonomes [WYL+13].

I.4.1 Définition

Il existe d'innombrables définitions pour les CPS. Une qui convient à notre travail est la suivante :

Definition I.1. *Les **systèmes cyber-physiques (CPS)** sont une nouvelle génération de systèmes avec des capacités computationnelles et physiques qui peuvent interagir avec les humains grâce à différentes modalités.* [24] *[BG11]*

20. CrashOverride Malware. US-CERT. Consulté le 13 Juillet 2017. `https://www.us-cert.gov/ncas/alerts/TA17-163A`

21. 'Industroyer' virus could bring down power networks, researchers warn. The Guardian. Consulté le 13 Juillet 2017. `https://www.theguardian.com/technology/2017/jun/13/industroyer-malware-virus-bring-down-power-networks-infrastructure-wannacry-ransomware-nhs`

22. Vault 7 : CIA Hacking Tools Revealed. Wikileaks. Consulté le 13 Juillet 2017. `https://wikileaks.org/ciav7p1/`

23. Réseau électrique intelligent.

24. The term cyber-physical systems(CPS) refers to a new generation of systems with integrated computational and physical capabilities that can interact with humans through many new modalities.

Ces systèmes sont souvent classifiés comme **systèmes critiques**. Cette catégorisation est donnée à tous les systèmes pour lesquels une faille peut produire des conséquences inacceptables [Kni02]. En fonction de l'application, ces conséquences peuvent varier : dégâts humains, dégâts matériels ou dégâts environnementaux. La plupart des pays ont créé un cadre législatif et organisationnel pour les protéger. En France plusieurs secteurs ont été identifiés comme critiques, par exemple l'énergie, la santé et le transport.

Les défis des CPS sont nombreux [SGLW08]. Un ensemble d'exemples est présenté par la suite. Souvent leurs applications critiques donnent l'obligation d'avoir des réponses en **temps réel**. De plus, cette contrainte s'étend à tous les modules qui composent le système par exemple aux sous-systèmes de calcul et aux sous-systèmes de communication. Aussi, une **modélisation** de la structure du système et de son comportement est nécessaire parce que cela va permettre l'utilisation de méthodes analytiques et de cartographier les capacités du système. Un autre problème est la **confiance** qu'il faut donner à ces systèmes sans avoir de certitude sur leur fiabilité. La **qualité de service** ou *Quality of Service* (QoS) est une finalité importante. La **robustesse**, la **fiabilité** et la **sécurité** sont de dimensions qui contribuent à la QoS des CPS qui est un point d'importance cruciale dans les systèmes critiques.

I.4.2 Types de CPS

Il y a plusieurs classifications des CPS en fonction de leurs connexions, interactions et fonctionnalités. Normalement un système multi capteurs fait partie d'un CPS pour avoir une architecture complète de surveillance et contrôle ce qui permet l'automatisation des processus. Généralement ces systèmes sont construits en fonction des besoins et par conséquent ils ont des caractéristiques transversales qui ne permettent pas de les classer selon un type. C'est pour cela que parfois dans l'industrie, le terme SCADA est utilisé de façon générale. Une description et comparaison des quatre types plus génériques de CPS est donnée ci-dessous. Un aperçu des différentes configurations disponibles accompagne chaque terme [BW03].

1. Un **SCADA** est un système de RTUs (Remote Terminal Unit) qui collecte les informations de l'environnement ou d'un système et l'envoie à une station centrale à l'aide d'un système de communication. La station centrale affiche les données prises et permet à l'opérateur de prendre des décisions à partir de celles-ci. Le débit est en général réduit et il y a une forte dépendance avec les moyens de communication choisis.

2. Dans un **DCS** (Distributed Control System), les systèmes d'acquisition et les fonctions de contrôle sont réalisés par des unités distribuées de microprocesseurs qui sont à côté des instruments contrôlés. Ce sont des systèmes très évolués avec des fonctionnalités avancées et des interfaces de contrôle intégrées qui permettent d'avoir une paramétrisation et des opérations de contrôle simples. Le débit des connexions de ce type est assez haut (plus de 1Mops).
3. Un **PLC** (Programmable Logic Controller) ou en français un API (Automate Programmable Industriel) est une solution bon marché pour remplacer les interactions avec des relais. La solution se base sur la programmation du dispositif qui fait un traitement séquentiel d'un processus. L'appareil a deux éléments : les entrées qui viennent d'une partie commande ou des capteurs et une partie opérative de préactionneurs.
4. Les ***smart components*** ou composants intelligents sont des dispositifs qui peuvent être connectés aux autres types de systèmes. Ce sont des capteurs avec un microprocesseur qui fait un prétraitement des données pour envoyer plus tard ces informations à un système informatique.

Une classification est faite aussi à partir de la distance qu'il y a entre la base de contrôle et les opérations surveillées. Par exemple, dans un système SCADA normalement la distance d'opération est considérée comme « assez grande », mais cela reste très subjectif. Typiquement, la façon de faire la différence entre un système de contrôle de processus et un système SCADA est de voir respectivement si la connexion câblée entre les éléments est faisable directement ou s'il faut utiliser des réseaux ou des technologies sans fil.

I.4.3 Description d'un CPS

Les CPS sont composés par plusieurs sous-systèmes avec des objectifs précis. Le standard ANSI/ISA-95 [ISA] propose un langage commun pour les décrire. Selon ce standard, les CPS peuvent se décomposer en cinq niveaux qui sont décrits dans la Table I.2.

Les travaux réalisés pendant cette thèse s'inscrivent dans la détection d'anomalies et d'attaques correspondent aux niveaux 0, 1 et 2. Les PLC sont des sous-systèmes qui servent à faire le pont entre les niveaux 1 et 2 et ils sont donc aussi objets d'étude. Les niveaux 3 et 4 ne sont pas étudiés même si la propagation des effets des décisions prises dans ces niveaux vers les niveaux inférieurs peut également déclencher des alarmes. Quelques réflexions à ce

Niveau	Nom	Description
0	Physique	Processus de production
1	Capteurs et automates	Activités de percevoir et d'agir sur les processus physiques.
2	Surveillance et contrôle	Activités de surveillance et de contrôle des processus physiques.
3	Opérations de fabrication	Définition des activités du flux de travail (*workflow*) pour la réalisation d'une tâche ou d'un produit désiré.
4	Haute gestion	Management et réalisation des plans d'activité pour gérer la production.

TABLE I.2: Niveaux d'un CPS selon le standard ANSI/ISA-95.

sujet sont proposées dans la section I.4.6.

I.4.4 Différences entre un CPS et un Système d'Information classique

Il y a une tendance à unifier certaines activités entre les CPS et les SI classiques dans le but d'une collaboration entre leurs fonctionnalités. Par exemple, les CPS sont d'habitude intégrés aux systèmes informatiques de l'entreprise. La motivation principale de l'intégration des différents systèmes, plateformes, réseaux, logiciels et outils de maintenance est économique. Cette intégration va permettre de réaliser une organisation de l'entreprise plus efficace ce qui est l'intérêt de l'industrie 4.0. De plus, dans les CPS, les systèmes standards des SI sont chaque jour plus utilisés. Cette standardisation permet d'avoir plus de fonctionnalités et de facilités de connexion en évitant de gros développements. Même avec cette convergence, il y a des grandes différences entre un SI traditionnel et un CPS. Celles-ci sont montrées dans le tableau I.3.

Ces différences font que les architectures de détection et de réaction aux cyber-menaces des CPS doivent respecter quelques contraintes. Par conséquent, les méthodes des SI traditionnels ne peuvent pas être reprises par les CPS, le plus souvent à cause de la limitation du temps réel. La plupart des différences montrées signalent la criticité de ce type de systèmes et la nécessité d'éviter que les architectures de détection ne perturbent leur fonctionnement. La perte de données et les temps de récupération ne seront pas tolérés. La durée de vie et les mises à jour peu fréquentes font que ces architectures doivent prévoir des attaques qui

SI Traditionnel	**CPS**
La perte de données et les interruptions sont habituellement tolérées s'il y a une récupération avec sauvegardes de sécurité et redémarrages.	La perte de données et les interruptions ne sont pas tolérées parce qu'il peut y avoir des conséquences sérieuses.
Le débit doit être très haut et les retards peuvent être adaptés.	Réponses déterministes et en temps réel.
Une récupération par redémarrage n'a pas de conséquences dangereuses.	Il n'y a pas de grande tolérance avec les récupérations parce qu'elles peuvent être très dangereuses et entraîner éventuellement des pertes humaines.
Les antivirus sont très utilisés.	Les antivirus ne peuvent pas être appliqués parce qu'ils introduiraient des perturbations dans le système.
Connaissances et tests de sécurité.	Connaissances et tests de sécurité réduits.
Message chiffré.	Certains systèmes d'automates envoient les messages en clair.
Tests d'intrusion utilisés fréquemment	Les tests d'intrusion sont rarement réalisés et il faut les faire avec attention.
Mises à jour fréquentes.	Mises à jour peu fréquentes, bien cadrées et ayant besoin de faire intervenir souvent les fabricants des composants.
Audits de sécurité requis et réalisés périodiquement.	Audits de sécurité rarement réalisés.
Durée de vie entre trois et cinq ans.	Durée de vie beaucoup plus grande, jusqu'à vingt ans.

TABLE I.3: Différences entre un SI traditionnel et un CPS.

n'ont pas encore été développées. De plus, il faut penser aux CPS comme des systèmes qui n'ont pas été conçus en prenant en compte la sécurité et que des audits et tests de sécurité sont rarement réalisés. Ainsi, la Table I.4 représente la modification de l'ordre des priorités à l'heure de protéger les CPS.

À cause de ces différences, la recherche dans ce domaine est complexe surtout parce qu'il y a une absence énorme d'information. Les architectures des systèmes et les caractéristiques des capteurs ou des automates sont habituellement confidentielles malgré la standardisation. De plus, les entreprises cachent les détails des anomalies, plus particulièrement ceux des

Priorité	SI Traditionnel	CPS
1	Confidentialité	Disponibilité
2	Intégrité	Intégrité
3	Disponibilité	Confidentialité

TABLE I.4: Priorités dans la protection des systèmes.

cyberattaques subies.

En relation avec la qualité des données et de l'information, il y aura quelques attributs qui ne seront pas mesurables à cause de ces limitations, par exemple par la restriction des réponses en temps réel. Ces limitations peuvent également être elles-mêmes des attributs pour certains éléments de la qualité. Par exemple, les réponses déterministes et en temps réel peuvent être des attributs de la dimension « pertinence ». Une mauvaise qualité due à des données non pertinentes correspondant à des réponses qui n'ont pas étés données en temps réel peut être le signal d'une intrusion. En résumé, ces restrictions vont permettre d'avoir des flux de données avec une qualité qui peut être bornée, mais qui pourront par contre limiter énormément les architectures de détection.

I.4.5 Systèmes navals

Le travail réalisé pendant cette thèse est orienté vers les systèmes navals à cause de leur importance et criticité. Les bâtiments navals sont des systèmes avec des caractéristiques très particulières. Cette section explique en détail ces systèmes pour comprendre l'intérêt de leur étude.

Importance des systèmes navals

Concrètement, cette recherche se focalise sur les systèmes navals de par leur criticité et de par l'importance économique et stratégique du transport maritime. Le transport maritime de marchandises a toujours eu une importance majeure. La CNUCED (Conférence des Nations Unies sur le Commerce Et le Développement) publie tous les ans une étude sur les transports maritimes dans laquelle il est fait une analyse de leur état actuel et de leur évolution en fonction de différents critères [Uni16]. La figure I.8 montre que le transport d'une grande partie du commerce mondial est maritime en lien avec les autres moyens de transport disponibles. Le transport maritime est privilégié par rapport aux autres pour les

marchandises volumineuses tandis que l'avion est choisi pour les plus coûteuses. Sa croissance a été en général positive montrant une relation forte avec le produit intérieur brut. Son importance est restée proportionnelle par rapport aux autres moyens de transport dans le temps.

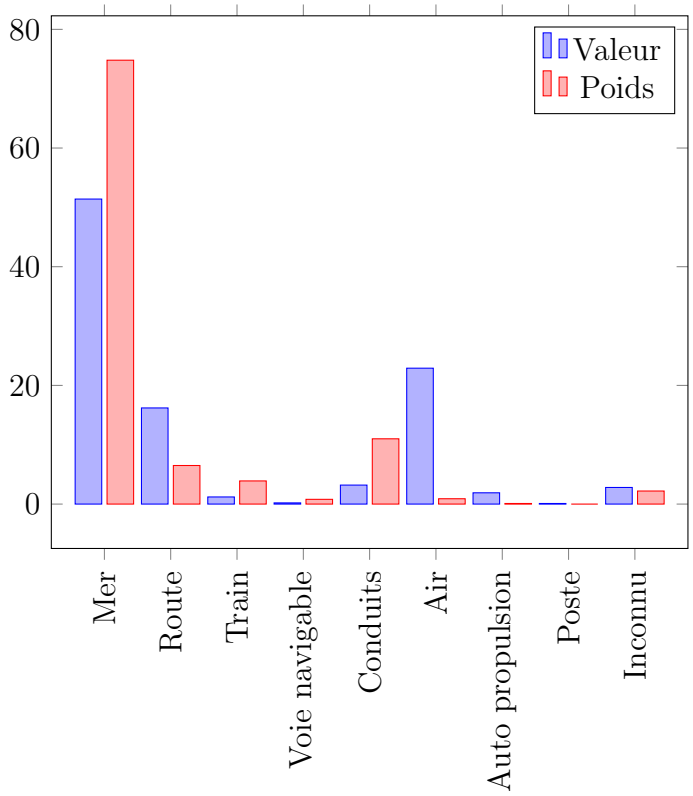

FIGURE I.8: Commerce des 28 pays européens par moyen de transport en 2014[Com16].

Pour la première fois, plus de 10 milliards de tonnes de marchandise ont été transportées en 2015 par porte-conteneurs [Uni16]. En outre, une partie considérable des porte-conteneurs est capable de transporter plus de 5000 TEU (*Twenty-foot Equivalent Units*), ce qui représente une grande valeur par envoi. La tendance est de faire des navires de plus en plus volumineux. Par exemple, les nouveaux porte-conteneurs OOCL Hong Kong [25] et Madrid Maersk [26] ont respectivement une capacité de 21.413 et de 20.568 TEUs . Dans ce contexte, une grande quantité d'anomalies, et en particulier de cyberattaques, peuvent engendrer des conséquences significatives.

25. OOCL reaches milestone with the christening of the OOCL Hong Kong. OOCL. Consulté le 27 Septembre 2017. `http://www.oocl.com/eng/pressandmedia/pressreleases/2017/Pages/12may17.aspx`

26. World's Largest Boxship Joins Maersk Line's Fleet. World maritime news. Consulté le 27 Septembre 2017. `http://worldmaritimenews.com/archives/217566/worlds-largest-boxship-joins-maersk-lines-fleet`

Mais l'importance maritime n'est pas seulement due au transport de marchandises, il y a d'autres intérêts comme l'extraction de pétrole et la pêche. Une technologie actuellement en plein développement est la réalisation de travaux marins à l'aide de *rovers* ou drones de plus en plus autonomes. Par exemple, le sous-marin autonome de l'université japonaise d'Okayama [27] permet d'exécuter des tâches qui pourraient être utiles pour des missions de surveillance d'installations sous-marines ou environnementales.

Cette importance économique fait de la mer un terrain qu'il faudra surveiller et protéger en prenant en compte que les menaces peuvent à présent provenir du cyberespace. Ce nouveau terrain apparaît du fait de la croissance de l'informatisation des systèmes et en particulier à bord des navires militaires et civils.

Les systèmes navals et les CPS

Grâce aux CPS, les nouvelles générations de systèmes navals évoluent constamment pour améliorer leurs performances, réduire l'équipage et simplifier leur utilisation. Un réseau de capteurs embarqués fournit plusieurs flux de données critiques permettant de naviguer en sécurité en suivant les routes optimales tout en surveillant le bon fonctionnement de tout le système. Ces tâches de contrôle et de surveillance sont effectuées depuis les postes placés dans la passerelle. Le contrôle distant et l'automatisation de processus font que la gouvernance des bâtiments a une forte synergie avec les réponses automatiques et l'aide à la décision. Parfois, ces systèmes ont des connexions vers l'extérieur permettant une surveillance et un contrôle à distance depuis un poste à terre.

De plus, il y a des systèmes indépendants qui réalisent des actions conjointes. Sur toutes les frégates, il y a au moins un hélicoptère donnant la possibilité de reconnaître le terrain depuis le ciel. À présent, l'utilisation de drones autonomes ou télécommandés à la place des hélicoptères pour certaines missions est privilégiée. Pour l'instant, ils sont seulement utilisés comme capteurs, mais dans le futur ils posséderont des systèmes d'armes. Quand un bateau est en flotte, il y a un partage d'information entre tous les bateaux faisant attention aux différents niveaux de confidentialité. Par exemple, une flotte peut être constituée par des navires de différents pays et il peut y avoir une communication restreinte entre les bateaux de chaque pays.

27. Intelligent deep sea robotics : Autonomous underwater robot with intelligent 3D cameras for high precision search and tracking in deep seas. Okayama University. Consulté le 16 Juillet 2017. `http://www.okayama-u.ac.jp/user/kouhou/ebulletin/feature/vol9/feature_001.html`

Il faut penser que dans un bateau plusieurs réseaux sont installés : comme le réseau de gestion et le *wifi* pour l'équipage qui est essentiel dans les missions longues afin que l'équipage ait une façon de communiquer avec leur famille. Tous ces réseaux sont isolés les uns des autres et contrôlés par la direction des services d'information. Il y a un poste à la passerelle dédiée à faire le tri des connexions, cherchant à réduire les risques d'espionnage ainsi que d'empêcher l'envoi d'information confidentielle et les sabotages depuis l'intérieur du bâtiment.

Du fait de ces caractéristiques, le domaine naval présente un cas particulier de CPS. Il faut prendre en compte le constat que la durée de vie des navires étant de plusieurs décennies, la technologie à bord peut être dépassée. Par exemple, le porte-avions Charles de Gaulle, un des bâtiments les plus célèbres de la marine nationale française, qui a été commandé en 1986 n'est pas entré en service avant 2001. De plus, un bâtiment de la sorte a une vie moyenne de 50 ans. Donc, en 2050, ce bateau en particulier, encore en service, aura une technologie des décennies 1980 et 1990 quand les réseaux et les protocoles pour la transmission de données seront sûrement complètement différents. Un exemple plus récent est le porte-avions HMS Queen Elizabeth de la marine anglaise, livré en 2017. Il a mis en question cette problématique avec l'utilisation de Windows XP, système d'exploitation considéré comme dépassé et vulnérable, pour la réalisation de certaines tâches [28]. Les mises à jour de ces systèmes ne seront certainement pas suffisantes pour empêcher le dépassement des technologies utilisées initialement. Ainsi, même si les sous-systèmes qui composent les CPS navals sont « marinisés » pour les protéger des effets de la mer, l'environnement maritime de fonctionnement est rude et incertain, pouvant provoquer de multiples anomalies dans ses composants.

Toutes les catégories de cyberattaques présentées dans la Section I.3.1 peuvent être extrapolées aux systèmes navals. Par exemple, des attaquants peuvent avoir l'intérêt de surveiller la position d'un bateau en réalisant une écoute passive, le détourner par différents moyens ou voler des informations industrielles ou militaires. La Table I.5 résume les différents systèmes qui se trouvent sur les navires et les scénarios potentiels de crise qu'un acte malveillant ou un dysfonctionnement peut produire.

Ces systèmes ont déjà démontré leur vulnérabilité à différentes attaques. Par exemple, Maersk, une des principales entreprises de transport maritime, a reconnu en 2017 une cyberattaque sur ses systèmes informatiques (Figure I.9). Son message mettait en perspective le

28. HMS Queen Elizabeth doesn't run on Windows XP. UK defence journal. Consulté le 16 Juillet 2017. `https://ukdefencejournal.org.uk/new-aircraft-carriers-dont-run-windows-xp/`

TABLE I.5: Systèmes navals et risques associés

Système	**Exemples**	**Risques**
Position	GPS GLONASS, Galileo, BeiDou, NAVSAT, LORAN	Détournement du navire, mauvais fonctionnement d'autres systèmes comme l'AIS
Navigation	Compas, loch/speedo, sondeur	Détournement du navire, prise de mauvaises décisions
Cartographie	ENC (*Electronic Navigational Charts*)	Calcul de routes pas optimales et potentiellement dangereuses
Météo	Centrale météo, prédictions	Calcul de routes pas optimales et potentiellement dangereuses
Repérage d'autres bateaux	AIS, Radar, sonar	Risque de collision
Communication	Radio VHF, téléphonie, réseau, Internet	Espionnage, sabotage, isolement
Contrôle	Propulsion, gouvernail	Perte de contrôle, détournement, réduction de durée de vie
Computation	Serveurs et ordinateurs de bord (pilote automatique, calcul de routes...)	Calcul de routes pas optimales et potentiellement dangereuses
Monitoring (cuves, moteur, groupe électrogène...)	Pression de fluides, niveau des cuves (ultrasonique, ondes sonores, sondes), température, tachymètre, flux de fioul, détection de flamme	Occultation de sabotages et pannes, prise de mauvaises décisions
Sécurité et surveillance du navire	Capteurs incendie, CO_2, voie d'eau, caméras, détecteur de glace, détecteur MOB (Homme à la mer)	Fausses alertes, pas de détection de situations de danger
Surveillance extérieure	Hélicoptère, drone aérien, drone de surface, drone subaquatique	Perte du système, surveillance irréalisable
Monitoring secondaire	Détecteur de portes ouvertes, contrôle de lumières, alarmes, chambre froide	Problèmes d'habitabilité et de santé

risque que cela pourrait représenter pour les employés, les opérations en cours et les affaires de ses clients.

FIGURE I.9: Message de Maersk sur Twitter en reconnaissant une cyberattaque affectant ses systèmes informatiques.

Ainsi, les premières cyberattaques contre les systèmes de position ont commencé a être reportées. Par exemple, fin Juin 2017, une vingtaine de bateaux ont détecté une attaque de type *spoofing* sur leurs GPSs dans la mer Noire [29].

En conséquence, les CPS équipés dans les bâtiments navals ont une importance majeure à cause de leur criticité et de leur importance stratégique et économique. Au vu de leurs vulnérabilités informatiques et physiques, les CPS navals doivent être protégés tout en respectant leurs contraintes.

I.4.6 Sécurité des CPS

À partir de la décomposition par couches présentée dans la section I.4.3, la sécurité des CPS peut être abordée indépendamment pour chaque niveau. Sur la Table I.6, les éléments à sécuriser par niveau sont indiqués.

29. Ships fooled in GPS spoofing attack suggest Russian cyberweapon. New scientist. Consulté le 11 Août 2017. `https://www.newscientist.com/article/2143499-ships-fooled-in-gps-spoofing-attack-suggest-russian-cyberweapon/`

Niveau	Nom	Éléments à sécuriser
0	Physique	Hardware et entourage
1	Capteurs et automates	Software
2	Surveillance et contrôle	Réseau
3	Opérations de fabrication	Physique et personnel
4	Haute gestion	Management

TABLE I.6: Couches de sécurité d'un CPS.

Depuis quelques années, plusieurs listes des bonnes pratiques ont été conçues par les professionnels de la sécurité. Par exemple, le département de l'énergie des États-Unis propose une liste de 21 pas pour sécuriser les CPS [Sha06]. La liste ci-dessous explique les différentes mesures de sécurité et leur relation avec ces systèmes pour les niveaux examinés dans cette thèse [Kru05]. Néanmoins, cette liste est d'application générale donc elle n'est pas adaptée pour une application directe aux CPS à cause de leurs caractéristiques, mais elle permet d'identifier les points importants de leur sécurité.

— **Logs de surveillance et audits :** Les journaux d'activité (log) sont utiles pour analyser les événements qui sont passés ou pour la surveillance en temps réel. Les anciens CPS n'ont pas la capacité de créer de tels fichiers de log et le coût d'installation peut-être trop élevé comparé aux bénéfices immédiats.
— ***Firewalls* :** Le blocage de messages entre différents réseaux est nécessaire pour protéger le système des intrusions à partir d'un autre système informatique. Mais il faut évaluer les retards que le *firewall* va introduire avec son traitement de messages. Il y a des *firewalls* spécifiquement conçus pour les CPS [NP16].
— **IDS (Intrusion Detection Systems) :** Les IDSs sont des systèmes pour la détection des intrusions dans un système informatique. Il y a différents types de IDS : NIDS (Network Intrusion Detection System) si l'analyse est faite dans le réseau ou HIDS (Host Intrusion Detection System) si elle est faite sur une machine. Ces deux types d'IDS sont complémentaires, l'un n'étant pas l'alternative de l'autre. Il n'y a pas d'IDSs pour tous les protocoles utilisés par les sous-systèmes, car ce sont souvent des protocoles propriétaires dont les spécificités précises sont inconnues. Une étude devra être réalisée pour savoir si l'utilisation d'un IDS peut potentiellement affecter le fonctionnement du système.
— **Détection et élimination de code malveillant :** La charge computationnelle nécessaire pour la détection et l'élimination des codes malveillants n'est pas toujours disponible sur un CPS sans affecter les réponses en temps réel. De plus les

mises à jour des logiciels et des bases de données exposent le système aux attaques extérieures.

— **Mots de passe :** Dans un CPS, les utilisateurs doivent utiliser des mots de passe pour avoir accès au système avec certains privilèges en cas d'urgence. En cas d'une situation d'oubli de mot de passe ou d'erreurs répétées, un système traditionnel bloque l'utilisateur, mais sur un CPS en temps réel, un refus d'accès peut être particulièrement dangereux. Ainsi, les mots de passe des systèmes locaux doivent être très simples ou éliminés. Au niveau de la supervision, il faut utiliser des systèmes d'authentification plus avancés, par exemple en deux étapes, et des systèmes cryptographiques dans les communications en cas d'interception.

— **Mesures biométriques :** L'accès par systèmes biométriques promet d'être le futur pour l'authentification et l'obtention des privilèges nécessaires permettant d'opérer dans le système. La biométrie est encore en développement, mais il y a déjà de nombreux exemples d'application. Ces systèmes peuvent également être utilisés pour prendre en compte le profil de l'utilisateur, par exemple avoir une console de contrôle personnalisée dans la passerelle d'un bâtiment en fonction de la personne qui est en train de l'utiliser.

— **Chiffrement par clé publique ou asymétrique :** Dans ces systèmes cryptographiques, il n'y a pas d'échange de mot de passe entre l'expéditeur et le récepteur. Il y a une correspondance entre un mot de passe publique et un autre privé qui est protégé sur le récepteur et impossible à trouver à partir de la clé publique. De plus, une clé publique permet de signer un document et de l'envoyer à n'importe qui ayant accès à cette clé. Ce type de cryptographie prend beaucoup de temps et de traitement et c'est pour cela que ces opérations sont généralement incompatibles avec les systèmes en temps réel comme les CPS.

— **Chiffrement par clé secrète ou clé symétrique :** L'expéditeur et le récepteur ont la même clé secrète qui sert à chiffrer et déchiffrer les messages. L'avantage de ce système est que le temps de traitement est plus rapide que le chiffrement asymétrique. Ces clés doivent être transmises à tous les émetteurs et récepteurs de messages de façon sécurisée, par exemple un chiffrement à clé publique. De cette façon, le système profitera du temps de traitement court, du chiffrement symétrique et d'une partie de la sécurité du chiffrement asymétrique. Sur les CPS cette méthode est utilisée uniquement dans le cas des SCADA à grandes distances ou dans quelques sous-systèmes très critiques.

— **Contrôle d'accès basé sur le rôle :** Ce type d'accès se répand de plus en plus dans

les services gouvernementaux et dans l'industrie. C'est un contrôle d'accès qui prend le rôle de l'utilisateur au lieu de son identité comme méthode d'authentification avec privilèges.

— **Redondance :** Une mesure de protection contre les attaques et anomalies dans tous systèmes est la duplication de sous systèmes. Les *backups*, les duplications et les systèmes éloignés géographiquement font que dans le cas où il y a une faille, un autre système pourra rapidement prendre sa place ou bien elle sera résolue à partir des *backups*. Normalement, un CPS est divisé en deux parties : une principale et une autre secondaire qui prendra le contrôle en cas d'indisponibilité de la première.
— Les ***honeypots*** sont des cibles piégées pour que les attaquants pensent avoir accès à une partie sensible du système, mais il s'agit uniquement d'une simulation. Ce système peut aider à prévenir une attaque sur une partie plus sensible du système ou bien à l'esquiver.

L'ensemble de ces points devra être pris en compte pour définir et mesurer la qualité. Dans une situation particulière, un *firewall* peut être considéré comme un système d'analyse de certains aspects de la qualité de données. En effet, un flux de données qui est passé par un *firewall* aura sa qualité vérifiée sur certaines dimensions. Un de ces éléments pourra aussi être utilisé comme un attribut, par exemple, la force d'un mot de passe en tant qu'attribut de la dimension sécurité. Néanmoins, il y aura certainement des mesures qui ne seront pas réalisables du fait des limitations du système.

I.5 Conclusion

La qualité des données et de l'information a été introduite avec l'objectif de résumer brièvement les études existantes et pouvoir analyser comment elles peuvent s'intégrer à notre problématique. La majorité de la recherche sur la qualité réalisée jusqu'aujourd'hui a été focalisée sur les MIS comme la plupart des articles cités dans ce chapitre.

La mesure de la qualité sur les CPS est depuis quelques années un défi émergent. Les méthodologies présentées sont souvent des applications sur problématique concrètes avec difficile réutilisation. Entre elles, aucune n'est optimisée pour la qualité des flux de données des sous-systèmes des CPS et plus précisément orientée vers la détection d'intrusions.

La détection d'anomalies est une problématique qui est étudiée depuis longtemps. C'est pour cela qu'il existe une multitude de techniques avec des caractéristiques particulières. La

catégorisation de ces techniques a été examinée pour avoir une vue d'ensemble du domaine. Ce panorama va permettre d'apprécier dans les chapitres suivants comment la qualité peut s'incorporer à ce domaine.

La cyberdéfense a été présentée comme une problématique qui devient de plus en plus importante. Les cyberattaques, étant des anomalies provoquées, peuvent être détectées grâce aux mécanismes de détection de celles-là. C'est pour cela que l'incorporation de la qualité pour la détection d'anomalies sera introduite dans les chapitres suivants en tant que potentielle solution à cette problématique.

Les CPS ont été présentés comme un cas particulier d'étude. Leurs différentes caractéristiques et particularités ont été examinées pour comprendre l'importance de la recherche réalisée. Aussi, les différences entre ce type de systèmes et les systèmes informatiques classiques ont été constatées. Par conséquent, les systèmes d'évaluation de la qualité ainsi que les systèmes de détection doivent respecter les contraintes définies pour ces systèmes. Ainsi, ces différences avec les systèmes classiques peuvent apporter des bénéfices pour la détection des actes malveillants et des dysfonctionnements. Les systèmes navals comme cas spécifique de CPS sont présentés avec une description de leurs systèmes et leurs risques associés. Les systèmes de sécurité existants pour les CPS ont été listés pour déterminer où la nouvelle méthodologie peut s'insérer au sein des systèmes existants.

Chapitre II

Méthodologie pour la mesure de qualité

Sommaire

II.1 Définitions contextuelles 48
II.2 Définition des entités 49
II.2.1 Données 49
II.2.2 Information 50
II.2.3 Connaissance 51
II.2.4 Intelligence 51
II.3 Mesure de la qualité à différents niveaux dans les CPS 52
II.3.1 Qualité des données 54
II.3.2 Qualité de l'information 55
II.3.3 Qualité de la connaissance 58
II.3.4 Qualité de l'intelligence 59
II.3.5 Méthodologie d'évaluation de la qualité 60
II.4 Détection des anomalies 62
II.4.1 Hypothèses 62
II.4.2 Approche de détection et catégorisation 63
II.4.3 Implémentation de l'approche 65
II.5 Conclusion 69

Ce chapitre introduit une nouvelle méthodologie pour la mesure de la qualité dans les CPS. Cette méthodologie va permettre d'intégrer la détection des anomalies en général, et les cyberattaques en particulier, appliquant des mesures adaptées de qualité. Tous les éléments qui ont mené à cette méthodologie sont présentés afin de comprendre la problématique traitée et la structure du modèle résultant.

II.1 Définitions contextuelles

Ce travail aborde la problématique de la mesure de la qualité d'un sous-système considéré comme un module élémentaire d'un système. C'est pour cela que la compréhension de cette thèse nécessite la définition d'un ensemble de termes. Plusieurs approches ont proposé les définitions de ces termes. Parmi elles, nous avons identifié celles de la littérature qui s'adaptent le plus à nos travaux. Cette étude se basera sur les définitions suivantes [Xia09] :

Definition II.1. *Un* ***système*** *est un ensemble d'éléments interdépendants qui interagissent entre eux de manière organisée et formant un ensemble unique. Parmi les éléments constituant un système, se trouvent des produits, des processus, des humains, des informations, des techniques, des installations, des services et d'autres éléments de support.*

Definition II.2. *L'****architecture d'un système*** *est un arrangement d'éléments et de sous-systèmes qui offrent des fonctions afin de respecter les spécificités techniques requises par le système.*

Dans l'architecture d'un système, il y a des éléments et des sous-systèmes qui peuvent être identifiés comme modules élémentaires. Pour qu'un module puisse être défini comme élémentaire, il doit remplir quelques caractéristiques. Sa définition peut être simplement : l'élément fonctionnel indivisible qui fait partie d'un système. Mais il y a plusieurs définitions plus élaborées pour les CPS [CYB00].

Definition II.3. *Un* ***module élémentaire*** *est défini comme étant une unité composée des éléments ayant des connexions très étroites entre eux et des connexions faibles avec les éléments des autres unités.*

Il faut aussi considérer qu'un système aura des limitations et il faudra déterminer les conditions qu'il doit respecter. Une fois que ses caractéristiques de base ont été définies, il faudra chercher un compromis entre ces caractéristiques et ce qui est faisable avec les ressources disponibles. Pour la conception et mise à jour d'un système, il faudra prendre en compte qu'un système doit [INC07] :

1. Satisfaire les prérequis.
2. Avoir une architecture fonctionnelle.
3. Être suffisamment proche de l'optimal en faisant attention à ses restrictions de temps, budget, connaissances et ressources.

4. Être cohérent avec la technologie existante et avec les risques des éléments disponibles.

Dans le cas des CPS, un prérequis essentiel sera la sécurité parce que ce sont généralement des systèmes d'importance critique. Cette sécurité s'étend comme prérequis à tous les modules élémentaires afin d'éviter l'apparition de points faibles.

II.2 Définition des entités

Afin d'étudier la qualité dans les CPS, la pyramide DIKW présentée dans la Section I.1.1 a été prise comme référence. Elle a été choisie parce que sa généralité et ses éléments correspondent avec les éléments identifiés lors de la définition des connaissances dans ce type de systèmes. Cela veut dire que les sous-systèmes s'échangent ou enregistrent des données qui ne sont pas directement compréhensibles, si celles-ci ne sont pas associées aux contextes du protocole de communication et du capteur. Quand ces données sont interprétées et que l'on connaît ce qu'elles représentent, il est possible d'obtenir des informations. À partir de ces informations, les CPS créent des réponses ou des résultats intermédiaires, en relation avec la connaissance, grâce à un processus associé à l'intelligence. Ensuite, tous ces éléments sont décrits en détail et accompagnés respectivement par une définition formelle.

II.2.1 Données

Deux types de données s'opposent : les données quantitatives et les données qualitatives. Les données quantitatives sont exprimées par des valeurs numériques tandis que les données qualitatives sont représentées par des termes ou des notions floues. Dans ce mémoire nous ne traiterons que les données quantitatives, majoritairement prédominantes dans les CPS.

Plusieurs définitions des données quantitatives existent, cependant elles sont relativement proches. Nous pouvons ainsi formuler une définition générale du terme "donnée".

Definition II.4. *Les **données** sont des représentations discrètes et objectives des événements et des entités qui ont été observés [Tod14].*

Les données sont le résultat de faire l'abstraction de la réalité de deux processus : modélisation et représentation. La modélisation consiste en l'obtention de triplets « entité,

attribut, valeur » qui permettent de définir le domaine d'observation. En résumé, c'est une mesure sur une entité d'un attribut qui donne comme résultat une valeur. La représentation consiste en l'affichage ou l'enregistrement de ces triplets. Le niveau d'abstraction des données sera déterminé par le contexte d'utilisation.

II.2.2 Information

L'extraction de l'information a été abordée dans plusieurs domaines et avec différents points de vue. En conséquence, les définitions de la communauté ont eu plutôt tendance à se différencier [Min96]. Actuellement, la définition la plus habituelle est que l'information est un ensemble de données qui ont été traitées pour les rendre utiles. Une autre définition construite à partir des différents domaines de recherche est résumée par l'équation suivante :

$$\text{Information} = \text{donnée} + \text{sens} \tag{II.1}$$

En définissant le sens comme le contexte nécessaire pour que les données soient compréhensibles et qu'elles puissent être interprétées dans l'objectif de définir des connaissances. Plusieurs études ont été réalisées pour définir le contexte de différentes problématiques [Pet10]. En fonction du cas d'étude, le contexte sera défini par différents éléments comme par exemple la localisation et le temps.

Dans notre cas d'étude, le contexte d'un flux de données sera composé par le contexte du sous-système qui crée la donnée et par le contexte du système dont il fait partie. Cela va permettre d'interpréter la donnée et de savoir ce qu'elle représente.

$$Information = \text{Donnée} + Contexte_{sous\text{-système}} + Contexte_{\text{système}} \tag{II.2}$$

Le contexte du sous-système comprend les spécifications techniques, par exemple le protocole utilisé pour la transmission de la donnée et sa représentation. Le contexte du système ajoute les spécifications techniques du système complet où il est installé, ainsi que de son entourage. En d'autres termes, toutes les variables qui peuvent affecter le système, quand les mesures de l'environnement du système sont comprises.

II.2.3 Connaissance

La connaissance est un terme difficile à définir. Normalement, une grande partie des définitions s'appuient sur le concept d'information pour la définir [RH08]. Par exemple, Zeleny explique la connaissance avec « know-how », cela veut dire, savoir comment utiliser l'information. Le point commun de toutes ces définitions est l'utilisation de l'information avec un objectif. Pour les CPS une définition qui peut être utilisée est la suivante :

Definition II.5. *La connaissance est l'information qui a été traitée, organisée, ou structurée d'une certaine façon, ou qui a été appliquée ou mise en action*[1] *[Pit11].*

À partir de cette définition, nous pouvons identifier dans les CPS plusieurs éléments comme des connaissances : l'adaptation d'un processus, la transformation d'une information pour la rendre utile ou le déclenchement d'une action.

II.2.4 Intelligence

L'intelligence (*wisdom* en anglais) est un terme qui, lui aussi, a différentes définitions en fonction de l'objectif recherché. Ainsi, sa définition a évolué pour certains auteurs. Par exemple pour Zeleny, au début l'intelligence pouvait être décrite avec « *know-why* » (savoir pour quoi) pour après étendre cette définition à « *why-do* » et « *know-what to do, act or carry out* » [Zel05]. En général, la définition suivante peut être utilisée :

Definition II.6. *L'intelligence est l'ensemble des éléments nécessaires qui expliquent un fonctionnement, un processus ou une décision.*

Quand ces définitions sont adaptées aux CPS, l'intelligence peut-être, entre autres, un principe mathématique, une loi physique ou une règle définie à partir de l'expérience. Ces éléments vont faciliter la définition de différents types de connaissances, par exemple une réaction du système. L'intelligence peut être vue aussi comme une fonction qui donne un résultat avec une ou plusieurs informations comme entrée. Nous pouvons la représenter de la façon suivante :

$$Connaissance = Intelligence(Information) \tag{II.3}$$

1. *Knowledge refers to information having been processed, organized or structured in some way, or else as being applied or put into action.*

II.3 Mesure de la qualité à différents niveaux dans les CPS

Parmi les nombreuses études publiées sur la qualité des données et de l'information, aucune méthodologie ne s'adapte aux CPS en général et plus particulièrement aux systèmes navals. Dans la Section I.1.2, nous avons vu comment les travaux réalisés sur les CPS sont orientés vers des applications concrètes et sont difficilement généralisables. Ainsi, seulement quelques niveaux de la pyramide DIKW sont pris en compte par ces analyses.

En raison de cela, dans ces travaux, seulement un ensemble de dimensions est étudié en fonction de l'objectif recherché. Lors de l'application générale de ces méthodes aux systèmes navals, nous avons identifié un manque de dimensions pour l'évaluation de la qualité de certains sous-systèmes. Jusqu'à maintenant, ces études ont toujours présenté l'humain comme consommateur final des connaissances. La détection d'anomalies dans le sous-système doit prendre comme référence le sous-système au lieu de l'humain. Ainsi, les résultats doivent être bien structurés et applicables à d'autres problématiques comme la propagation de la qualité. À partir de cette analyse, nous pouvons construire une liste de limitations et de déficiences à examiner qui permettra de les comparer plus tard aux solutions proposées par nos travaux. Cette liste est présentée ci-dessous, indiquant quelques points spécifiques à traiter :

1. Manque d'application générale pour tout sous-système d'un CPS
   (a) Besoin d'une identification générale d'éléments de la qualité
2. Évaluation partielle de la pyramide DIKW
   (a) Seulement les données et les informations sont examinées
3. L'humain a un rôle secondaire dans les CPS
   (a) L'humain n'est plus le consommateur final
   (b) L'humain ne peut pas toujours être un évaluateur de la qualité
4. Les résultats sont adaptés à une problématique précise
   (a) Réutilisation difficile des résultats
   (b) Il n'y a pas d'application à la détection d'actes malveillants et dysfonctionnements

En vue de résoudre ces limitations, une nouvelle méthodologie a été proposée adaptée à ces systèmes [MLBP16]. Dans notre définition, les caractéristiques des CPS sont prises en

compte et elles doivent être respectées lors de l'application de cette approche. Par exemple, toutes les opérations réalisées par un sous-système ayant besoin de temps réel doivent respecter cette contrainte. D'autre part, l'automatisation des CPS accorde à l'humain une position secondaire. Cela veut dire que l'humain n'est plus le consommateur principal de l'information parce que de plus en plus de sous-systèmes sont capables de produire des réponses automatiques. Ainsi, l'humain ne peut pas être l'évaluateur de la qualité à cause des contraintes des CPS sauf dans des cas très précis où il n'y a pas par exemple, besoin de réponses en temps réel. Cette évolution guide ainsi la construction de notre proposition.

La méthode proposée examine et caractérise la qualité des flux des capteurs pour les quatre entités définies dans la section précédente : *données*, *information*, *connaissance* et *intelligence*. Cette définition permet de faire une évaluation adaptative et sélective de la qualité à ces différents niveaux.

L'identification des éléments de la qualité de chaque niveau de la pyramide a été réalisée d'une façon globale pour les CPS. De cette manière, nous cherchons à concevoir une approche générale adaptée aux sous-systèmes des CPS. En fonction du sous-système auquel l'approche sera appliquée, les différents éléments présentés seront plus au moins pertinents. Une nouvelle notation à partir de vecteurs permet de structurer les résultats et les réutiliser éventuellement pour d'autres applications.

Les propositions introduites avec cette méthodologie sont résumées dans la liste ci-dessous. Cette liste cherche à répondre aux limitations identifiées dans les études de l'état de l'art.

1. Application générale pour tout sous-système d'un CPS
   (a) Identification exhaustive d'éléments de la qualité à partir des études précédentes de l'état de l'art et des cas d'étude
2. Introduction de la qualité des connaissances et de l'intelligence
3. L'humain est considéré dans les sous-systèmes d'IHM (Interfaces Humain-Machine)
   (a) Le sous-système est pris comme référence centrale par la méthodologie
   (b) Le besoin du temps réel relève l'humain de son rôle d'évaluateur de la qualité
4. Réutilisation des résultats
   (a) Résultats fournis sous forme de vecteurs
   (b) Définition des niveaux d'acceptation pour les éléments de la qualité (voir Section II.4.2)

Dans les sous-sections suivantes, l'évaluation de la qualité de chaque niveau est abordée. Les ensembles d'éléments identifiés pour chaque niveau sont accompagnés d'une notation basée sur des vecteurs.

II.3.1 Qualité des données

Cette mesure est composée à partir de quatre types d'imperfections dont un consensus relatif existe entre les chercheurs [MS96] :

1. Les données sont **erronées** (i_{err}) quand elles ont une valeur qui ne correspond pas avec la réalité. Dans cette catégorie se trouvent les *outliers*. Les *outliers* sont des valeurs qui sont anormales en relation avec le comportement observé dans la majorité des autres cas pour cet attribut. Il faudra faire une analyse pour déterminer si un *outlier* est une donnée erronée ou juste une valeur rare et correcte.
2. Une donnée est **incomplète** (i_{inc}) quand un attribut pertinent ou une valeur manque. Une valeur manquante peut être une valeur qui n'existe pas ou qui n'a pas été enregistrée.
3. Les données **imprécises** (i_{imp}) sont les valeurs qui ne peuvent être déterminées que par approximation : une disjonctive, une négation ou un intervalle. Aussi cela peut montrer une indisponibilité de la donnée, qui ne peut pas être considérée comme une valeur manquante.
4. Les données **incertaines** (i_{incer}) peuvent apparaître à cause de différentes sources, par exemple : des erreurs de mesure, l'intégration de données de multiples sources, des processus automatiques imprécis et le jugement humain imparfait. Le cas de la mesure de l'incertain est moins évident que les autres. Des théories mathématiques spécifiques ont été conçues pour mesurer l'incertain [Tod14].

Pour structurer les résultats de cette évaluation, toutes les imperfections mesurées (i_i) sont stockées dans un vecteur appelé $D\vec{Q}V$ (*Data Quality Vector*). L'équation II.4 représente ce vecteur.

$$D\vec{Q}V \in \{i_1 ... i_K\} \tag{II.4}$$

Ces quatre valeurs permettent de mesurer toutes les imperfections des données identifiées dans la littérature. Leur mesure doit être définie indépendamment pour chaque sous-

système étudié. Cela est dû à la grande diversité de types de données existantes dans les CPS. Présenter le résultat dans le vecteur $D\vec{Q}V$ permet de structurer les valeurs, de pouvoir opérer plus facilement avec elles ainsi que de les intégrer à l'étude de la propagation de la qualité.

II.3.2 Qualité de l'information

Une fois que l'information est extraite à partir des données, sa qualité peut être évaluée. Il y a différents points de vue sur cette problématique et en conséquence il n'y a pas une définition générale de la qualité de l'information [WP10] pour la communauté [2].

Quelques méthodologies se sont intéressées à la qualité de l'information des CPS, mais avec d'autres objectifs. Par exemple, la qualité de ces systèmes peut être utilisée comme indicateur pour le management [KL09b] ou dans le cadre de la fusion d'informations [RB10].

Comme aucune définition générale n'a été spécifiée pour les CPS, nous proposons de sélectionner les dimensions les plus adaptées parmi celles proposées auparavant par les méthodologies spécialisées dans d'autres domaines. Classiquement, les dimensions sont divisées en quatre groupes : intrinsèques, contextuelles, de représentation et d'accessibilité [WS96]. Des fois, d'autres groupes sont ajoutés pour résoudre des problématiques particulières, mais sans pour autant remettre en question ce premier classement [ZRM+16]. Cette catégorisation n'est pas adaptée à notre méthodologie puisque les IHM sont considérées comme des sous-systèmes au lieu de prendre en compte l'humain comme consommateur final de l'information. Par conséquent, le problème de la représentation des informations perd une partie de sa signification et de son importance. Ainsi, la référence pour la méthodologie est le sous-système étudié comme module élémentaire. C'est pour cela que les dimensions identifiées ont été catégorisées en trois groupes : intrinsèques (Table II.1), contextuelles (Table II.2) et extrinsèques (Table II.3).

Cette nouvelle catégorisation permet de considérer les dimensions utilisant le sous-système comme référence. Les dimensions intrinsèques représentent toutes les dimensions qui peuvent être mesurées quand seulement les spécifications du sous-système sont connues. Les dimensions contextuelles sont les dimensions qui peuvent être mesurées quand le sous-système est mis en contexte, c'est-à-dire que l'on connaît son application dans le système

2. Compilation de définitions pour les dimensions de la qualité de l'information. Consulté le 1 Octobre 2017. `http://dke.uqcloud.net/DataQualityPatterns/`

TABLE II.1: Dimensions Intrinsèques de la Qualité de l'Information dans les CPS

ID	**Nom**	**Description**
id_{sp}	Précision de la source	Mesure la variation du résultat pour des conditions identiques
id_{acc}	Exactitude	Mesure la différence entre l'information extraite et l'information réelle
id_{obj}	Objectivité	Mesure le degré selon lequel l'information est non biaisée et impartiale
id_{rep}	Réputation	Mesure le crédit de l'information en termes de contenu et de source
id_{obs}	Obsolescence	Mesure la validité de l'information dans le temps
id_{fre}	Fraicheur	Mesure à quel point l'information est récente
id_{tru}	Confiance	Mesure la confiance dans l'information
id_{acq}	Coût d'acquisition	Mesure le coût d'acquisition de l'information
id_{rea}	Lisibilité	Mesure combien la donnée utilisée pour obtenir l'information est non bruitée et compréhensible
id_{res}	Résolution	Mesure la qualité de l'échantillonnage de la donnée utilisée pour obtenir l'information
id_{itg}	Intégrité	Mesure combien l'information est complète et la disponibilité du fournisseur
id_{cns}	Consistance	Mesure à quel point l'information est présentée toujours avec le même format
id_{uni}	Unicité	Mesure le degré d'unicité de l'information

où il est installé. Finalement, les dimensions extrinsèques prennent en compte la connexion du sous-système avec d'autres sous-systèmes. La Figure II.1 explique graphiquement cette catégorisation pour le cas d'un capteur dont le contexte est donné par où il est installé et ce qu'il mesure.

Le résultat de l'évaluation de la qualité de l'information est stocké dans un vecteur appelé $I\vec{Q}V$ (*Information Quality Vector*) représenté par l'équation II.5. Les composantes (d_i) de ce vecteur sont les dimensions évaluées.

$$I\vec{Q}V \in \{d_1...d_L\} \tag{II.5}$$

Cette structure en forme de vecteur va nous permettre d'organiser les résultats et de pouvoir réaliser des opérations facilement. Dans le cas de $I\vec{Q}V$, les composantes peuvent être

TABLE II.2: Dimensions Contextuelles de la Qualité de l'Information dans les CPS

ID	Nom	Description
cd_{rp}	Précision réelle	Mesure la variabilité d'une mesure avec des conditions identiques à cause de l'utilisation du système
cd_{cla}	Clarté	Mesure la compréhensibilité de l'information entre autres sources d'information
cd_{val}	Valeur ajoutée	Mesure les avantages et bénéfices que l'information apporte
cd_{tim}	Opportunisme	Mesure à quel point l'information est attendue par le système
cd_{err}	Erronée	Mesure si l'information représente un état possible du sous-système
cd_{cmt}	Complétude	Mesure si l'information est connue dans un contexte complet
cd_{cnc}	Concision	Mesure à quel point l'information est représentée de forme compacte
cd_{vol}	Volume	Mesure si le volume d'information est approprié pour la tâche à accomplir.
cd_{bel}	Crédibilité	Mesure le degré dans lequel l'information est regardée comme vraie et crédible

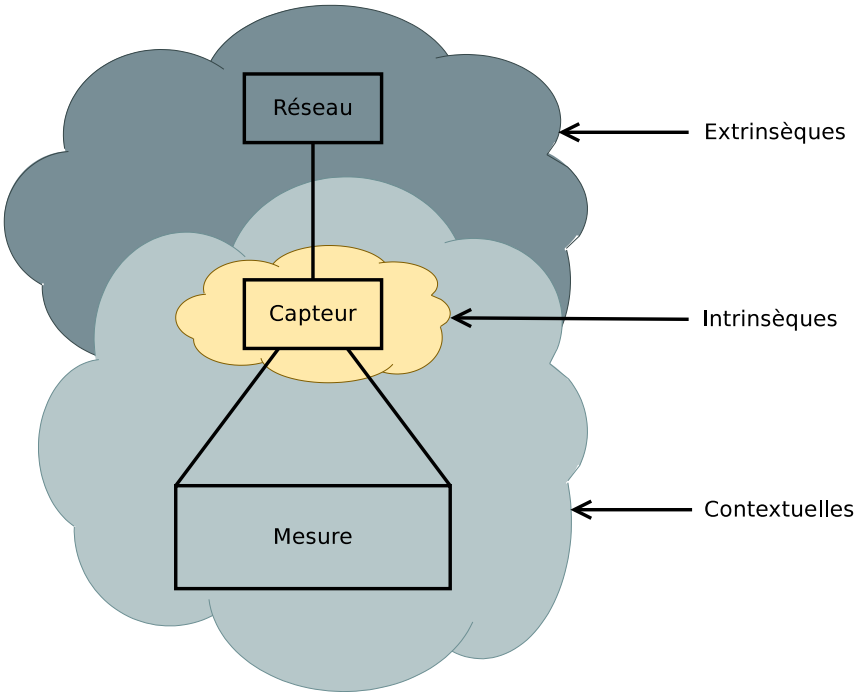

FIGURE II.1: Schéma des catégories des dimensions de la qualité de l'information.

TABLE II.3: Dimensions Extrinsèques de la Qualité de l'Information dans les CPS

ID	Nom	Description
ed_{acc}	Accessibilité	Mesure la disponibilité ainsi que la facilité et la vitesse de récupération de l'information
ed_{sec}	Sécurité	Mesure le niveau de protection de l'information
ed_{eu}	Facilité d'utilisation	Mesure la facilité d'utilisation de l'information pour différents objectifs
ed_{man}	Manipulation	Mesure la facilité de manipulation de l'information
ed_{int}	Interprétable	Mesure si l'information est dans la langue appropriée et que les symboles, unités et définitions soient clairement exprimés
ed_{cmp}	Compatibilité	Mesure combien l'information est compréhensible pour les différents sous-systèmes
ed_{for}	Format	Mesure combien l'information respecte le format attendu
ed_{und}	Compréhension	Mesure la facilité de compréhension de l'information
cd_{red}	Redondance	Mesure la quantité de sous-systèmes qui produisent la même information
ed_{coh}	Cohérence	Mesure à quel point l'information est logique par rapport aux autres informations

séparées dans les trois catégories qui ont été définies afin de faciliter leur identification. Cette distinction sera utilisée dans les cas d'étude pour faciliter la compréhension de l'application (Équation III.2 et Équation III.12).

II.3.3 Qualité de la connaissance

La qualité des connaissances et de l'intelligence pour les CPS n'a pas été traitée directement dans la littérature. Jusqu'à présent, aucune méthodologie n'a évalué la qualité pour l'ensemble de la pyramide DIKW. À partir de ce constat, nous proposons également une définition adaptée de cette qualité.

La connaissance est définie par l'application directe de l'intelligence sur l'information. Cela implique que la qualité de la connaissance est considérablement influencée par les qualités de l'information et de l'intelligence. Pour cela, nous proposons la définition de la qualité de la connaissance en utilisant un vecteur $K\vec{Q}V$ (*Knowledge Quality Vector*) représenté dans

l'équation II.6.

$$K\vec{Q}V \in \{I\vec{Q}V_1 ... I\vec{Q}V_M, W\vec{Q}V, f_1 ... f_N\} \tag{II.6}$$

Ce vecteur comporte la qualité des flux d'information ($I\vec{Q}V_i$) et de l'intelligence ($W\vec{Q}V$) utilisées pour sa définition en plus de trois autres facteurs complémentaires (f_i) qui ont été identifiés. La qualité de l'intelligence est introduite dans la section suivante (Section II.3.4). Ces facteurs (f) sont présentés et définis de la manière suivante :

- **Complétude (f_{com}) :** Mesure si les informations nécessaires par l'intelligence sont disponibles.
- **Coût d'erreur (f_{err}) :** Mesure le coût qu'une connaissance erronée peut produire.
- **Pertinence (f_{rel}) :** Mesure à quel point la connaissance est applicable et utile pour la tâche traitée.

Grâce au vecteur $K\vec{Q}V$, la qualité des entités - information et intelligence - utilisées pour créer une connaissance précise peut être tracée. La traçabilité de la qualité permet entre autres d'identifier les sources des informations qui ont produit une connaissance de « mauvaise » qualité ou si l'intelligence doit être améliorée. Cette structure permet, comme dans les autres cas, d'opérer facilement avec les résultats des mesures analysées. Ces calculs peuvent avoir différents objectifs comme la détection d'anomalies ou la représentation de la qualité agrégée avec une valeur globale de qualité.

II.3.4 Qualité de l'intelligence

Du fait que la qualité de l'intelligence n'a pas été traitée auparavant dans la littérature, une proposition de sa mesure est introduite par notre travail. Pour la définition de sa qualité, six aspects (a) ont été identifiés (Table II.4).

Le résultat de l'évaluation de ces aspects est stocké, de la même façon que pour les autres éléments, sur un vecteur WQV (*Wisdom Quality Vector*) dont les composantes sont les aspects évalués (a_i) :

$$W\vec{Q}V \in \{a_1 ... a_P\} \tag{II.7}$$

TABLE II.4: Aspects de la Qualité de l'intelligence dans les CPS

ID	Nom	Description
a_{exp}	Expérience	Mesure combien de fois l'intelligence a prouvé sa validité avec son utilisation
a_{con}	Confidence	Mesure le crédit de l'intelligence en termes de contenu et de source
a_{acc}	Accessibilité	Mesure la disponibilité et la facilité de récupération de l'intelligence et si elle est modifiable
a_{int}	Interprétable	Mesure à quel point l'intelligence est définie clairement avec le langage. les symboles et les unités appropriées
a_{sec}	Sécurité	Mesure le niveau de protection de l'intelligence et ses restrictions d'accès
a_{com}	Complétude	Mesure à quel point l'intelligence prend en compte les variables qui l'influencent

Le vecteur $W\vec{Q}V$ permet d'intégrer la qualité de l'intelligence à la qualité de la connaissance à travers le vecteur $K\vec{Q}V$. Grâce aux aspects identifiés, les améliorations de l'intelligence peuvent avoir un objectif plus précis. Par exemple, quand un algorithme a été créé avec une seule expérience, un indice pour améliorer sa qualité peut être la réalisation d'une plus grande quantité de tests.

II.3.5 Méthodologie d'évaluation de la qualité

À partir de tous les termes décrits dans la Section II.1 et des mesures de qualité pour les niveaux de la pyramide DIKW introduits dans la Section II.3, nous proposons une méthodologie d'évaluation de la qualité. Son domaine d'application concerne les flux créés par un sous-système appartenant à l'architecture d'un CPS. Cela veut dire, une unité communiquant avec d'autres en utilisant des connexions qui respectent les spécificités techniques du système. Par exemple, dans un navire où un CPS assiste la navigation, un capteur qui communique avec d'autres sous-systèmes comme un écran dans la passerelle utilisé comme IHM peut être objet d'étude par cette méthodologie.

En partant des quatre niveaux de la pyramide DIKW - Donnée, Information, Connaissance et Intelligence -, nous pouvons identifier les transformations qu'une donnée peut subir jusqu'à la définition d'une connaissance. Sur le schéma de la Figure II.2, le processus

de définition des connaissances est représenté en gras. Les données deviennent information quand elles sont mises en contexte et une information a comme résultat une connaissance grâce à l'application de l'intelligence.

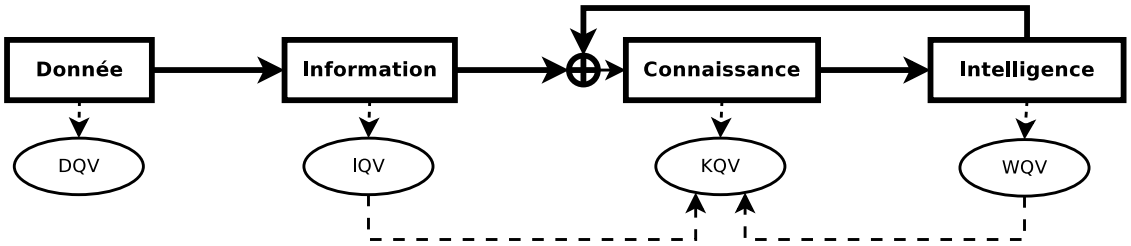

FIGURE II.2: Méthodologie de mesure de la qualité dans le processus de définition d'une connaissance.

En groupant les définitions de qualité correspondant à chaque entité de la pyramide, une méthodologie complète de la mesure de la qualité des flux créés par un sous-système est spécifiée. Chaque entité de la pyramide est accompagnée par un vecteur représentant sa qualité. La qualité d'une donnée est résumée par un vecteur $D\vec{Q}V$, pour l'information par un vecteur $I\vec{Q}V$, pour la connaissance par un vecteur $K\vec{Q}V$ et pour l'intelligence par un vecteur $W\vec{Q}V$. La méthodologie complète pour la mesure de la qualité est représentée dans la Figure II.2.

Chaque élément évalue indépendamment sa qualité respective sauf la connaissance puisque sa qualité sera fortement influencée par la qualité de l'information et par la qualité de l'intelligence qui la définissent. La répercussion de $I\vec{Q}V$ et $W\vec{Q}V$ sur $K\vec{Q}V$ est aussi indiquée sur le schéma.

Les données qui peuvent être traitées varient en fonction du système et du sous-système. Par exemple, les protocoles de communication utilisés dans les navires sont variés et souvent coexistent dans une installation. Ainsi, un sous-système peut créer des données avec différents formats par exemple comme une chaîne de caractères, une valeur entière ou avec une valeur de virgule flottante.

L'intelligence dans les CPS est souvent associée à des algorithmes ou guides qui permettent de définir des connaissances. Cette intelligence peut varier énormément en fonction de ce qu'elle représente et de la manière dont elle traite le problème. Par exemple, une intelligence qui transforme un volume de liquide en poids est basée sur des principes physiques tandis qu'une autre qui indique les mesures de sécurité à prendre en fonction de la route de navigation est basée sur une loi et l'expérience. De la même façon que pour les autres entités, les connaissances varient. Elles peuvent donner comme résultat par exemple une réaction,

une valeur ou une information complémentaire.

La variabilité des entités à tous les niveaux de la pyramide fait que la manière de mesurer chaque élément de qualité n'est pas unique. Pour chaque cas précis, il faut l'intervention d'un humain afin d'étudier le système et choisir la technique la plus adaptée à chaque élément de la qualité.

Vu que chaque unité de donnée, d'information, de connaissance et d'intelligence est accompagnée d'un vecteur représentant sa qualité, cette technique d'évaluation de la qualité peut être adaptée aisément à plusieurs problématiques. Deux des domaines les plus intéressés par les mesures de qualité sur les CPS sont la fusion d'information et l'aide à la décision. Dans ce qui suit, nous présentons comment la méthodologie proposée peut être utilisée pour la détection d'anomalies et des cyberattaques en tant que situations particulières.

II.4 Détection des anomalies

Les CPS étant des cas très particuliers de systèmes de par leur nature critique, il devient nécessaire de les protéger. Comme cela a été présenté dans la Sous-section I.4.4, les systèmes de protection pour les SI classiques ne sont pas applicables dû à leurs caractéristiques. D'autres solutions sont nécessaires.

Dans les dernières années, le domaine de la qualité a commencé à s'intéresser aux CPS. Des approches spécifiques ont été créées pour résoudre des problématiques précises. Les travaux de cette thèse proposent une méthodologie d'application générale qui estime des mesures de qualité adaptées aux données et aux informations. Ces éléments groupent des caractéristiques qui peuvent être intéressantes pour trouver des indices de détection d'anomalies. En conséquence, les méthodologies de mesure de la qualité se présentent comme des candidats pour faire partie des systèmes de protection des CPS. Leur principal objectif est alors : l'identification d'anomalies basée sur la détection et la catégorisation des altérations de la qualité.

II.4.1 Hypothèses

Normalement les CPS, par leur architecture fixe et leur nature, ont une quantité finie d'états de fonctionnement. Cette particularité fait que différentes valeurs de qualité peuvent

être définies comme étant normales. De ce fait, un changement dans ces attributs pourra toutefois définir les flux comme anormaux et déclencher une alarme.

Ainsi, une catégorisation des anomalies détectées est désirée. Pour cela, il faudra faire une analyse sur les anomalies connues et leur effet sur la qualité. Les différents éléments qui définissent la qualité pourraient aider à atteindre cet objectif.

Avec les points identifiés, nous pouvons énoncer deux hypothèses :

Hypothesis 1. *Les anomalies ont un impact sur la qualité et cela va permettre leur détection.*

Hypothesis 2. *Les éléments de la qualité affectés par les anomalies sont un indice pour les catégoriser.*

Ces deux hypothèses seront examinées par la suite. Deux cas d'étude permettront d'analyser leur pertinence dans le Chapitre III.

II.4.2 Approche de détection et catégorisation

Pour prouver les hypothèses présentées, il faut examiner chaque sous-système identifié comme module élémentaire, à la fois séparément et comme composant d'un système plus large pour déterminer leurs faiblesses et les anomalies qu'ils peuvent subir. Chaque scénario sera examiné pour déterminer s'il affecte la qualité et quels éléments sont affectés. Pour automatiser ce travail, un protocole doit être défini. Par conséquent, une approche de détection d'anomalies à partir des mesures de qualité est présentée.

Cette approche est illustrée dans la Figure II.3 et elle est constituée de trois parties : la mesure de la qualité des flux des données produites par un sous-système, la création des AL (*Agreement Levels*) et l'évaluation d'acceptation des AL.

La première partie consiste à appliquer la méthodologie décrite dans la section II.3.5 aux données et à l'information d'un sous-système. Nous observons également comment d'autres flux d'information ou connaissances peuvent s'avérer nécessaires pour la mesure de certaines dimensions. Par exemple, la mesure de la cohérence tient compte de l'état d'autres sous-systèmes.

La deuxième partie consiste à fixer des AL pour la qualité avec un mécanisme d'apprentissage. Cette partie prend comme entrée le résultat de l'évaluation de la qualité et

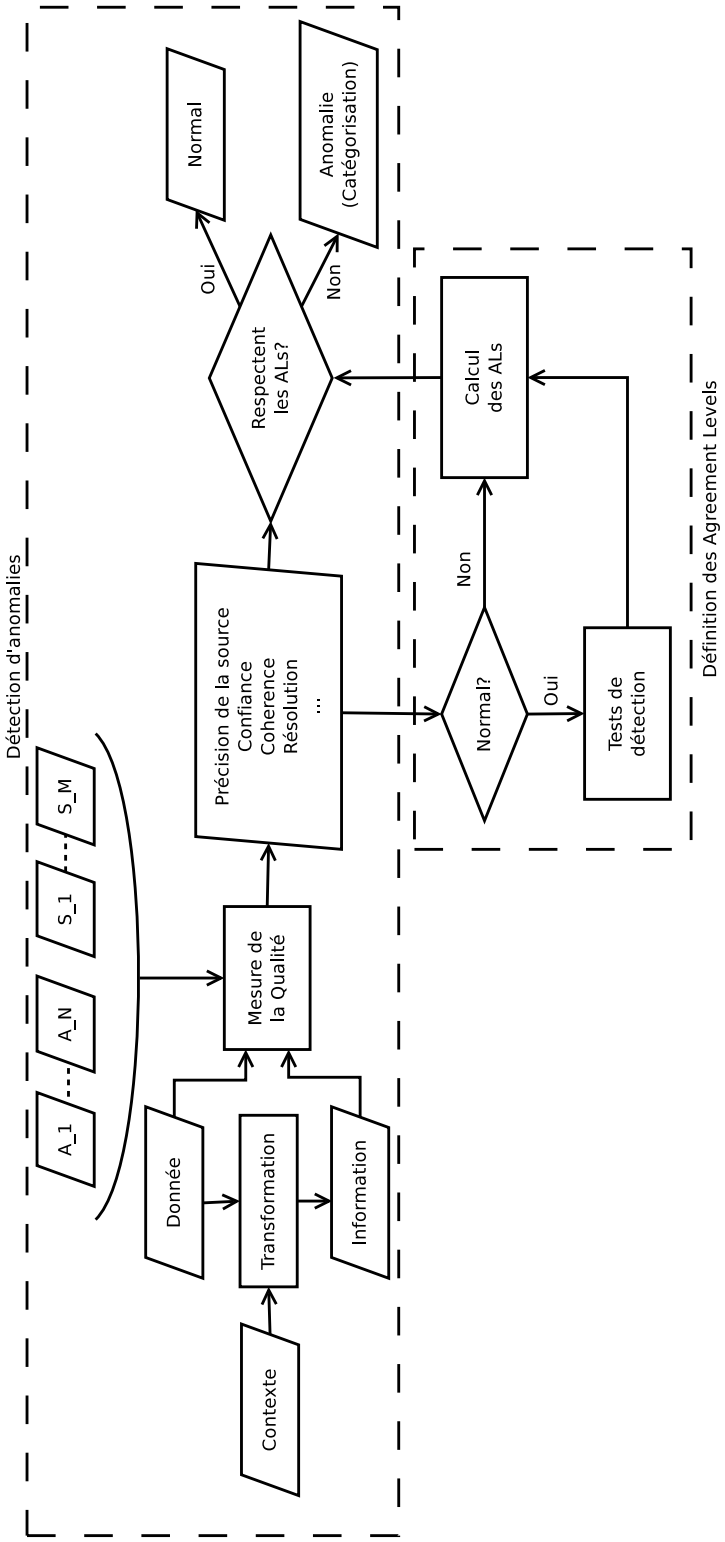

FIGURE II.3: Schéma explicatif de l'approche proposée pour la détection et catégorisation d'anomalies.

cherche à fixer des valeurs d'acceptation appelées AL afin de séparer les valeurs de qualité représentatives des scénarios normaux et des scénarios anormaux. Le calcul applique les méthodes présentées dans la Section I.2. En fonction des mesures définies pour chaque élément, la pertinence des méthodes varie.

Finalement, la troisième partie identifie les flux anomaux en regardant si les mesures de qualité réalisées à un moment précis de temps sont vérifiées par les ALs. Ainsi pour chaque détection, une catégorisation du scénario accompagne l'alerte. Cette étape est possible grâce à l'analyse des éléments affectés.

Tous les algorithmes utilisés dans chaque étape devront prendre en compte les restrictions du sous-système et le système où ils sont implémentés. La sécurité dans les CPS est examinée à plusieurs niveaux comme cela a été présenté dans la Section I.4.6 et la méthodologie peut se mettre en place au moyen de différents choix. Par exemple, elle pourra s'appliquer directement avec des flux des données ou à travers des logs de surveillance. Ainsi, elle pourra être implémentée sur des *firewalls* ou dans des IDS (*Intrusion Detection System*).

II.4.3 Implémentation de l'approche

Avant de pouvoir se servir de l'approche proposée pour la détection d'anomalies, il faut modéliser le système étudié et établir plusieurs paramètres. L'utilisation de l'approche dans un système concret peut être décrite en plusieurs étapes :

1. Identification des sous-systèmes
2. Enregistrements de données
   - Quantité de données représentatives
   - Structure des logs
3. Mesure de la qualité
   - Spécifications (Data-sheets)
   - Étude du système
4. Identifications des ALs
   - Identification des états de normalité
   - Calcul des ALs avec différentes techniques
5. Conception des systèmes d'alertes basées sur les ALs
   - Logiciel

— Visualisation des alertes

6. Validation de l'installation
   — Identification de risques/vulnérabilités
   — Test de détection
   — Adaptation des ALs
   — Évaluation de l'impact sur le système

Les trois premières étapes décrivent la méthodologie pour la mesure de la qualité tandis que les trois derniers groupes représentent les pas d'application nécessaires pour l'approche de détection d'anomalies basée sur les mesures de qualité. Tous ces points sont expliqués ci-dessous avec un exemple.

Les systèmes navals ont été identifiés comme des systèmes qui pourraient bénéficier de cette méthodologie. Classiquement, les navires utilisent différents modules qui se connectent dans un bus pour communiquer entre eux. Dans la Figure II.4, un exemple de comment ces modules peuvent être connectés est montré.

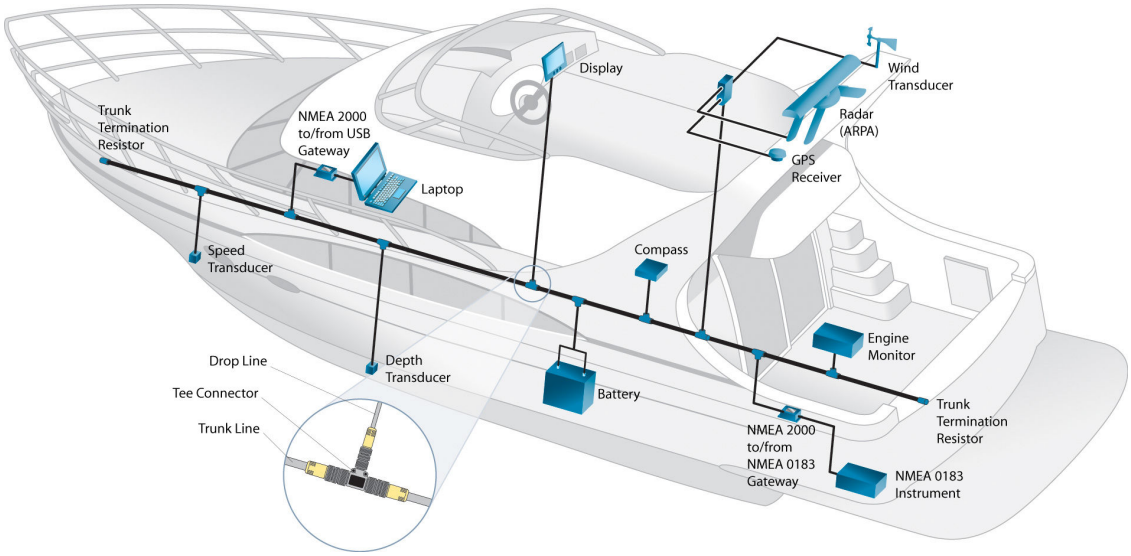

FIGURE II.4: Exemple de connexion de modules CPS d'un bateau [Nat04].

Chacun de ces modules crée et reçoit des flux de données et d'informations. Ainsi, la décomposition en modules doit être associée à l'identification cohérente des sous-systèmes. Parfois, ces modules sont composés de plusieurs éléments. Par exemple, la station météo est composée par plusieurs capteurs qui génèrent des informations différentes. De ce fait, il est nécessaire d'analyser chaque module afin d'identifier s'ils peuvent être décomposés en plusieurs sous-systèmes.

L'étape d'enregistrement de données est essentielle avec le but de les analyser « *off-line* ». Une quantité représentative de données doit être contenue dans les logs pour qu'ils puissent décrire les différents comportements du système. La structure des logs doit faciliter la génération de *datasets*. Ces datasets doivent être bien documentés pour pouvoir les étudier en détail. Ainsi, la machine qui enregistre ces logs possède des caractéristiques similaires au système où l'approche de détection sera implémentée. Dans les navires qui ont une certaine taille, un module appelé VDR (*Voyage Data Recorder*) fonctionne en mode de « boîte noire ». Les logs créés par le VDR pendant différentes situations pourront être utilisés.

À partir des spécifications des sous-systèmes, les caractéristiques du système général et des *datasets*, il est possible de faire une analyse exhaustive de chaque flux de données et d'information. Quant au cas du navire, il faut regarder les *data-sheets* des modules, les caractéristiques du navire et les données enregistrées par le VDR. Ainsi, il faut prendre en compte où est faite la détection parce que les conditions de mesure peuvent varier en fonction du point d'évaluation. Cette étude détaillée permet d'identifier tous les éléments de la qualité qui pourront être mesurés et comment le faire. Ce travail doit être réalisé pour chacun des sous-systèmes identifiés dans la première étape. Par exemple, nous pouvons étudier la mesure de profondeur produite par un capteur à ultrason qui est installé dans la coque d'un navire. Après avoir examiné le *datasheet* et où le capteur est installé, il est possible de déterminer si les données sont erronées et comment calculer la précision du capteur en fonction de la profondeur.

L'évaluation de la qualité sur les *datasets* enregistrés dans le deuxième pas permet d'étudier les valeurs sur l'ensemble du système. Cette analyse sert à faire une identification des valeurs normales et anormales des éléments de la qualité. Afin de séparer ces deux cas, nous utiliserons les ALs qui seront calculés avec les techniques présentées dans la Section I.2. Pour l'exemple de la transmission de la mesure de profondeur dans le réseau du navire, il est approprié de limiter les données erronées autorisées dans un intervalle de temps.

Une fois que les mesures et les ALs ont été définis, il faut développer des solutions logicielles que l'on pourra installer dans le système en respectant ses contraintes. Ces solutions peuvent avoir une partie logicielle et une partie matérielle. L'objectif du développement est la détection et visualisation des alertes dans le système. Pour un navire, cela correspond aux logiciels des ordinateurs de bord et aux systèmes d'alerte qui sont installés dans la passerelle et différents points du navire.

Finalement, des scénarios de danger du système étudié doivent être identifiés. Des tests

de détection dans le contexte de ces scénarios doivent être réalisés afin de pouvoir valider la méthodologie. Une adaptation des ALs est possible à ce point pour améliorer leur détection et réduire les faux positifs. Afin de valider complètement l'application de l'approche, une analyse de l'impact qu'elle a dans le système doit être réalisée pour éviter de nuire aux caractéristiques du système à cause de ses contraintes. Un problème récurrent de la mesure de la profondeur est que le capteur peut être bloqué par différents objets comme des coquilles ou des algues ce qui produit des mesures fausses. C'est pour cela que ce scénario doit faire partie des tests de détection pour un navire. Même s'il n'est pas étudié directement dans cette thèse, il existe une forte similitude avec le blocage du capteur à ultrason analysé dans le Chapitre III.

Cette approche peut ainsi s'intégrer au processus d'amélioration de la qualité proposé par la méthodologie TDQM et présenté dans la Section I.1.3 avec quelques petites modifications. Nous observons la façon dont l'étape de définition de la qualité a été décomposée en deux sous-étapes. Ainsi, les définitions des éléments ont été adaptées aux besoins. Le processus de l'approche présentée est introduit dans la figure II.5.

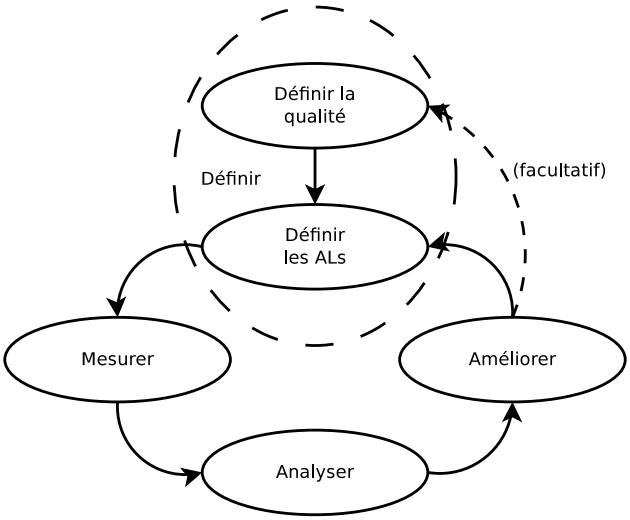

FIGURE II.5: Processus d'amélioration de la qualité de la méthodologie proposée.

De manière similaire au processus classique, il consiste en 4 étapes qui se répètent en boucle :

1. La première étape de **définition** est décomposée en deux sous-étapes :
   (a) La première sous-étape consiste à définir la **qualité** pour notre cas de système particulier ainsi que la façon de la mesurer.

   (b) La deuxième sous-étape définit des ***Agreement Levels*** pour les mesures identifiées.

2. La deuxième étape consiste à **mesurer** tous les attributs identifiés de la qualité.
3. Ensuite, une **analyse** est réalisée à partir des mesures et des AL définis. Quand une détection se déclenche, elle doit être catégorisée pour identifier la source de l'anomalie.
4. Finalement, une fois identifiées l'anomalie et sa source, elles seront corrigées pour **améliorer** la qualité actuelle et future.

La principale différence par rapport au processus d'amélioration de la qualité classique, présenté dans la Section I.1.3, est que cette approche prend comme point de départ des mesures de qualité prédéfinies et qu'elle introduit l'idée des ALs. Seulement dans certains cas de notre cas d'étude, les mesures de qualité pourront changer et elles le feront en cas de nécessité. C'est pour cela que le passage par la sous-étape de définition de la qualité est facultatif.

II.5 Conclusion

Il existe une grande variété de termes et définitions qui varient en fonction du domaine d'application. C'est pour cela que ce chapitre commence avec l'identification des termes nécessaires pour la compréhension du contexte de la thèse et avec le choix des définitions qui s'adaptent le plus à nos travaux.

Le chapitre I affirme que les CPS ont besoin de solutions spécifiques à cause de leurs caractéristiques distinctives. L'état de l'art nous a montré que la mesure de la qualité sur les CPS est un défi actuel pour la recherche. Il existe plusieurs méthodologies adaptées aux problématiques spécifiques, mais il n'existe aucune application générale pour n'importe quel sous-système d'un CPS.

À partir des travaux existants ainsi de notre cas d'étude, nous avons analysé les définitions afin de choisir celles qui s'ajustent aux sous-systèmes des CPS. Par la suite, nous avons parcouru les quatre niveaux de la pyramide DIKW - Donnée, Information, Connaissance et Intelligence - avec comme objectif d'identifier les éléments qui décrivent leur qualité. Quatre ensembles de 4 imperfections, 32 dimensions, 3 facteurs et 6 aspects ont été identifiés et décrits dans ce chapitre. La définition des quatre vecteurs pour le résultat des mesures

respectives - $D\vec{Q}V$, $I\vec{Q}V$, $K\vec{Q}V$ et $W\vec{Q}V$ - permet une analyse simple de la qualité. Chaque unité de donnée, information, connaissance ou intelligence peut de cette façon être toujours accompagnée de sa qualité correspondante.

Ainsi, tous les éléments utilisés dans le travail ont été définis avant d'introduire une nouvelle approche pour la mesure de la qualité qui devient une approche de détection d'anomalies avec l'application des deux hypothèses. Un guide avec les étapes d'implémentation de la méthodologie pour l'évaluation de la qualité et l'approche pour la détection basée sur elle a été présenté. Cela permet de suivre une procédure d'application sur un système réel.

Dans le chapitre suivant, ces méthodologies seront appliquées à deux cas d'étude différents. Faisant partie de notre problématique, l'application navale est toujours mise en contexte. Cela permettra d'évaluer leur pertinence, leur validité et confirmer les deux hypothèses présentées.

Chapitre III

Application de la méthode

Sommaire

III.1 Introduction 72

III.2 Système cyber-physique de deux cuves 72

III.2.1 Description 73

III.2.2 Mesure de la qualité 79

III.2.3 Détection d'anomalies (actes malveillants) 83

III.2.4 Catégorisation d'anomalies (actes malveillants) 88

III.3 Drones aériens 89

III.3.1 Description 92

III.3.2 Mesure de la qualité 95

III.3.3 Détection d'anomalies (actes malveillants) 101

III.4 Comparaison et positionnement par rapport à d'autres méthodes 105

III.4.1 Comparaison avec d'autres méthodes de mesure de la qualité . . . 105

III.4.2 Positionnement de l'approche par rapport aux autres méthodes de détection d'anomalies 107

III.5 Conclusion 108

Ce chapitre présente deux cas d'étude qui servent à valider l'application de l'approche présentée dans le chapitre précédent. Tout d'abord, un système SCADA avec deux cuves est étudié en tant que sous-système critique. Un deuxième cas d'étude concerne les données produites par les capteurs embarqués dans un drone aérien. Ce type de drones est de plus en plus utilisé lors des missions opérationnelles navales, ce qui représente de nouveaux risques.

III.1 Introduction

Une première étape a consisté à rechercher des *datasets* publiques mettant à disposition des données de navigation comprenant différents scénarios qui nous permettraient de valider notre méthodologie. Seulement deux *datasets* de systèmes SCADA ont été trouvés. Le premier a été créé par la simulation d'une canalisation de gaz [MTT15] et le deuxième grâce à une sonde installée dans une usine de traitement d'eau [GAJM16]. Aucun de ces deux jeux de données n'est en lien avec le milieu naval et un seul contient des données réelles. Le *dataset* de la plate-forme réelle n'est pas accompagné d'informations détaillées sur les capteurs ce qui complique sa réutilisation. De plus, celui-ci ne prend pas en compte ni les sabotages ni les autres types d'anomalies.

L'inexistence de *datasets* adaptés à notre problématique a motivé la création de notre propre jeu de données. Après avoir vérifié l'impossibilité de récupérer des logs réels de différents navires, d'autres alternatives ont été analysées. Le besoin d'avoir des données réelles et d'éviter les simulations informatiques ont restreint les solutions possibles. Deux sous-systèmes fortement liés aux navires actuels ont été identifiés. Dans un premier temps, nous avons travaillé avec un sous-système de deux cuves, pour plus tard nous intéresser aux drones aériens. Ces deux sous-systèmes sont décrits et étudiés en détail dans les sections suivantes.

III.2 Système cyber-physique de deux cuves

Un système composé par deux cuves [1] constitue notre premier cas d'étude [MLBP17b]. Dans le contexte naval, cette configuration peut représenter différents systèmes critiques. De grandes cuves sont souvent utilisées comme châteaux d'eau ou comme réservoirs de combustible dans les ports. D'autre part, les navires nécessitent plusieurs cuves pour stocker différents liquides nécessaires à leur fonctionnement. Par exemple, elles peuvent être utilisées dans les systèmes de propulsion comme réservoir d'essence ou d'huile ou avec d'autres objectifs comme le stockage d'eau potable pour l'équipage.

1. Plate-forme utilisée pour le challenge Cyberdef 2015.

III.2.1 Description

La présentation de la plate-forme est réalisée en trois parties :

1. la description du matériel employé,
2. son fonctionnement
3. et les risques que la plate-forme peut subir.

Description

Le système est composé de deux cuves : une cuve principale de 7 litres et une cuve secondaire d'un volume de 9 litres. Pour le cas d'étude, elles sont remplies d'eau. Un trou au fond de la cuve principale permet de simuler une consommation constante d'eau. La cuve secondaire sert à incrémenter la capacité totale du système. Pour cela, lorsque la pompe 1 est utilisée pour faire le transfert de liquide vers la cuve principale, la pompe 2 remplit la cuve secondaire à partir d'une cuve de récupération (Figure III.1).

Le contrôle et le monitoring du niveau d'eau se font grâce à différents capteurs et automates qui sont connectés en réseau. Ces sous-systèmes sont connectés à un PLC, permettant d'interconnecter tous les composants et d'accéder à la plate-forme depuis un réseau externe (Figure III.1).

Deux types de capteurs sont utilisés pour la mesure du volume du liquide. Dans la cuve principale, un capteur à ultrason est installé sous le couvercle de la cuve. Ce capteur est largement utilisé dans l'industrie pour la mesure de distances ou la détection de présence. Il a été calibré avec les dimensions de la cuve, en donnant une valeur de référence, comportant 10 000 pas. À cet égard, les valeurs récupérées varient entre 10 000 pour une cuve vide et 0 pour une cuve pleine. Dans la cuve secondaire, seulement quatre niveaux peuvent être mesurés : 1,25L, 3,35L, 8L et 9L. Ces mesures sont faites avec quatre flotteurs installés à différentes hauteurs qui activent un interrupteur différent pour chacun.

Les deux cuves possèdent deux trous sur le dessus qui servent de système de sécurité. Ceci permet d'évacuer l'eau sans causer le moindre dégât, par exemple abîmer le capteur à ultrason. Dans la cuve secondaire, ce niveau de sécurité correspond au niveau le plus haut, mesuré par le flotteur 3.

L'ensemble du système est contrôlé par un PLC avec un module d'extension qui permet d'ajouter des entrées analogiques au modèle de base. Le contrôle du PLC peut se faire depuis

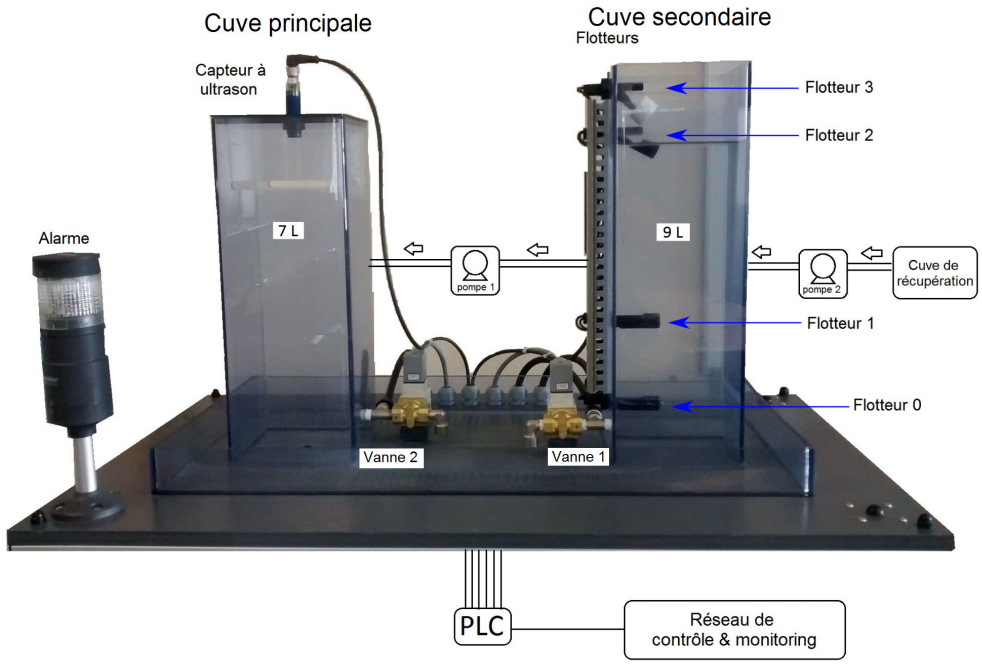

FIGURE III.1: Schéma de la plate-forme des cuves.

un écran tactile installé sur la plate-forme ou bien depuis une machine connectée au réseau. La communication avec le PLC doit se faire avec le protocole Modbus/IP [MOD96]. Pour la mise en place du réseau, un commutateur réseau (en anglais *switch*) est utilisé. Un schéma complet du réseau est présenté dans la Figure III.2.

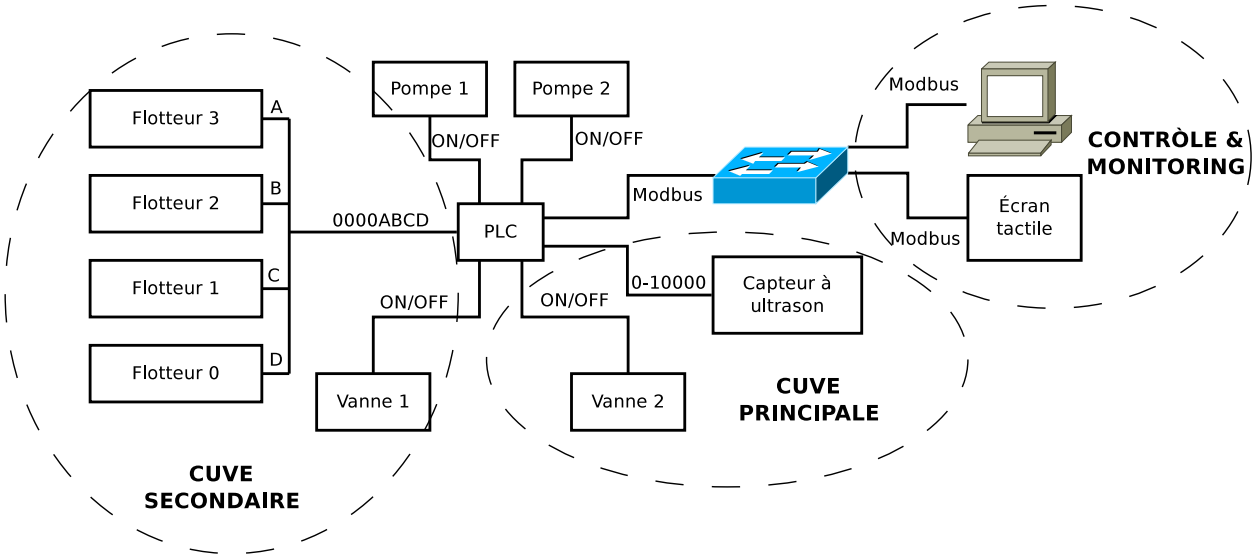

FIGURE III.2: Diagramme du réseau de la plate-forme.

Fonctionnement

La plate-forme a deux modes de fonctionnement : un mode manuel et un mode automatique. Le mode manuel permet d'activer et de désactiver les automates avec des commandes envoyées depuis une machine connectée en réseau ou depuis l'écran tactile. Seulement quelques commandes sont bloquées le cas échéant pour raison de sécurité, en général afin d'éviter qu'une pompe ne fonctionne à vide et qu'elle puisse s'abîmer.

Le mode automatique, quant à lui, met en route un fonctionnement préprogrammé. Le niveau dans la cuve principale sera maintenu entre 1,4 et 5,6 litres, ce qui correspond respectivement aux pas 8000 et 2000 mesurés par le capteur à ultrason. Ce fonctionnement sera altéré seulement dans le cas où le niveau de la cuve secondaire n'assurerait pas un fonctionnement en sécurité. De la même façon, le niveau de la cuve secondaire est maintenu entre les niveaux 3,3 et 8 litres.

Un exemple de signaux capturés pendant un fonctionnement en mode automatique est montré dans la Figure III.3. Les deux premiers graphiques regroupent les signaux correspondants à la cuve principale tandis que les deux derniers correspondent à la cuve secondaire. Grâce à la courbe du capteur à ultrason, nous observons que le point de départ est une cuve vide (la distance entre le capteur et le liquide est maximale). Les graphiques des pompes permettent de visualiser quand celles-ci sont en fonctionnement et que les cuves se remplissent. Le troisième graphique avec les états des flotteurs permet d'apercevoir le niveau du liquide présent dans la cuve secondaire. Les vannes, celle installée dans la cuve principale et celle installée dans la cuve secondaire, ne sont pas utilisées dans ce mode de fonctionnement.

Anomalies, sabotages et dysfonctionnements étudiés

Afin de pouvoir évaluer la validité de la méthode, plusieurs scénarios représentant des risques à différents niveaux pour le système ont été étudiés. Afin de représenter le mode *offline* de la plate-forme, un *dataset* a été créé. Ce *dataset* est décrit en détail dans l'Annexe A. Les signaux collectés du *dataset* correspondent aux risques classifiés en deux catégories : dysfonctionnements et sabotages.

Les dysfonctionnements sont tout type de situations dans lesquelles le système ne fonctionne pas de la façon attendue et peuvent être à l'origine de risques de dysfonctionnement systémique. Un exemple de scénario analysé est le cas de gouttes d'eau qui se créent sur la surface du capteur à cause de l'humidité. Celles-ci peuvent réduire la durée de vie du capteur

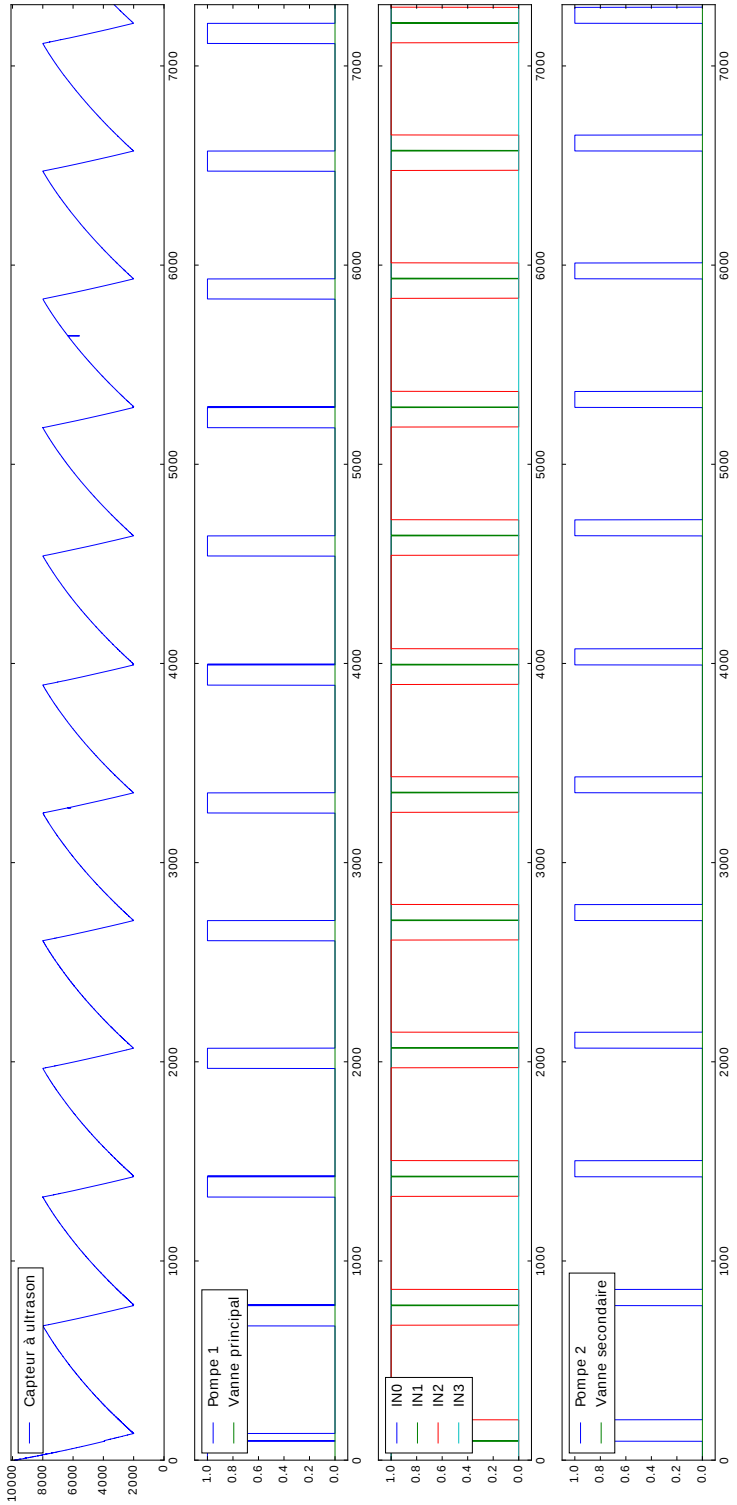

FIGURE III.3: Représentation des données pendant un fonctionnement normal.

et modifier la mesure. Un autre exemple est un fil abîmé, qui produit une perte d'information lors de la transmission.

Les sabotages sont des dysfonctionnements provoqués dans le but d'altérer le fonctionnement normal du système. Les cyberattaques sont un type particulier de sabotage et elles ont été identifiées dans le Tableau A.3 comme une attaque par déni de service (DoS) et une attaque d'usurpation d'identité (*spoofing*).

Les scénarios et leur description sont présentés dans la liste suivante :

- Dans un navire, les cuves sont un excellent endroit pour cacher des objets. De plus, différentes matières peuvent également tomber par accident dans les cuves. Ces deux situations sont représentées dans notre cas par les deux scénarios étudiés : quand un **sac en plastique** et une quantité variable d'**objets flottants** apparaissent sur la surface du liquide. Ce scénario peut produire des dommages dans le système comme le blocage d'une pompe.
- Une **mesure bloquée** peut empêcher le capteur d'obtenir une valeur de volume correcte. Dans le cas d'étude, une feuille de papier bloque le capteur à ultrason.
- Comme les cuves sont remplies d'eau, une certaine **humidité** peut s'accumuler sur la surface du capteur. Cela peut produire une altération des mesures ainsi qu'une réduction de la durée de vie du capteur.
- Plusieurs situations peuvent provoquer des **pannes sur les flotteurs** par exemple, des objets qui bloquent leur fonctionnement et favorisent l'oxydation du mécanisme.
- Une des cyberattaques les plus utilisées est l' **attaque par déni de service ou *DoS*** qui peut bloquer complètement un réseau informatique. Le cas où un ordinateur connecté au réseau de monitoring attaque le PLC est étudié.
- Une autre cyberattaque classique est le ***spoofing***. Dans ce cas, l'attaquant modifie les données envoyées sur le réseau avec l'objectif de nuire au système. De cette façon il pourrait prendre le contrôle de la plate-forme.
- Une panne ou sabotage commun sont les **mauvaises connexions** à cause d'un câble qui a été déconnecté ou abîmé.
- Une attaque physique pourrait être réalisée par une personne qui essaye d'abîmer les cuves pour produire des fuites. Le cas où une personne donne des **coups sur les cuves** avec différentes intensités est étudié.

Pour mettre en perspective la Figure III.3 où un fonctionnement normal est représenté, deux exemples de scénarios analysés sont montrés par les Figures III.4 et III.5. Dans la

Figure III.4, nous pouvons observer qu'une feuille de papier qui bloque la mesure du capteur à ultrason fait que le signal enregistré soit plus lisse que la normale. Par contre, dans la Figure III.5, nous voyons comment un groupe d'objets qui flottent sur l'eau produisent des perturbations plus importantes comparées à un fonctionnement normal (Figure III.3).

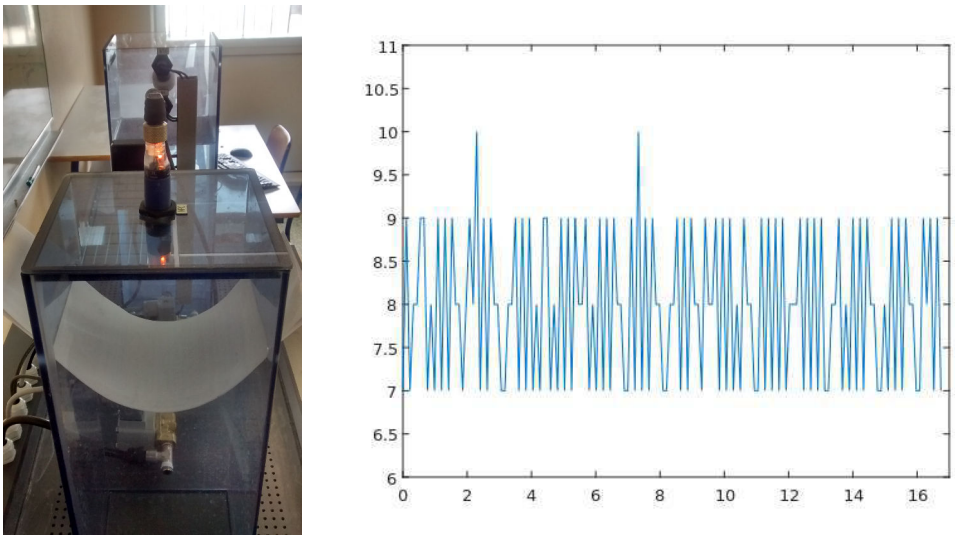

FIGURE III.4: Photo du scénario de la mesure du capteur à ultrason bloquée et son impact sur le signal.

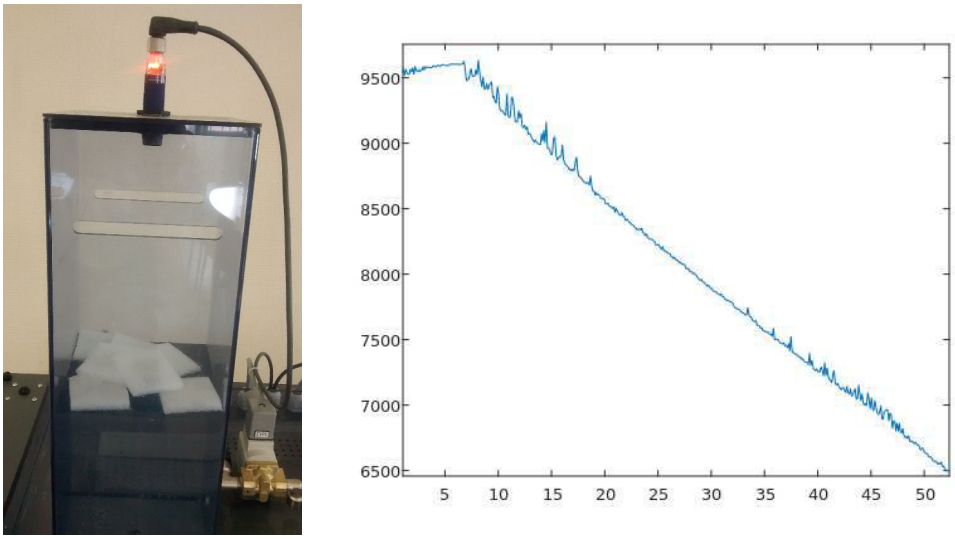

FIGURE III.5: Photo du scénario de 7 objets flottants introduits dans la cuve principale et son impact sur le signal.

III.2.2 Mesure de la qualité

Chaque composant de la plate-forme est identifié comme un module élémentaire. Tous ces modules peuvent interagir entre eux grâce au PLC qui les interconnecte. Le PLC peut se comporter en même temps comme un autre module élémentaire quand il produit un flux de données. Ce comportement apparaît en réaction à une situation déterminée par exemple le déclenchement d'une alarme.

Il n'y a qu'un ensemble d'imperfections, de dimensions, d'aspects et de facteurs qui peuvent être évalués pour chacun de ces modules. Cette limitation peut être causée par plusieurs raisons, par exemple un manque d'information sur le module, parce que la mesure n'est pas accessible ou parce qu'elle ne peut pas être créée automatiquement par une machine à cause de sa subjectivité. Ainsi, la mesure de la qualité du flux de données et des informations déduites du capteur à ultrason est montrée comme exemple. La connaissance sur l'autonomie créée à partir de ce flux et l'intelligence nécessaire pour la construire sont ainsi analysées.

Dans ce cas d'étude, il y a des dimensions statiques puisqu'elles ne sont calculées qu'une fois pour le système, comme la réputation. La mesure de la réputation est une mesure subjective basée sur l'expérience des autres modules avec le sous-système. Ces mesures statiques sont intéressantes pour l'aide à la décision ou pour fixer des AL plus ou moins restrictifs. Cependant, dans le cadre de la détection d'anomalies, seules les mesures dynamiques ont un intérêt. Ainsi, certaines mesures ont été écartées parce qu'elles n'étaient pas évaluables. Par exemple, de par l'impossibilité de connaître la valeur réelle, l'exactitude n'a pas été conservée, car seule la mesure donnée par le capteur est connue.

Ainsi, un ensemble de dimensions ont été identifiées comme intéressantes, soit pour une application d'aide à la décision soit pour la détection d'anomalies. L'intérêt des mesures d'aide à la décision a été identifié en fonction de l'application à notre cas d'étude, le calcul de l'autonomie. Pour la détection d'anomalies, les dimensions ont été choisies soit parce que les scénarios étudiés vont les impacter soit parce qu'elles peuvent être utiles pour définir les seuils des AL. Conformément aux définitions de la section II.3, le résultat de ces mesures est donné en forme de vecteurs. Une description de tous ces éléments ainsi qu'une explication de la manière dont ils sont mesurés sont présentées dans la suite. Un exemple de mesure réalisée à un instant de temps accompagne l'explication.

L'évaluation de la qualité des données est réalisée de façon similaire pour tous les sous-systèmes installés dans les cuves puisqu'ils partagent des caractéristiques communes comme

les standards de communication. Seulement deux imperfections sont mesurées. Les deux autres imperfections existantes ne sont pas pertinentes, car dans les sous-systèmes étudiés ni les données incertaines ni les données imprécises n'existent. Les imperfections étudiées sont :

— Les données **erronées** sont détectées et rejetées par le protocole de communication grâce à un code CRC (Contrôle de Redondance Cyclique) qui accompagne les données. Les statistiques du protocole Modbus [MOD96] fournissent directement la valeur de données erronées qui est traduite par une valeur « *true/false* » pour les données erronées et les données correctes, respectivement.
— L'**incomplétude** des données est corrélée avec la quantité de paquets qui n'a pas été reçue. Les statistiques de Modbus vont également fournir directement le résultat de cette mesure.

L'évaluation pour une donnée particulière est présentée comme exemple :

$$D\vec{Q}V = \{i_{err} = false, i_{inc} = 0\, messages\}; \tag{III.1}$$

Dans ce cas, le protocole Modbus n'a pas détecté de données erronées ($i_{err} = false$) ni de paquets IP perdus ($i_{inc} = 0\, messages$).

Pour le sous-système du capteur à ultrason, sept dimensions ont été évaluées : 3 dimensions intrinsèques, 2 contextuelles et 2 extrinsèques. Chacune d'elles est décrite ci-dessous. Les dimensions qui ne sont pas présentées n'ont pas montré un intérêt particulier après avoir été examinées. Ainsi, certaines dimensions ont été écartées soit parce qu'elles ne pouvaient pas être évaluées, soit à cause de leur subjectivité qui ne permet pas d'automatiser leur mesure avec une machine. La crédibilité est un exemple de mesure subjective qui ne peut être réalisée que par un humain.

— Dimensions intrinsèques :
  — La **précision de la source** est l'erreur occasionnée par le fonctionnement du capteur. Elle est donnée par son *datasheet* et est variable en fonction de la distance mesurée. Cette valeur est calculée directement avec l'information fournie par le fabricant et la distance mesurée par le capteur jusqu'à la surface du liquide. Cette précision est mesurée avec l'unité utilisée dans le système, nommée u [2]. Comme nous le verrons dans la suite, la distribution de cette mesure est gaussienne avec une moyenne de zéro. Afin que sa représentation soit plus claire,

2. 1u = 10^{-4} parts d'une cuve pleine.

nous avons créé une fonction qui prend la valeur absolue de la précision réelle et qui est filtrée avec un filtre du type FIR (*Finite Impulse Response*).

- L'**intégrité** du sous-système est déterminée à partir des observations faites sur le système. Le résultat est un pourcentage qui indique à quel point le sous-système fonctionne correctement. Chaque partie du sous-système représente une pondération sur ce pourcentage en fonction de son importance dans le sous-système. Par exemple, une LED abîmée peut faire descendre cette valeur à 10%.
- L'**unicité** a une valeur « *true* » quand l'information est unique et une valeur « *false* » quand la même information est répétée plusieurs fois.

— Dimensions contextuelles :

- La **précision réelle** est définie comme la différence entre le signal mesuré et le signal filtré. Cela permet de séparer le bruit qui est produit par les vagues et autres perturbations plus rapides qu'un changement de volume. Comme la précision de la source, elle est mesurée en unités de la cuve (u).
- L'information est considérée comme **erronée**, et donc indiquée avec une valeur « *true* » dans cette dimension, quand elle est en dehors de l'intervalle de valeurs possibles de volume, entre 0 et 10000. Dans le cas contraire, cette dimension a la valeur « $false$ ».
- L'**opportunisme** est défini comme la différence de temps entre l'arrivée de deux unités d'information. Elle est mesurée en secondes.

— Dimensions extrinsèques :

- Le **format** est une valeur « *true/false* » qui indique si l'information respecte ou pas le format attendu, respectivement.
- La **cohérence** est définie par la différence entre la mesure attendue et la valeur mesurée en u. La mesure attendue est calculée à partir d'une valeur mesurée précédemment, fixée tous les 10 échantillons et le volume de liquide que théoriquement la pompe déplace par fraction de temps. Cette valeur précédente est mise à jour avec une fréquence déterminée.
- La **confiance** dont la valeur maximale de 100 se dégrade avec la détection d'un dysfonctionnement dans le sous-système peut ensuite augmenter si le fonctionnement de celui-ci s'améliore. Les valeurs de dégradation se fixent en fonction de la faille détectée et de son impact.

Par exemple, une unité d'information peut avoir un résultat d'évaluation comme le suivant :

$$I\vec{Q}V = \left\{ \begin{array}{ll} \text{intrinsèques} & \{id_{sp} = 24u., id_{itg} = 100\%, id_{uni} = true\}, \\ \text{contextuelles} & \{cd_{rp} = 5u., cd_{err} = false, cd_{tim} = 11ms\}, \\ \text{extrinsèques} & \{ed_{for} = true, ed_{coh} = 2u., ed_{con} = 100\} \end{array} \right\} \quad \text{(III.2)}$$

Pour les dimensions intrinsèques, la précision de la source est de $24u$ ($id_{sp} = 24\,u$) calculée avec l'information du *datasheet* du capteur, le système fonctionne au complet ($id_{itg} = 100\,\%$) et les informations sont uniques à l'instant observé ($id_{uni} = true$). Les dimensions contextuelles indiquent que la précision réelle est de $5u$ ($cd_{rp} = 5\,u$), que la valeur mesurée est possible pour la cuve et qu'aucune information n'a été transmise depuis 11ms ($cd_{tim} = 11ms$). Les dimensions extrinsèques indiquent que l'information a le format attendu ($ed_{for} = true$), qu'il y a une différence de $2u$ par rapport à la mesure espérée ($ed_{coh} = 2u$) et que le système de détection d'anomalies n'a rien détecté dernièrement ($ed_{con} = 100$).

Dans ce cas, la connaissance de l'autonomie est créée par les flux d'information des deux sous-systèmes qui mesurent le volume contenu dans les deux cuves. Le résultat de l'évaluation de sa qualité est formé par les vecteurs $I\vec{Q}V$ des informations utilisées et par deux autres facteurs décrits ci-dessous :

— Le **coût d'erreur**, calculé à partir du risque assumé avec l'utilisation de la connaissance définie. Ce risque est représenté par une valeur entre 1 et 5.
— Une valeur entre 1 et 5 évalue l'**importance** de la connaissance pour le système au moment de sa définition.

Le calcul de ces deux facteurs se fait à partir du contexte d'utilisation. Un exemple de résultat de cette évaluation pour une navigation en haute mer est le suivant :

$$K\vec{Q}V = \{I\vec{Q}V_{cuve_principale}, I\vec{Q}V_{cuve_secondaire}, W\vec{Q}V, f_{err} = 5, f_{rel} = 2\}; \quad \text{(III.3)}$$

Comme nous pouvons l'apprécier dans l'Équation III.3, le résultat de la qualité de la connaissance est composé par la qualité de l'information et de l'intelligence à partir de laquelle elle a été définie. Deux facteurs complètent ce résultat. Ces deux éléments nous indiquent qu'il y a une importance de niveau 2 ($f_{rel} = 2$) établie à partir de la mission en cours et un coût d'erreur de niveau 5 ($f_{err} = 5$) parce qu'une erreur occasionnée par un faux

calcul de l'autonomie peut laisser le bateau à la dérive.

L'intelligence qui transforme le volume des cuves en autonomie est un algorithme qui utilise une fonction basée sur la consommation du navire. Deux aspects définissent sa qualité :

— L'**expérience**, donnée par la quantité d'expériences qui ont été utilisées pour la définition des concepts de l'intelligence.
— La **complétude**, interprétée comme la quantité de variables qui sont utilisées pour la définition de la connaissance.

$$W\vec{Q}V = \{a_{exp} = 38, a_{com} = 2\} \tag{III.4}$$

La qualité de l'intelligence nous indique qu'elle a été définie avec 38 expériences ($a_{exp} = 38$) seulement en considérant les niveaux des cuves comme entrées ($a_{com} = 2$). Un cas plus complet serait l'inclusion de la vitesse du navire et l'état de la mer comme variables de l'autonomie.

Toutes ces mesures peuvent évoluer dans le temps. Par exemple, nous évaluons la qualité de l'information pour chaque unité d'information reçue. Cette évolution peut être représentée sous la forme de séries temporelles. Dans la Figure III.6 l'évolution de plusieurs composantes du vecteur $I\vec{Q}V$ du capteur à ultrason est montrée comme exemple.

Les résultats obtenus dans cette section évaluent la qualité des quatre niveaux impliqués dans la définition d'une connaissance. Le maximum d'éléments a été évalué même si, comme cela a été décrit, certains éléments de la qualité n'ont pas pu être évalués. Les quatre vecteurs résultants vont servir comme base pour la détection d'anomalies.

Par la suite, les scénarios présentés dans la sous-section III.2.1 seront analysés en détail. Ces situations variées vont donner différents résultats de qualité qui seront étudiés de manière exhaustive.

III.2.3 Détection d'anomalies (actes malveillants)

Pour la détection d'anomalies, nous prenons comme point de départ les mesures réalisées dans la section précédente. Seules les mesures objectives sont d'intérêt, car elles peuvent être réalisées par un algorithme sans recours à un expert. D'autre part, les mesures subjectives pourront affecter l'application des AL en les rendant plus restrictifs.

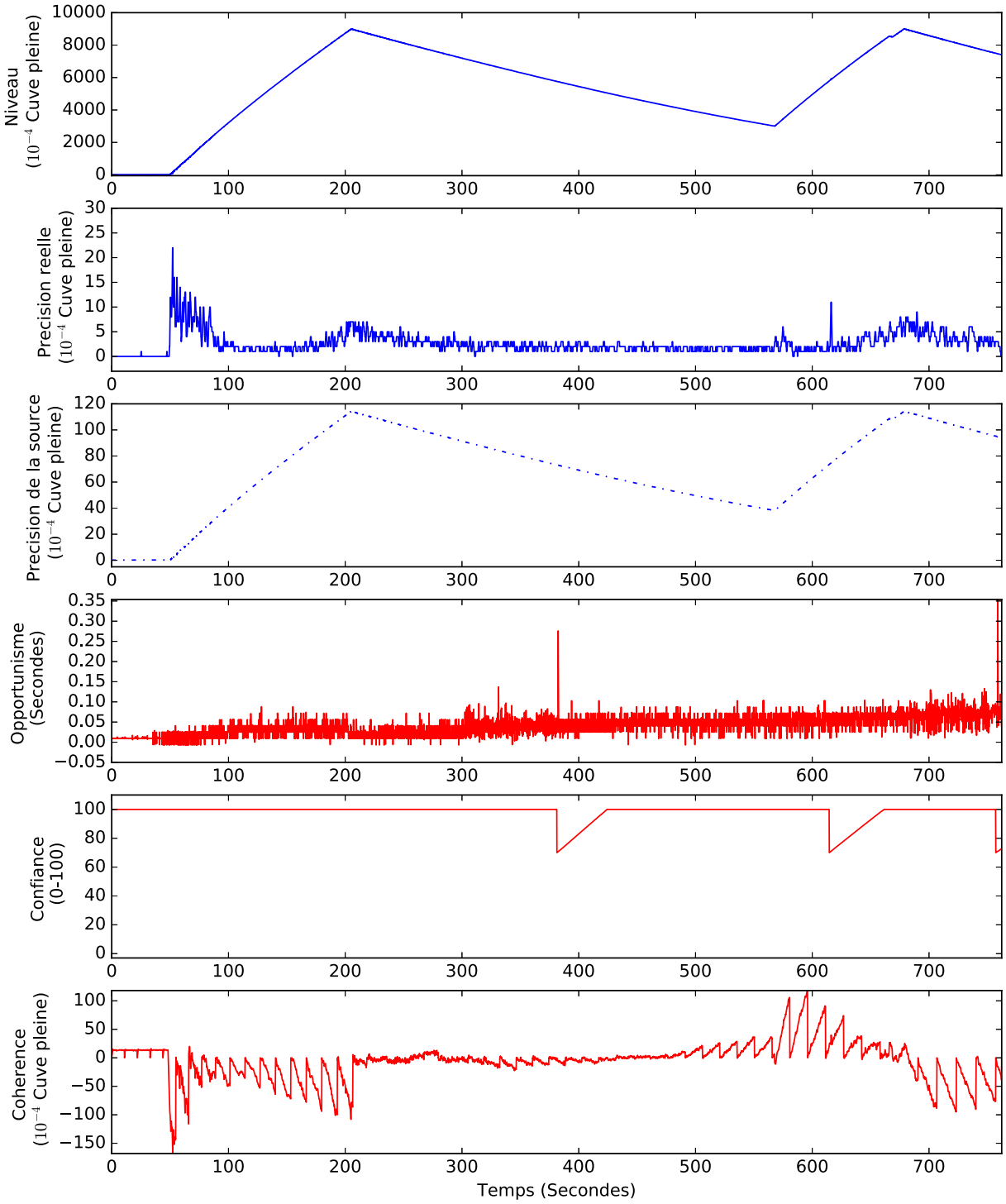

FIGURE III.6: Évolution dans le temps de la mesure réalisée par le capteur à ultrason et de ses mesures de qualité respectives (précision réelle, précision de la source, opportunisme, confiance et cohérence).

Lorsque l'effet des objets flottants sur les mesures de qualité est examiné, la précision réelle est une des dimensions fortement affectées. L'impact d'une quantité variable d'objets est représenté dans la Figure III.7.

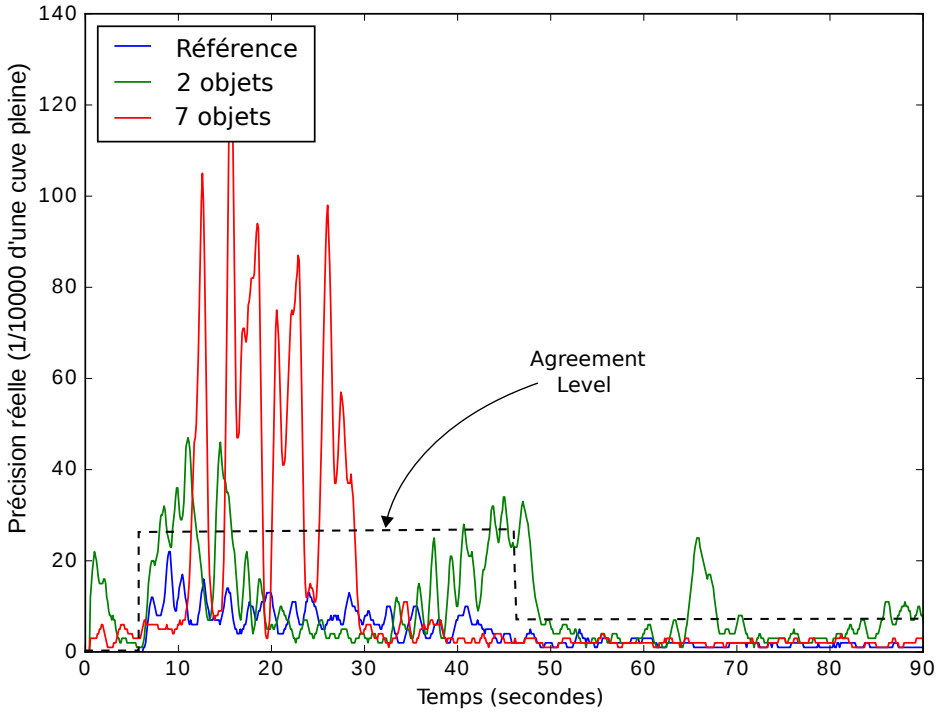

FIGURE III.7: Effet des objets flottants sur la précision réelle du capteur à ultrason et représentation de l'AL.

Dans les trois cas - référence, 2 et 7 objets - nous observons, quand la cuve commence à se remplir (entre les secondes 5 et 45), comment le bruit dans les trois cas est supérieur aux valeurs mesurées une fois que la hauteur de la pompe est franchie (vers la seconde 45). Ceci permet de fixer deux niveaux de signal afin de différencier le cas de normalité des scénarios anomaux. Ainsi, quand la cuve est vide pour le cas de référence, la mesure est quasiment nulle tandis que pour les autres scénarios ce n'est pas toujours le cas, conduisant à fixer un troisième niveau de différentiation.

Pour le calcul des ALs, nous avons observé que la précision réelle présentait une distribution gaussienne dans les trois intervalles définis - $x = 0$, $0 < x < 3000$ et $x > 3000$ - où x représente la valeur mesurée. Dans la Figure III.8, l'histogramme du troisième intervalle est montré comme exemple. Cette représentation nous guide vers les systèmes de détection adaptés aux systèmes qui génèrent des fonctions de densité de probabilité de type gaussienne (présentées dans la Section I.2.2).

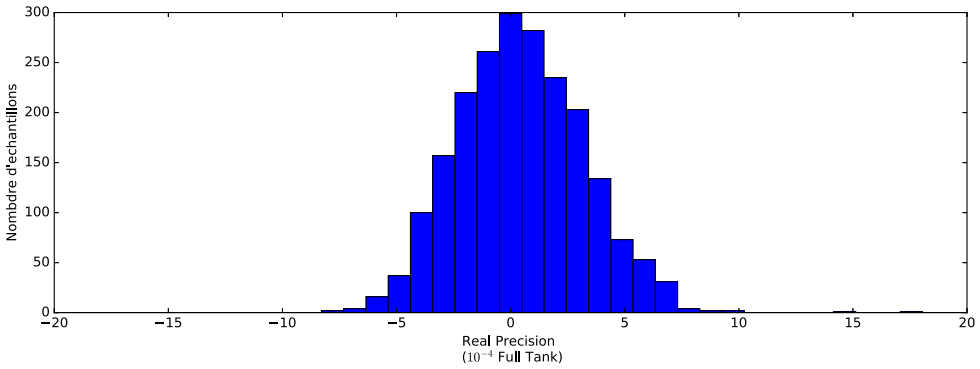

FIGURE III.8: Histogramme de la précision réelle lorsque la valeur mesurée est plus grande que $3000u$.

Ainsi, après avoir construit les diagrammes de ces 3 cas, nous avons utilisé la règle 68-95-99.7 pour définir les ALs de chaque intervalle avec un seuil de 3σ. Au moyen de ces trois valeurs, il est possible de définir complètement un AL (Équation III.5) pour la précision réelle, composé par trois valeurs qui changeront en fonction du volume de liquide présent dans la cuve. Cet AL a été ainsi représenté sur la Figure III.7.

$$AL_1(x) \equiv \begin{cases} D_{rp} < 1 & \text{quand } x = 0 \\ D_{rp} < 25 & \text{quand } 0 < x < 3000 \\ D_{rp} < 8 & \text{quand } 3000 < x \end{cases} \tag{III.5}$$

Par contre, lorsque la mesure du capteur à ultrason est bloquée, la mesure de la précision réelle est affectée, ayant des valeurs proches de zéro pendant une longue durée (Figure III.9). Cela permet à un AL de valeur 0.5 dans la précision réelle représentée de détecter ce comportement (Équation III.6). Cette limite est ainsi directement applicable à la mesure de la précision réelle originale puisque même si la mesure zéro - très habituelle - est considérée comme anormale, cela affecte seulement la confiance. Donc, seule une mesure de zéro en continu fera déclencher la détection de l'anomalie. Cet AL est limité aux mesures qui sont réalisées quand la cuve n'est pas vide, car le comportement observé est très similaire.

$$AL_2(x) \equiv D_{rp} > 0.5 \text{ quand } x > 0 \tag{III.6}$$

L'opportunisme est également affecté par un scénario de cyberattaque de déni de service (DoS). Pour cette dimension, deux ALs bornent l'état normal de l'opportunisme (Équations III.7 et III.8). Dans la Figure III.10, une première perturbation est détectée quand l'attaque

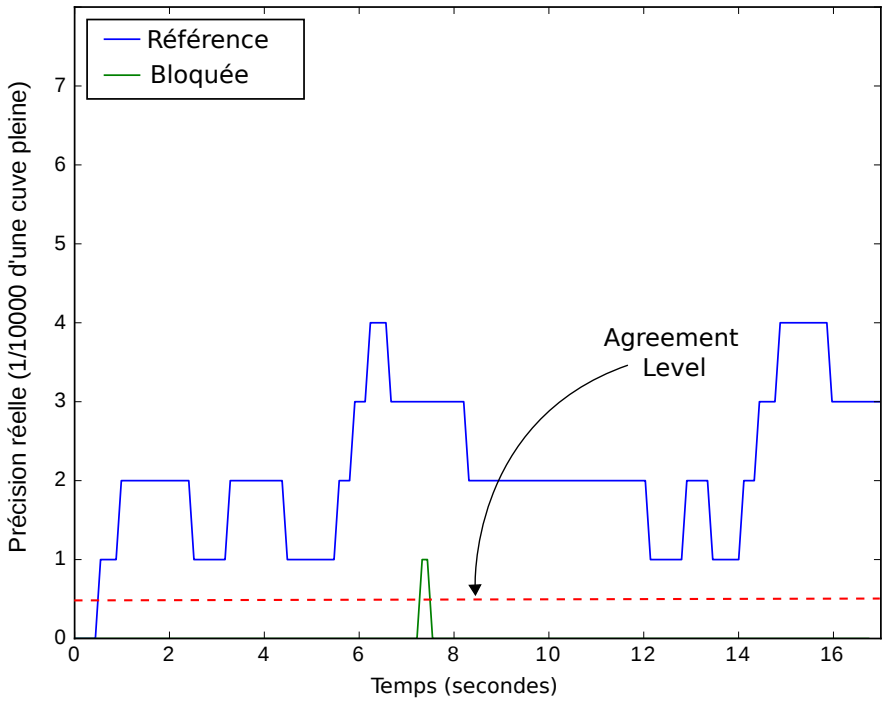

FIGURE III.9: Effet du capteur à ultrason bloqué sur la précision réelle et représentation de l'AL.

est déclenchée (vers l'échantillon 90). Lorsque l'attaque réussit à bloquer la transmission des données (vers l'échantillon 160) l'opportunisme devient indéterminé (arrêt indéfini du CPS), tant que l'attaque n'est pas contrôlée ou stoppée. Une fois que l'attaque est terminée (à partir de l'échantillon 165), un temps de récupération est nécessaire avant que l'opportunisme revienne à un état normal. Les méthodologies de détection d'anomalies basées sur des histogrammes ont aidé à fixer les valeurs des ALs dans un premier temps. Ensuite, l'expérience mise en place a démontré qu'étendre l'intervalle entre les deux ALs réduit les fausses détections sans dégrader la détection de l'attaque. Finalement les valeurs choisies pour ces ALs sont les suivantes :

$$AL_3 \equiv D_{tim} > 0.105 \tag{III.7}$$

$$AL_4 \equiv D_{tim} < 0.11 \tag{III.8}$$

Toutes ces détections peuvent fréquemment déclencher de fausses alarmes à cause de

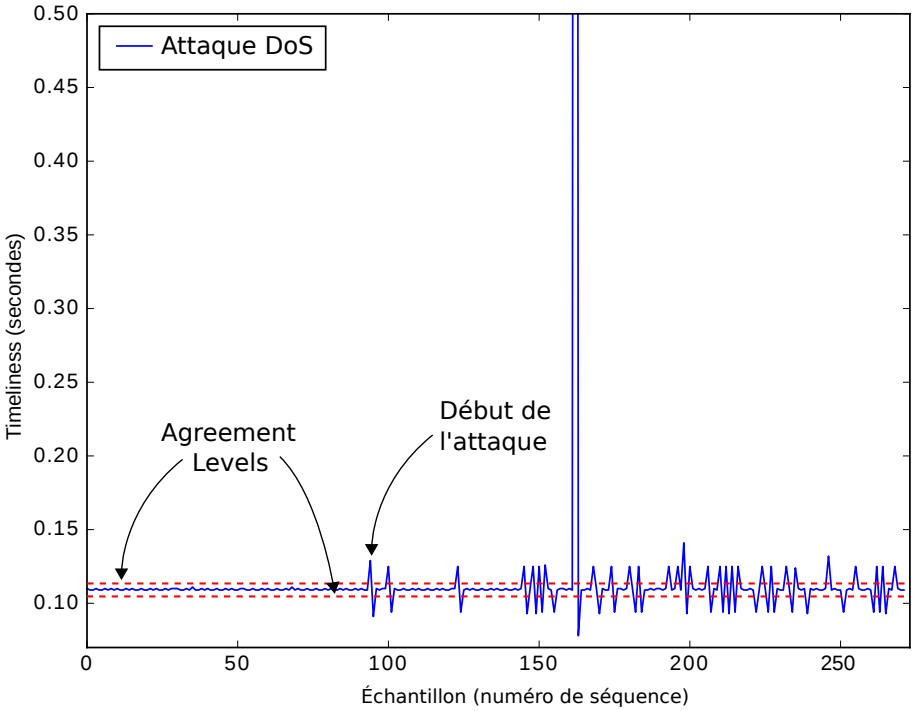

FIGURE III.10: Effet d'une attaque DoS sur l'opportunisme et représentation des ALs.

valeurs anormales isolées. Pour éviter ce problème, la dimension confiance se dégrade à chaque détection. De cette façon, elle fonctionne comme une mémoire qui se souvient des anomalies pendant un certain temps, ce qui permet d'identifier une suite de valeurs anormales et réduire les fausses alarmes.

Les alarmes permettent d'indiquer que le système est entré dans un fonctionnement anomal. Cependant, l'identification du scénario subi à partir des mesures de la qualité peut devenir compliquée . L'équipage du navire n'est pas familiarisé avec l'approche et donc, il faudra traduire la détection à son langage technique. Une étape de catégorisation est réalisée après chaque détection avec ce but.

III.2.4 Catégorisation d'anomalies (actes malveillants)

Afin de catégoriser les détections, nous avons compilé toutes les imperfections et dimensions impactées dans quelques scénarios typiques identifiés pour chaque sous-système. Au total, 17 scénarios illustrent cette démarche. Pour le capteur à ultrason 10 scénarios ont été étudiés (Table III.1), 5 pour les flotteurs (Table III.2) et 2 scénarios d'attaques réseau

(Table III.3).

Dans ces tableaux, plusieurs détails sont représentés. L'origine de l'anomalie est indiquée avec « M » pour une anomalie produite dans la mesure, avec « S » quand elle est produite par le capteur et « N » quand l'origine est le réseau. Le facteur qui peut générer l'anomalie est représenté à partir d'une nomenclature de catégorisation créée avec cet objectif et qui est affichée dans la Table III.4.

En pensant toujours à l'application navale, le coût d'erreur qu'une décision prise à partir de ces informations anomales peut avoir est expliqué. Finalement, les dimensions impactées pour chaque scénario sont marquées avec « x » et avec « (x) » quand l'impact est seulement potentiel. La nomenclature utilisée est celle introduite dans la Section II.3.

Ces trois tableaux permettent d'observer que chaque scénario impacte différentes mesures de qualité. Ainsi, la catégorisation des anomalies peut parfaitement identifier la source de l'anomalie ou bien réduire les possibilités. Dans certains cas, la mesure de qualité peut être affectée de différentes manières. Par exemple, les cas des objets flottants et de la mesure bloquée du capteur à ultrason produisent une précision réelle trop mauvaise et trop bonne respectivement par rapport aux AL définis.

III.3 Drones aériens

Les systèmes navals s'intéressent de plus en plus aux drones aériens pour la réalisation de différentes tâches comme la reconnaissance de cibles, la surveillance ou le transport de petits objets. Leur prix réduit et leur déploiement facile font que les hélicoptères utilisés jusqu'à nos jours sont remplacés au fur et à mesure par ces systèmes pour certaines missions [3].

La sécurisation de ces systèmes est un enjeu majeur dû à l'importance et criticité de leurs tâches. Les drones aériens sont contrôlés par un CPS qui prend différentes mesures comme la position et l'altitude et qui lui permet de manœuvrer en conséquence. La connexion à distance et l'utilisation de systèmes standards font des drones une cible des cyberattaques, en plus du risque d'autres anomalies qui peuvent nuire au système.

3. DCNS et Airbus Helicopters préparent un drone naval ultra-digital. IT Industrie & Technologies. Consulté le 08 Août 2017. https://www.industrie-techno.com/dcns-et-airbus-helicopters-prepare nt-un-drone-naval-ultra-digital.46563

TABLE III.1: Catégorisation des anomalies identifiées pour le capteur à ultra-son à partir des éléments [a] impactés.

Scénario	**Type : Origine**	**Type : Facteur**	**Facilité de détection**	**Coût d'erreur**	I_{err}	I_{inc}	id_{sp}	id_{uni}	id_{tru}	cd_{err}	cd_{tim}	cd_{rp}	ed_{coh}
Environnemental	M-S	1.b	Elle dépend (de l'effet)	Il dépend (de l'effet)			x		x			x	
Humidité	M-S	1.b	Difficile	Durée de vie réduite			x						
Mauvaise connexion	N	1.b/2.a	Facile	Perte d'information	x						x		
Mauvaise calibration	N	1.a/2.a	Elle dépend (de la mesure)	Décisions basées sur mesures fausses						x			
Capteur endommagé	S	1	Elle dépend (des dégâts)	Perte d'information - Mesure fausse	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)
Mesure bloquée	M	2.a	Facile	Perte d'information - Mesure fausse			x					x	
Replacement de liquide	M	2.a	Elle dépend (du liquide)	Navire à la dérive - Composants endommagés			x					x	
Objets flottants	M	1.b/2.a	Elle dépend (de l'objet)	Lectures erronées - Composants endommagés			x					x	
Coups sur les cuves	M	2.a	Facile	Cuves endommagées						x		x	x
Fuite	M	1/2.a	Elle dépend (de la taille)	Navire à la dérive et risque d'incendie								x	x

a. Définition des acronymes dans la Section II.3

TABLE III.2: Catégorisation des anomalies identifiées pour les flotteurs à partir des éléments [a] impactés.

Scénario	**Type : Origine**	**Type : Facteur**	**Facilité de détection**	**Coût d'erreur**	I_{err}	I_{inc}	id_{sp}	id_{uni}	id_{tru}	cd_{err}	cd_{tim}	cd_{rp}	ed_{coh}
Mauvaise connexion	N	1/2.a	Facile	Perte d'information	x						x		
Capteur endommagé	S	1.a/2.a	Elle dépend (des dégâts)	Perte d'information - Mesure fausse	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)
Capteurs bloqués	M	2.a	Elle dépend (de la mesure)	Décisions basées sur mesures fausses						x			x
Coups sur les cuves	M	2.a	Elle dépend (de l'intensité)	Cuves endommagées						x		x	x
Fuite	M	1/2.a	Elle dépend (de la taille)	Navire à la dérive et risque d'incendie								x	x

TABLE III.3: Catégorisation des anomalies réseau identifiées à partir des éléments [a] impactés.

Scénario	**Type : Origine**	**Type : Facteur**	**Facilité de détection**	**Coût d'erreur**	I_{err}	I_{inc}	id_{sp}	id_{uni}	id_{tru}	cd_{err}	cd_{tim}	cd_{rp}	ed_{coh}
Attaque de déni de service	N	2.b	Facile	Perte d'information		x					x		
Spoofing	N	2.b	Elle dépend (de la sophistication)	Décisions basées sur mesures fausses	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)

a. Définition des acronymes dans la Section II.3

TABLE III.4: Catégorisation de types d'anomalies en fonction du facteur d'origine.

ID	Description
1	Détérioration naturelle
1.a	Défaut / Endommagement
1.b	Facteur externe
2	Détérioration causée
2.a	Sabotage (accès physique)
2.b	Attaque externe

III.3.1 Description

Pour la réalisation du cas d'étude, deux drones ont été choisis (Figure III.11) : un « *parrot rolling spider* » et un « Crazyflie 2.0 »[4].

FIGURE III.11: « *Parrot rolling spider* » (à gauche) et « Crazyflie 2.0 » (à droite)

Du fait des différents buts de chaque produit, les drones ont certaines particularités. Le Parrot est conçu pour être ludique et possède une assistance au pilotage. À cet égard, il a des fonctionnalités comme le décollage/atterrissage automatique. La quantité de capteurs embarqués et ses autres possibilités ont motivé son utilisation dans l'enseignement et la recherche[5]. La récupération des logs de navigation des vols réalisés est une option qui apparaît par défaut. Ces fichiers permettent de réaliser une analyse *a posteriori* des données produites par les capteurs. Les logs sont au format JSON (JavaScript Object Notation) dont les variables qui font objet de ce travail sont affichées ci-dessous.

$$vol_{parrot}\{temps, batterie, \text{état}, alerte, v_x, v_y, v_z, \phi, \theta, \psi, altitude, vitesse\} \quad \text{(III.9)}$$

4. Crazyflie 2.0. Bitcraze. Consulté le 13 Juillet 2017. `https://www.bitcraze.io/crazyflie-2/`

5. Feedback Control Systems. An MIT Feedback Control Systems Class that Teaches with Palm-size Drones. MIT. Consulté le 13 Juillet 2017. `http://fast.scripts.mit.edu/dronecontrol/`

Dans ce vecteur $\{v_x, v_y, v_z\}$ représentent la vitesse décomposée dans les trois axes, $\{\phi, \theta, \psi\}$ sont les angles mesurés par le gyroscope et l'état est une valeur entre 0 et 4 qui indique le mode de fonctionnement (1 = décollage, 2-3 = en vol et 4 = atterrissage). La Figure III.12 montre les données récupérées pendant un vol d'exemple. Au cours de ce vol, aucune alerte n'a été déclenchée et n'est donc pas représentée.

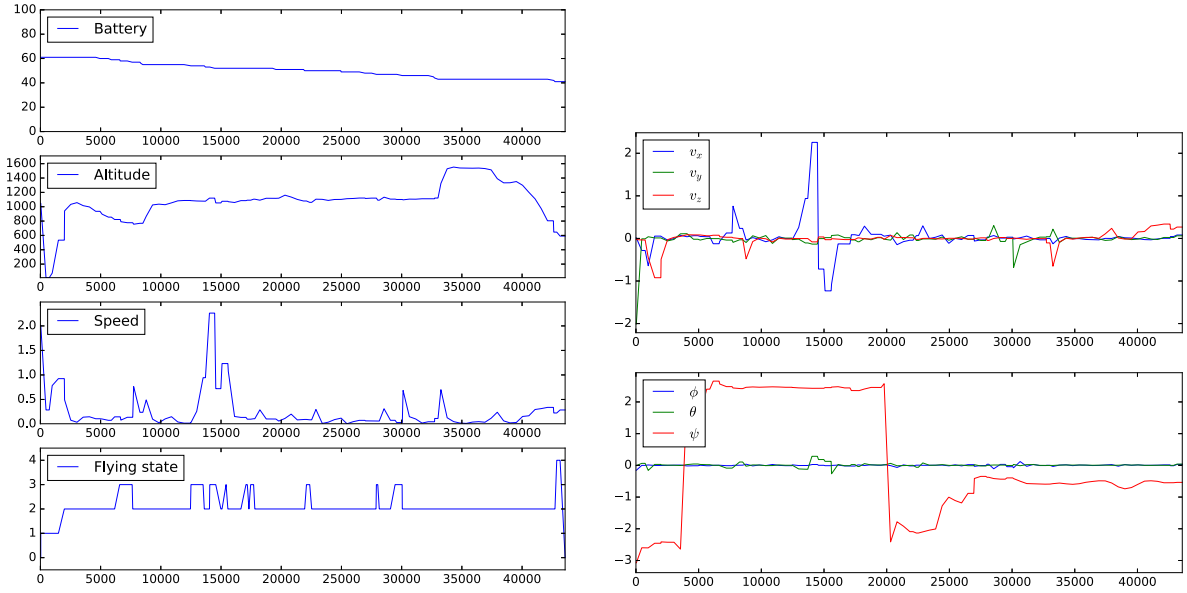

FIGURE III.12: Exemple de séries temporelles des mesures produites par les capteurs du drone Parrot.

Contrairement au drone Parrot, le Crazyflie ne possède pas d'aide au pilotage. Il s'agit d'un drone conçu par une équipe de recherche avec pour but l'adaptabilité à différents scénarios afin de servir de plate-forme de test. Un système de broches permet l'installation de modules d'expansion avec par exemple des lumières, un GPS ou un capteur laser pour mesurer l'altitude. Ce dernier, appelé « Z-ranger deck », a été utilisé afin de faciliter le pilotage pendant la création des scénarios.

L'interface graphique pour le pilotage du Crazyflie inclut l'activation de logs de différents paramètres. Entre les paramètres étudiés, nous trouvons le niveau de batterie, les angles du gyroscope (*roll*, *pitch*, *yaw*), le baromètre et la puissance de chaque moteur (m_i). Ces valeurs sont enregistrées dans des fichiers CSV (*Comma-separated Values*). Ainsi, le vol d'un Crazyflie peut être représenté de la façon suivante :

$$vol_{crazyflie}\{temps, batterie, m_1, m_2, m_3, m_4, \text{poussée}, v_x, v_y, v_z, \phi, \theta, \psi, \text{baromètre}\} \quad \text{(III.10)}$$

Anomalies, sabotages et dysfonctionnements étudiés

Les drones aériens sont vulnérables à différents dysfonctionnements, sabotages et cyberattaques. Un travail d'identification des risques a été réalisé afin de les étudier en détail. Les scénarios identifiés sont les suivants :

- La **prise de contrôle** du drone par un attaquant présente un risque majeur. La détection de ce type d'attaque peut éviter la perte de l'appareil et donc ne pas compromettre la mission effectuée.
- La mesure d'altitude permet de réaliser une navigation en sécurité. Si la **mesure d'altitude est bloquée**, la réalisation de certaines manœuvres entraînera un risque important.
- Une caméra peut améliorer la stabilité du vol du drone ainsi que permettre la réalisation de missions de surveillance. Si l'image de la **caméra est bloquée**, la stabilité du drone peut être altérée et les missions de surveillance compromises.
- La **perte de puissance** d'un moteur peut être compensée quand son effet n'est pas important, mais le problème doit être détecté et corrigé pour éviter un endommagement du système. Ce problème peut survenir à cause d'une cyberattaque, un sabotage ou un manque de maintenance.
- L'ajout d'un **poids** sur le drone peut affecter son autonomie et sa stabilité. Ce scénario peut être causé par des objets qui ne devraient pas être transportés par le drone.
- Les hélices sont un composant vital pour les drones. De ce fait, plusieurs cyberattaques essayent de les saboter [BYTE16]. La détection des **hélices mal équilibrées** peut être un indice pour détecter une cyberattaques et ce avant qu'une faille majeure n'apparaisse.
- La **perte intermittente de connexion** avec un composant peut générer des situations de risque. Par exemple, un niveau bas de batterie peut générer ce problème. Ainsi, le contrôle à distance peut produire des déconnexions à cause du bruit dans les signaux ou une grande distance entre l'opérateur et le drone.
- Une **mauvaise calibration** des capteurs peut provoquer de fausses mesures qui peuvent entraîner des décisions incorrectes.

— Les **batteries** sont un composant particulièrement sensible qui s'abîme rapidement au cours de leur utilisation. La détection d'un comportement anormal de ce composant permet de le remplacer avant que des dégâts n'apparaissent.
— Les drones utilisent fréquemment des systèmes standards dont différentes vulnérabilités connues peuvent être exploitées. C'est pour cela qu'il faut prévoir les attaques les plus communes. Une attaque d'**usurpation d'identité ou *spoofing*** pourrait réussir à prendre le contrôle du drone en injectant de fausses données dans le système.

Tous ces scénarios ont été analysés pour les deux drones testés. La création de logs de vol pendant la réalisation de ces scénarios a permis de les étudier en détail « *off-line* ».

III.3.2 Mesure de la qualité

Chaque flux de données enregistré dans notre cas d'étude représente un sous-système, correspondant à chaque capteur installé sur les drones. Comme pour le cas précédent, seul un ensemble d'imperfections, de dimensions, d'aspects et de facteurs peut être mesuré.

L'évaluation subjective des éléments de la qualité devient difficile puisqu'elle ne peut pas être réalisée par une machine. De plus, l'ajout de l'humain dans la boucle d'évaluation ne permet pas de donner des résultats en temps réel. Cette contrainte est nécessaire pour réagir pendant les vols. Éventuellement, quelques résultats peuvent être améliorés avec l'ajout de certaines dimensions même si elles ne peuvent pas être utilisées dans notre cas d'étude. Par exemple, la crédibilité des valeurs mesurées peut être évaluée par l'expérience du pilote, mais son évaluation ne respecterait pas la contrainte du temps réel.

L'étude des données se fait *a posteriori* avec les fichiers de logs. Cela facilite leur traitement et permet de travailler « off-line » avec un ensemble de scénarios définis. Leur évaluation est faite afin de pouvoir l'appliquer à un flux de données reçu en continu. Pour que l'évaluation puisse être réalisée en temps réel, plusieurs contraintes doivent être satisfaites comme l'utilisation d'un système d'exploitation qui fonctionne en temps réel (par exemple Linux-RT) et le développement d'algorithmes qui s'adaptent à ce besoin.

Certaines mesures seront calculées au démarrage du drone et resteront statiques pendant le vol tandis que d'autres mesures dynamiques évolueront dans le temps. Les éléments étudiés, identifiés comme intéressants et analysés en détail, sont présentés ci-dessous :

Pour le cas d'étude des cuves, les mesures statistiques du protocole ont servi à évaluer certaines dimensions de la qualité des données. Les mesures réalisées par les protocoles utilisés par le Parrot et le Crazyflie ne sont pas accessibles pour réaliser cette évaluation. En conséquence, la mesure de la qualité des données est définie à partir des données enregistrées au lieu des données reçues.

— Les données **erronées** dans la communication sont détectées par le protocole de communication et rejetées. En conséquence, ce type de données n'apparaîtra pas dans les logs étudiés même si cette contrainte est évaluée par le sous-système de communication. Les données erronées auront un impact sur la dimension d'opportunisme de la qualité de l'information. Dans notre cas cette imperfection est représentée par une valeur « true/false » qui indique si les données contenues dans le log sont corrompues ou pas.

Un exemple d'évaluation où les données n'ont pas été identifiées comme erronées est le suivant :

$$D\vec{Q}V = \{i_{err} = false\}; \tag{III.11}$$

Pour l'évaluation de la qualité de l'information, 9 dimensions - 4 intrinsèques, 2 contextuelles et 3 extrinsèques - ont été identifiées comme intéressantes. Elles sont décrites par la suite :

— Dimensions intrinsèques :
  — La **précision de la source** est estimée comme le bruit présent dans les signaux si la mesure est connue. Pour ce cas d'étude, cette dimension peut uniquement être calculée quand les drones ne volent pas, parce que les valeurs à mesurer ne sont connues avec certitude qu'à ce moment.
  — La **confiance** est formée par deux attributs. Le premier est une valeur booléenne qui indique si le sous-système a passé le test au démarrage du drone. Le deuxième attribut est une valeur qui se dégrade avec la détection d'anomalies et qui s'améliore avec un fonctionnement correct du système. Cette valeur est plafonnée à 100.
  — L'**intégrité** est mesurée comme le pourcentage du système qui fonctionne correctement.
  — L'**unicité** est une valeur « *true/false* » qui évalue si la même information est reçue plusieurs fois dans un intervalle de temps.

— Dimensions contextuelles :
  — La **précision réelle** est mesurée en filtrant toutes les valeurs qui n'ont pas été produites par un comportement normal du système. Le fonctionnement normal permet de spécifier les fréquences du filtre.
  — L'**opportunisme** est défini comme la différence de temps entre la lecture de deux échantillons. Pour les drones la fréquence d'échantillonnage est fixée à une valeur prédéfinie.
  — Les informations **erronées** correspondent aux valeurs impossibles du sous-système. Quand elles sont identifiées, cette dimension a la valeur « *true* » au lieu d'avoir la valeur « *false* » qu'elle a par défaut.
— Dimensions extrinsèques :
  — Le **format** est une valeur booléenne qui indique si l'information enregistrée est conforme au format attendu.
  — La **redondance** indique la quantité des sources comparées au cours du processus d'extraction d'information. Pour tous les sous-systèmes analysés la redondance a une valeur de 1 sauf pour l'altitude du Crazyflie, qui peut être calculée avec le capteur laser et la mesure de pression atmosphérique.
  — La **cohérence** est mesurée comme la différence entre la valeur théorique attendue et la valeur mesurée. Pour le calcul de ces valeurs, l'interaction entre les sous-systèmes doit être analysée. Par exemple, l'ajout de puissance dans les moteurs produit un incrément des mesures de l'accéléromètre dans la direction de poussée des moteurs.

Un exemple d'évaluation de la qualité de l'information d'un échantillon de la hauteur, prise pendant un vol du Parrot est présenté dans l'Équation III.12.

$$\vec{IQV} = \left\{ \begin{array}{ll} \text{intrinsèques} & \{id_{sp} = null, id_{tru} = \{true, 95\}, id_{itg} = 98\%, id_{uni} = true\}, \\ \text{contextuelles} & \{cd_{rp} = 5mm, cd_{tim} = 520ms, cd_{err} = false\}, \\ \text{extrinsèques} & \{ed_{for} = true, ed_{red} = 1, ed_{coh} = 2u\} \end{array} \right\} \tag{III.12}$$

Dans ce vecteur IQV, nous pouvons observer que la précision de la source n'a pas pu être évaluée parce que le drone était en vol, l'altitude n'était pas fixe et en conséquence inconnue. Quant à la confiance, le drone a passé les tests réalisés au démarrage et elle a perdu 5 points par rapport au maximum à cause d'anomalies détectées dans le passé. Ces

anomalies ont fait que l'intégrité a été réduite à 98% par les doutes que les sous-systèmes ont générés. L'information enregistrée est unique ($id_{uni} = true$). Concernant les dimensions contextuelles une précision réelle de $5mm$ a été mesurée à partir du bruit filtré, la dernière information est valide pour le système et elle a été enregistrée $520ms$ avant. Les dimensions extrinsèques nous indiquent que le format a été respecté, qu'il n'y a pas de redondance dans la mesure d'altitude du parrot et la cohérence, qui est la différence avec l'altitude attendue, est de $2u$.

Pour la surveillance d'un navire, un drone doit décoller et rester statique à une hauteur déterminée. Dans notre cas d'étude, nous fixons pour le parrot cette hauteur à $1m$ au-dessus du sol. Afin de définir la connaissance sur la manière de réagir pour maintenir une altitude déterminée, deux facteurs ont été identifiés :

— Le **coût d'erreur**, dont la valeur entre 1 et 5 qui représente le risque d'utiliser la connaissance définie.
— La **pertinence**, évaluée avec une valeur entre 1 et 5 en fonction de l'importance de la connaissance pour le système au moment de sa création.

À un instant précis, nous extrayons une connaissance indiquant qu'il faut réduire la puissance des moteurs pour descendre et s'approcher de la hauteur définie. L'évaluation de la qualité de cette connaissance produit le résultat suivant :

$$K\vec{Q}V = \{I\vec{Q}V, W\vec{Q}V, f_{err} = 4, f_{rel} = 3\}; \tag{III.13}$$

Ce vecteur $K\vec{Q}V$ contient les évaluations de qualité de l'information et l'intelligence utilisées pour la définition de la connaissance. Ainsi, nous pouvons observer que réduire la puissance des moteurs a un risque assez élevé ($f_{err} = 4$), du fait que le drone peut chuter si la hauteur mesurée est incorrecte. La pertinence de cette réaction est moyenne parce que la surveillance d'un navire a une grande importance, or celle-ci n'est pas critique.

L'intelligence utilisée pour définir cette connaissance est un algorithme qui détermine la réaction à partir d'une fonction avec une hystérésis de $\pm 10cm$. Ceci évite de faire varier la vitesse des moteurs constamment. Cet algorithme est évalué lors de sa mise en place et son évaluation ne nécessite pas d'être réalisée en temps réel. En conséquence, nous pouvons exceptionnellement utiliser des mesures subjectives. Dans l'évaluation de la qualité 3 aspects sont utilisés :

— L'**expérience**, mesure subjective basée sur le travail réalisé par Parrot et Bitcraze.
— La **confidence**, évaluée subjectivement à partir de l'expérience de l'équipe de développement du drone.
— La **complétude**, quantité de variables prises en compte pour la définition de la connaissance. Le facteur de complétude est représenté en utilisant un pourcentage qui indique l'approche théorique associée au comportement réel.

L'évaluation de la qualité de l'intelligence reste statique pendant que celle-ci n'est pas modifiée. Le $W\vec{Q}V$ au moment de la définition de la connaissance montrée comme exemple est la suivante :

$$W\vec{Q}V = \{a_{exp} = 15, a_{con} = 3, a_{com} = 1\} \tag{III.14}$$

Grâce à ce vecteur, nous apercevons que l'analyse de 15 expériences a permis la modélisation du comportement souhaité avec un algorithme. Ainsi, basée sur des tests avec l'algorithme développé, une confiance de 3 sur 5 a été établie. Cet algorithme prend en considération une seule variable comme entrée ($a_{com} = 1$), la mesure de la hauteur avant de calculer une réponse. Avec la définition de ce vecteur, toutes les transformations d'information pour la génération de connaissances peuvent être analysées.

L'opportunisme est une dimension de grande importance pour l'aide à la décision et la détection d'anomalies. En fonction du système étudié, son comportement peut varier énormément. Deux évaluations des flux créés par les deux drones sont montrées comme exemple. La Figure III.13 présente l'évaluation d'un vol du Parrot dans lequel il décolle ($0 - 2100ms$), fait un survol statique ($2,1s - 17s$), réalise un circuit (17s - 94s), reste statique ($94s - 104s$) et atterrit ($104s$- fin). Nous observons que pendant que le Parrot reste statique la mesure à une valeur d'environ $500ms$, tandis que lorsque le drone manœuvre - décollage, atterrissage et déplacement - la densité d'unités d'information augmente au cours du temps. Par contre l'opportunisme du Crazyflie suit un motif différent. La Figure III.14 montre que les logs du Crazyflie enregistrent l'information avec une période parfaite de $200ms$ qui est altérée seulement lorsqu'il y a des données perdues. Si cette mesure est représentée dans un histogramme, nous voyons que la majorité des valeurs sont égales à $200ms$. Cela veut dire que normalement aucune donnée n'est perdue. Il est plus rare qu'une donnée soit perdue ($400ms$) et encore plus rare que 2 ou plus des données soient perdues ($600ms$ ou plus).

.

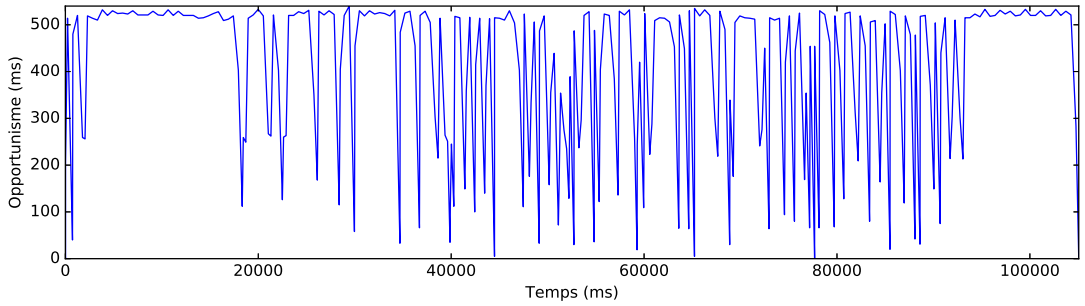

FIGURE III.13: Mesure de l'opportunisme d'un vol du parrot.

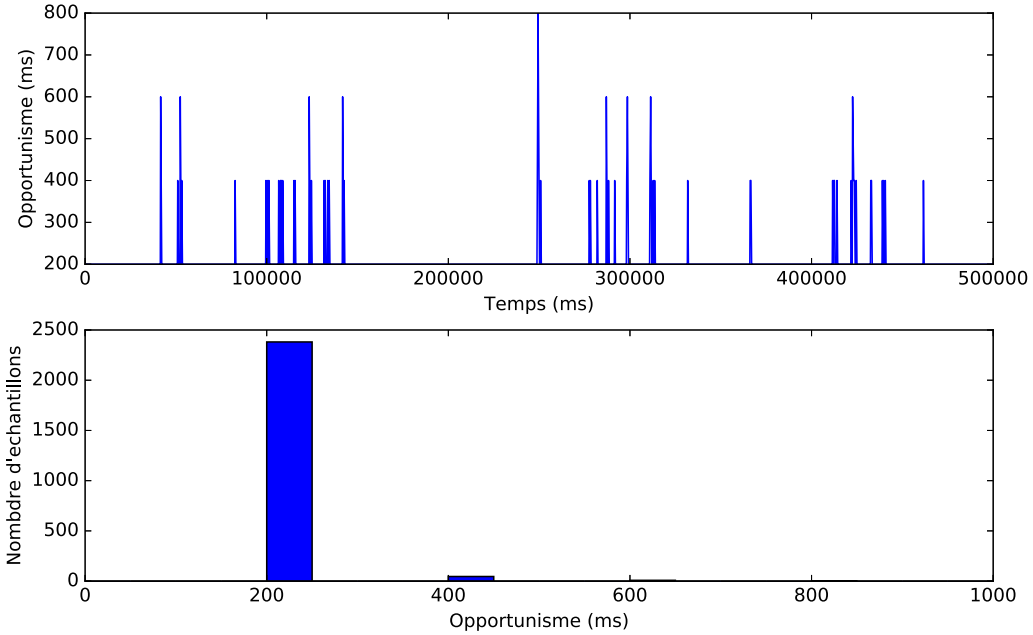

FIGURE III.14: Mesure de l'opportunisme pendant un vol du crazyflie (en haut). Histogramme de l'opportunisme mesuré pendant le vol (en bas).

Comme nous le verrons dans la sous-section suivante, la précision aura une importance majeure pour la détection des scénarios étudiés. La Figure III.15 montre deux vols avec un comportement complètement différent comme exemple. Dans le premier vol (à gauche), le drone reste à la même altitude après avoir décollé. Dans le deuxième (à droite), le drone monte et descend plusieurs fois. Cette seconde mesure permet d'apprécier que la précision réelle dans les deux cas soit un bruit avec un comportement similaire, sauf les pics qui apparaissent dans le deuxième vol à cause des changements brusques d'altitude. Grâce à l'histogramme, nous pouvons affirmer que cette dimension a une distribution gaussienne. Au fur et à mesure que le temps de vol augmente, l'histogramme s'approche de plus en plus d'une

courbe gaussienne. Nous observons ce comportement dans le deuxième cas, car le deuxième vol a une durée beaucoup plus longue que le premier.

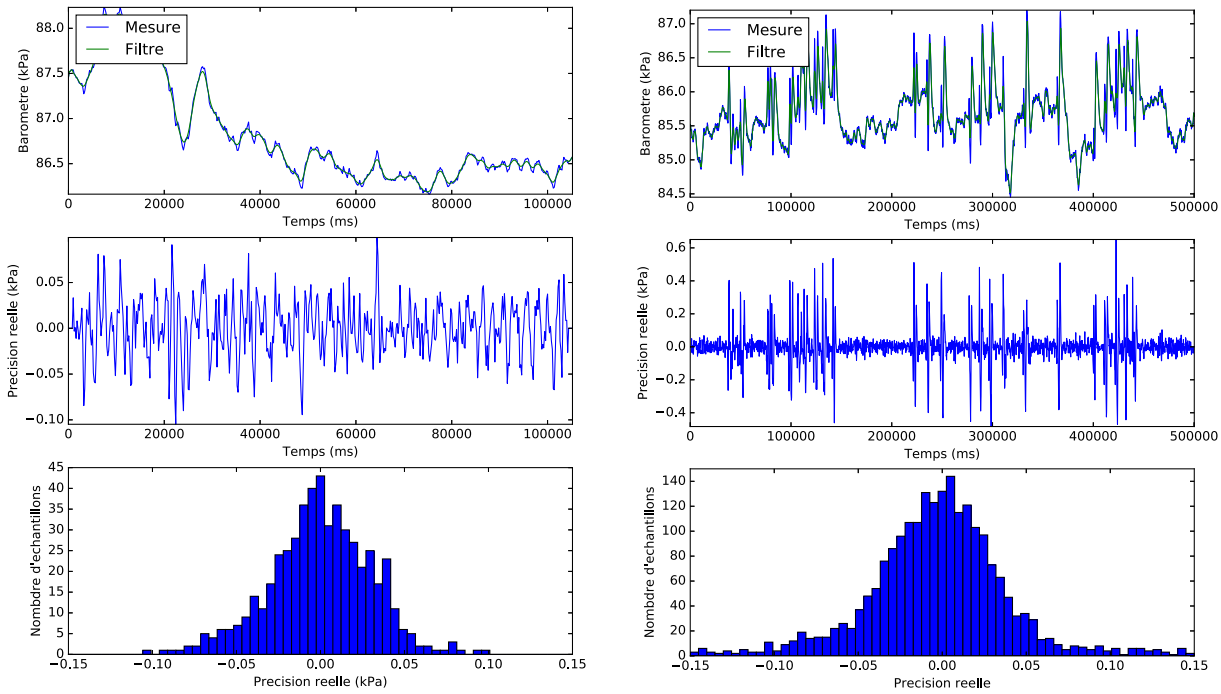

FIGURE III.15: Mesures du baromètre pour deux vols (un vol par colonne) et leur évaluation de la précision réelle (dans le temps et histogramme).

III.3.3 Détection d'anomalies (actes malveillants)

Les scénarios présentés précédemment ont été analysés pour identifier les imperfections et dimensions qui peuvent être impactées. Le Tableau III.5 illustre les résultats de ces expérimentations. Les colonnes nous donnent différentes informations : où le scénario s'est produit, le facteur qui l'a produit, la facilité de détection, le coût qu'il peut produire et les éléments impactés de la qualité. Ce tableau ne représente pas l'étude d'un seul sous-système. Il intègre les éléments impactés dans différents sous-systèmes avec l'objectif de présenter tous les éléments qui peuvent avoir une importance dans chaque scénario. La notation utilisée pour la deuxième et troisième colonne est la même que pour le cas des cuves (Sous-section III.2.4).

Dans ce tableau, nous voyons que chaque scénario a des caractéristiques bien spécifiques. La facilité de détection ainsi que le coût de ne pas réagir pour les corriger sont présentés dans leurs colonnes respectives. Les scénarios étudiés impactent différents éléments de la

TABLE III.5: Catégorisation des anomalies identifiées pour les drones à partir des éléments [a] impactés.

Scénario	**Type : Origine**	**Type : Facteur**	**Facilité de détection**	**Coût d'erreur**	I_{err}	I_{inc}	id_{con}	id_{uni}	cd_{rp}	cd_{err}	cd_{tim}	ed_{for}	ed_{coh}
Mesure d'altitude bloquée	M	2.a	Facile	Contrôle chaotique					x	x			x
Camera bloquée	M	2.a	Difficile	Contrôle difficile / Pas de reconnaissance du terrain					x				x
Perte de puissance	S	2.a	Facile	Chute ou perte du drone			x		x				x
Mauvaise calibration	S	1/2.a	Elle dépend (du trajet)	Décisions basées sur des fausses informations						x			
Batterie abîmée	S	1/2.a	Facile	Fonctionnement aléatoire / Chute du drone	x	x	x		x	(x)	(x)		x
Sous-système endommagé	S	1/2.a	Elle dépend	Il dépend	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)
Ajout de poids	S	2.a	Facile	Autonomie réduite			x		x				x
Hélices mal équilibrées	S	1/2.a	Difficile	Chute ou perte du drone					x				x
Perte de connexion	N	1/2.a	Facile	Drone à la dérive	x	x					x		
Prise de contrôle	N	2.b	Facile	Perte du contrôle / Perte du drone	(x)			(x)		x	(x)		x
Spoofing	N	1/2.a	Elle dépend (des données)	Perte de contrôle / Décisions basées sur de fausses informations	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)	(x)

a. Définition des acronymes dans la Section II.3

qualité. En conséquence, à partir des éléments impactés lors d'une détection, nous pouvons catégoriser le scénario subi. Pour automatiser ce mécanisme, nous pouvons utiliser un ensemble de fonctions conditionnelles formées avec les éléments marqués dans le tableau. Toutefois, une catégorisation complète est impossible du fait de l'existence de scénarios qui impactent les mêmes éléments de la qualité et que plusieurs scénarios peuvent avoir lieu au même moment. Dans ces cas, un ensemble de solutions peut être proposé. Ceci permettra de donner une réponse en accord avec la catégorie identifiée.

Deux scénarios peuvent avoir lieu en même temps, être indépendants ou bien être la conséquence l'un de l'autre. Par exemple, une cyberattaque de type *spoofing* peut avoir comme objectif la prise de contrôle du drone. C'est pourquoi les dimensions correspondantes aux deux scénarios peuvent être impactées simultanément.

Comme pour le cas d'étude précédent, nous observons comment certaines dimensions sont impactées par plus de scénarios que d'autres. Ainsi certaines dimensions sont plus pertinentes pour la détection que pour la catégorisation et inversement. La précision réelle et la cohérence sont deux exemples de dimensions qui sont impactées par presque tous les scénarios. En revanche, les données erronées et l'opportunisme sont des dimensions qui permettent de catégoriser les détections parce qu'elles correspondent à des scénarios particuliers.

La précision réelle, étant une des dimensions les plus impactées par les scénarios étudiés, a une grande importance. Si nous observons l'évaluation de cette dimension pour le baromètre du Crazyflie (Figure III.15), l'histogramme nous indique qu'elle a une distribution gaussienne. Ceci explique pourquoi la règle $68 - 95 - 99,7$ est utilisée. Nous avons obtenu l'AL résultant suivant :

$$AL_1 \equiv |cd_{rp}| < 0,12 \tag{III.15}$$

Dans les deux drones testés, nous pouvons paramétrer des limites pour la navigation. Cela veut dire qu'il est possible de limiter la puissance des moteurs et l'inclinaison des drones pour faciliter leur contrôle et assurer la sécurité des manœuvres. Les limites paramétrées et les contraintes physiques des drones forment des ensembles de valeurs qui bornent les états possibles des sous-systèmes. Toutes les valeurs hors de ces ensembles seront considérées comme erronées. Par exemple, les moteurs du Crazyflie sont limités par défaut à 80% de leur puissance maximale. Donc, leur AL sera celui de l'Équation III.16.

$$AL_2 \equiv cd_{err} < 0.8 \tag{III.16}$$

L'opportunisme pour les drones analysés a un comportement particulier que nous pouvons borner avec des ALs. Dans la Figure III.13, deux motifs ont été reconnus : un pour les états statiques et un autre pour les états de manœuvre. C'est pour cela qu'un AL comme le suivant peut limiter cette mesure :

$$AL_3 \equiv \begin{cases} 515 < cd_{tim} < 538 & \text{si statique} \\ cd_{tim} < 538 & \text{si en manœuvre} \end{cases} \tag{III.17}$$

Cet AL a été calculé grâce aux techniques d'histogramme. Pour le cas de manœuvre, aucun motif n'a pu être détecté. Néanmoins, une variance minimale sur des fenêtres temporelles pourrait borner l'aspect aléatoire de cette mesure.

En ce qui concerne le Crazyflie, nous voyons sur la Figure III.14 que les valeurs de l'opportunisme sont toujours multiples de $200ms$. De ce fait, la règle de l'Équation III.18 pourrait être adéquate pour décrire ce comportement :

$$AL_4 \equiv \frac{cd_{tim}}{200ms} \in \mathbb{N} \tag{III.18}$$

Grâce à ces ALs et d'autres, nous avons résumé dans le Tableau III.5 les éléments de la qualité impactés pour chaque scénario. Par exemple, un attaquant qui prend le contrôle du drone par une attaque de *spoofing* peut ne pas connaître les limites réglées pour le drone. Dans ce cas, il est possible que le système produise des données erronées, en injectant par exemple plus de puissance aux moteurs que celle qui est permise.

Chaque scénario analysé a ses particularités. Par exemple, une batterie abîmée ou qui n'est pas assez chargée pour une manœuvre peut produire des comportements inattendus et aléatoires. Cela peut déclencher d'autres scénarios comme la perte de puissance des moteurs. Plusieurs éléments de la qualité peuvent être liés entre eux comme l'opportunisme et l'apparition de données erronées et de données incomplètes à cause de problèmes dans la communication entre le drone et le récepteur. Ainsi, la cohérence entre les sous-systèmes est affectée quand un comportement aléatoire non corrélé apparaît. Chaque fois que les mesures de qualité dépassent leurs AL correspondants, la confiance dans le sous-système se réduit.

III.4 Comparaison et positionnement par rapport à d'autres méthodes

Cette section positionne les travaux réalisés dans les domaines explorés. Tout d'abord, la méthodologie pour la mesure de la qualité sera placée par rapport à l'état de l'art présenté dans la Section I.1. Nous expliquerons après où s'inscrivent dans notre travail les méthodes classiques de détection d'anomalies.

III.4.1 Comparaison avec d'autres méthodes de mesure de la qualité

En 2015, une réflexion menée dans le domaine de la qualité souligne que les limitations et caractéristiques des CPS nécessitaient des solutions particulières pour la mesure et l'amélioration de la qualité [SZ15]. Les travaux de cette thèse vont dans cette direction avec l'objectif de détecter les anomalies dans les CPS. Ces dernières années, plusieurs travaux avec des objectifs variés ont été réalisés sur différentes parties de ces systèmes. Ils ont dû faire face à ces limitations. Dans le Tableau III.6, ces travaux sont présentés et comparés.

TABLE III.6: Comparaison avec d'autres méthodes pour la mesure de la qualité

Méthode	**Domaine**	**Objectif**	**Niveaux étudiés**	**Quantité d'éléments**
[KL09b]	Management	Amélioration de la qualité des décisions du management	DI	5
[GL15]	Capteurs sans fil	Détection d'anomalies	I	5
[RB10]	Systèmes multi-source	Fusion d'information	I	28
[TLKC15]	Systèmes multi-source	Propagation de la qualité	DI	20
Cette thèse [MLBP17a]	CPS navals	Détection de cyberattaques, sabotages et anomalies	DIKW	45

Nous observons que les travaux examinés concernent différents domaines d'application. Le premier [KL09b] utilise des CPS pour prendre des décisions dans le management tandis que les autres cherchent à améliorer la performance des réseaux de capteurs. Le deuxième [GL15] a des applications environnementales et écologiques sur des ressources hydriques. Le

troisième [RB10] travaille sur les systèmes de fusion d'information et le quatrième [TLKC15] les applique aux systèmes radars, tandis que notre recherche est ciblée sur les systèmes navals.

Les objectifs sont variés également. L'objectif du premier travail [KL09b] est l'amélioration de la qualité des décisions prises pour le management à partir des mesures réalisées par des capteurs. Le deuxième [GL15] - comme notre travail - cherche une méthodologie pour la détection de certaines anomalies sur un réseau de capteurs sans fil. Les autres [RB10, TLKC15] ont comme objectif l'utilisation des mesures de qualité pour obtenir un meilleur résultat dans les sous-systèmes de fusion de flux d'information.

Une différence importante entre ces méthodologies concerne les niveaux de la pyramide DIKW analysés. Comme nous pouvons l'observer dans le tableau, seule notre recherche explore la pyramide DIKW au complet. De plus, les quatre études comparées cherchent à améliorer la qualité des connaissances sans entrer dans le détail d'évaluer leur qualité.

Une comparaison quantitative est impossible pour ces travaux. Seulement la quantité d'éléments de la qualité analysés dans chaque étude peut être mesurée. Nous parlons du total d'éléments parce que la différence entre donnée et information est floue. La première [KL09b] et la quatrième [TLKC15] étude utilisent 5 et 20 éléments respectivement sans aucune catégorisation particulière. Le deuxième travail [GL15] présente 5 dimensions qui groupent 14 attributs. Le troisième [RB10] réalise une hiérarchisation de 28 éléments catégorisés en trois groupes. La méthodologie proposée dans cette thèse propose 45 éléments au total pour les 4 niveaux de la pyramide avec une catégorisation en trois groupes pour les dimensions de la qualité. L'identification d'une grande quantité d'éléments de la qualité permettra une application plus générale de la méthode.

Dans toutes les études présentées, le résultat de l'évaluation de la qualité est donné comme un ensemble d'éléments qui la composent. Une question récurrente est comment ces valeurs peuvent être résumées avec une seule mesure du type « bonne/mauvaise qualité ». Ces travaux proposent d'utiliser des qualités multi variables pour les appliquer dans des méthodologies multicritères *a posteriori*. Cependant, d'autres approches [KS16] essayent d'intégrer ces valeurs avec les théories de fusion d'informations pour obtenir une valeur globale de qualité.

Dans la Section II.4.3, nous avons présenté la manière selon laquelle l'approche proposée peut s'intégrer au processus classique d'amélioration de la qualité introduit par la méthodologie TDQM [WS96] avec certaines modifications. Seul l'autre travail sur la détection d'anomalies [GL15] suit un comportement similaire. Le reste des travaux comparés dans cette

section ont choisi de suivre la définition classique.

III.4.2 Positionnement de l'approche par rapport aux autres méthodes de détection d'anomalies

Notre approche applique des outils du domaine de la détection d'anomalies pour démontrer la première hypothèse présentée dans cette thèse *c.-à-d.* les actes malveillants et dysfonctionnements ont un impact sur les mesures de qualité. Ces systèmes de détection d'anomalies s'inscrivent dans deux étapes de l'approche : pour l'évaluation de la qualité ou pour le calcul des ALs.

Dans les cas d'étude, nous avons utilisé des règles pour l'évaluation des différentes dimensions. Par exemple, l'information du volume est considérée comme correcte quand une règle prédéfinie avec les mesures possibles est respectée. À cause de l'hétérogénéité des imperfections et des dimensions, certaines techniques sont plus pertinentes que d'autres pour le calcul de leurs ALs respectifs. Nous avons observé que souvent le bruit mesuré par la précision réelle corrèle avec une distribution gaussienne, et donc les méthodes adaptées à ce type de signaux seront préférées. Dans les CPS, on trouve constamment des sous-systèmes avec des états de fonctionnement limités et en conséquence les méthodes basées sur les histogrammes sont plus appropriées pour leur exploration. Ceci facilite d'autre part l'application de règles afin de définir des ALs.

L'approche proposée n'a pas le but de concurrencer d'autres méthodes de détection d'anomalies. Dans l'état des travaux réalisés, les techniques de détection d'anomalies enrichissent les méthodologies pour l'évaluation de la qualité. Ainsi, ce travail peut être vu comme le rapprochement entre les domaines de la détection d'anomalies et de la qualité dans le but de détecter et catégoriser les cyberattaques et les dysfonctionnements.

Les méthodologies de détection d'anomalies et dysfonctionnements analysent des scénarios ou sous-systèmes précis. Avec l'application navale comme référence, la détection d'attaques et d'anomalies dans l'AIS [MVA+17] et dans le GPS [SK14] sont des questions qui sont actuellement étudiées. Ces travaux sont difficilement applicables à d'autres sous-systèmes dû à leur spécificité, contrairement à l'approche proposée qui est d'application générale en principe pour les sous-systèmes des CPS. Certaines dimensions de la qualité peuvent intégrer les mesures ou des techniques particulières pour les calculs des ALs.

Ainsi, certains travaux ont étudié des dimensions très précises. Pour la détection d'in-

trusions réseau, la dimension « opportunisme » a été exhaustivement étudiée [BSP12]. La création des *whitelistings* avec l'application de règles [BSP13] peut être complètement intégrée à notre approche à travers de la dimension « réputation ».

III.5 Conclusion

Deux sous-systèmes critiques présents dans les systèmes navals ont été analysés en détail : une plate-forme avec deux cuves et deux drones aériens. Ils ont servi à examiner les deux hypothèses introduites dans le Chapitre II.

Pour chacun des cas d'étude, plusieurs scénarios de risque ont été identifiés, testés, évalués et analysés. À partir de ces expérimentations, nous avons confirmé que les anomalies et les cyberattaques considérées comme anomalies provoquées ont un impact sur la qualité. De surcroît, la catégorisation automatique de certains scénarios peut être réalisée en examinant les éléments affectés de la qualité. Ces résultats ont conduit à la validation de la méthode de détection d'anomalies à partir des mesures de qualité proposées dans cette thèse.

Ainsi, nous avons pu observer comment certaines mesures de qualité ont une importance majeure dans les CPS. Par exemple, la précision réelle ainsi que la cohérence ont été des indicateurs récurrents pour la détection de la plupart des scénarios étudiés. D'autres mesures comme les informations erronées ont démontré leur utilité pour la catégorisation plus précise des détections.

Chapitre IV

Conclusion générale et perspectives

Sommaire

IV.1 Problématique 109
IV.2 Travaux réalisés 110
IV.3 Discussion 112
IV.4 Travaux futurs 114

IV.1 Problématique

Les systèmes navals représentent actuellement une énorme valeur stratégique due entre autres au transport de marchandises. Ces systèmes sont chaque jour de plus en plus informatisés pour réaliser leurs missions d'une manière plus efficace, plus sûre et avec un équipage réduit. C'est pour cela qu'ils doivent être toujours plus protégés et surveillés. Cette protection passe par des mesures de cyberdéfense améliorées notamment en termes de détection de cyberattaques.

Les systèmes navals sont un cas particulier de systèmes cyber physiques (CPS). Leurs caractéristiques font que les systèmes classiques de détection d'intrusions ne sont pas adéquats. Par exemple, à cause du besoin de réponses en temps réel. Ainsi, l'environnement marin complexifie la détection. La durée de vie des navires pouvant aller jusqu'à 50 ans, les systèmes et les sous-systèmes navals doivent être protégés à long terme, quel que soit l'âge de ceux-ci. C'est pourquoi de nouvelles méthodologies de détection sont nécessaires. De plus, une catégorisation des détections est nécessaire pour pouvoir réagir comme il se doit aux me-

naces dans les meilleurs délais.

Dans cette thèse, nous avons proposé les mesures de qualité comme indices permettant de détecter les anomalies. Les méthodologies d'évaluation de la qualité jusqu'à nos jours ont été fortement intéressées par d'autres domaines comme le management. Récemment, des recherches sur des CPS spécifiques ont été réalisées, mais il n'existe pas d'approches génériques qui puissent être appliquées directement aux systèmes navals. En conséquence, cela ajoute un handicap à la résolution de la problématique.

IV.2 Travaux réalisés

Cette thèse a été le résultat de la recherche d'une méthodologie pour la détection des actes malveillants et dysfonctionnements sur les systèmes navals. Les scénarios ont été catégorisés en : cyberattaques, sabotages et anomalies environnementales. Cette première catégorisation permettra de les différencier et de pouvoir réagir à une anomalie en fonction de sa catégorie.

En vue de résoudre ce problème, nous avons réuni deux domaines qui n'étaient pas encore liés entre eux : la qualité et la détection d'anomalies. Quant au domaine de la qualité, la détection des données frauduleuses et des anomalies n'était pas identifiée comme un défi même si certains travaux ont obtenu ce résultat comme effet secondaire dans des scénarios très précis. À présent, la détection de fraudes et d'anomalies a été identifiée comme un défi propre au domaine de la qualité. De la même façon, la détection d'anomalies n'avait pas identifié les mesures de qualité comme un indice associé à ses méthodologies. En conséquence, nous avons décidé d'explorer cette symbiose avec l'objectif de profiter des caractéristiques des méthodologies existantes dans les deux domaines. L'utilisation des méthodes classiques de détection d'anomalies utilisant des éléments de qualité rapproche ces deux domaines jusqu'à présent assez éloignés.

Comme point de départ, deux hypothèses ont été formulées. Elles supposent que les anomalies peuvent avoir un impact sur les mesures de qualité et que les différents éléments impactés servent à catégoriser les détections d'anomalies. Pour examiner ces deux hypothèses, une approche a été construite. Cette approche est constituée de quatre phases : la mesure de la qualité, le calcul des niveaux d'acceptation (AL) pour ces mesures, la détection des anomalies et leur catégorisation.

Comme cela a été identifié dans la problématique de la thèse, aucune méthodologie générale n'existe pour la mesure de la qualité des flux créés par les sous-systèmes intégrés dans les CPS et plus particulièrement pour les systèmes navals. À partir de l'état de l'art étudié, une méthodologie de mesure de la qualité des données et de l'information a été proposée. Ainsi, la qualité de la connaissance et la qualité de l'intelligence ont été explorées pour compléter la définition de la qualité de toute la pyramide DIKW. À l'égard de ce type d'évaluation, la mesure de la qualité à ces deux niveaux est traitée pour la première fois. Cette évaluation complète installe les bases pour d'autres études et pourra servir par exemple au domaine de l'aide à la décision et d'autres perspectives présentées ultérieurement dans ce chapitre.

De manière à prouver la validité et l'intérêt de la mesure de la qualité et de l'approche pour la détection d'anomalies, deux cas d'études ont été matérialisés. Ils s'intéressent à deux sous-systèmes critiques dans les systèmes navals : un sous-système de stockage de deux cuves et deux drones aériens. Les cuves peuvent symboliser un système critique comme le stockage de combustible. Les drones aériens sont de plus en plus utilisés dans les bâtiments navals pour des missions spécifiques comme la surveillance. Dans les deux cas, plusieurs scénarios d'attaques, de sabotages et d'anomalies environnementales ont été identifiés. Quant à leur étude statique, les données de tous les sous-systèmes ont été enregistrées pour chacun des scénarios ainsi que dans un fonctionnement normal afin de créer une référence. Tous ces enregistrements ont mené à la création des *datasets* d'étude.

Les *datasets* ont servi à évaluer la qualité des flux de données de chaque sous-système identifié. Cette étude a permis de comprendre que certains éléments de la qualité ne seront pas évaluables systématiquement pour tous les sous-systèmes et dans toutes les conditions. Ainsi, nous constatons l'intérêt de catégoriser les dimensions en intrinsèques, contextuelles et extrinsèques en fonction de la connaissance qu'il est possible d'obtenir sur un système. Ces enregistrements illustrent également un exemple d'évaluation complète de la qualité d'une part, et la définition et l'utilisation de la connaissance d'autre part.

L'application de la méthode dans les deux cas d'étude permet d'affirmer que certains éléments de la qualité seront prédominants dans les CPS. C'est pourquoi nous avons observé une hiérarchie selon leur importance pour la détection d'anomalies. Par exemple, la précision réelle apparaît comme une dimension essentielle pour atteindre ce but, tandis que d'autres jouent un rôle secondaire. Chaque jour, ces systèmes sont de plus en plus informatisés, réduisant l'interaction avec les humains. De ce fait, les flux sont générés par des machines et non par des humains. En conséquence, différents éléments de la qualité seront influencés et

d'autres peuvent disparaître comme la subjectivité et l'imprécision de l'information.

L'utilisation de ces deux *datasets* a servi à examiner et à confirmer les deux hypothèses présentées. Tous les scénarios étudiés ont pu être détectés. Seules les cyberattaques de type *spoofing* et des sous-systèmes endommagés pourraient poser des problèmes pour leur détection en fonction de l'impact qu'ils peuvent produire. Une autre limitation est que certains scénarios impactent exactement et de la même manière les mêmes éléments de la qualité. De ce fait, la catégorisation du scénario n'est pas complète et seulement un indicateur de possibilités peut être donné.

À partir de l'ensemble de résultats présentés, nous considérons que la méthodologie proposée apporte une approche originale et appropriée basée sur des mesures de qualité, permettant de détecter et d'identifier les anomalies des CPS navals étudiés. Ce travail a posé les bases pour des travaux futurs et proposé les premières approches pour la résolution des problématiques présentées.

IV.3 Discussion

Même si la solution proposée permet de répondre partiellement aux problématiques présentées, elle comporte des limitations. Dans cette section, l'ensemble des résultats obtenus sont discutés.

Initialement, la qualité a démontré être un défi à l'égard de sa définition ainsi que la création de ses ALs car elle est multidimensionnelle et hétérogène. La première étape consiste à identifier manuellement les éléments de la qualité pertinents pour chaque sous-système. Ainsi, l'analyse des mesures est réalisée manuellement avec l'objectif d'identifier les techniques les plus adéquates pour la définition de leurs ALs correspondants. Plus de travail est nécessaire sur ce point pour automatiser cette tâche.

L'approche a été validée grâce à deux *datasets* qui représentent deux sous-systèmes critiques navals. Ces *datasets* étudient une quantité réduite de scénarios qui ont été réalisés dans des environnements contrôlés. À cause du nombre limité de logs d'étude et des systèmes étudiés, la validation de la méthodologie n'est que partielle. La création d'autres scénarios et l'accès à d'autres *datasets* permettra de tester plus en profondeur le modèle proposé et d'améliorer les résultats obtenus. Un point difficile de cette étape sera l'obtention ou la génération de *datasets* réalistes contenant des anomalies et cyberattaques.

Dans l'application de l'approche, nous avons observé l'apparition de faux positifs et de faux négatifs. Afin de réduire le taux de fausses détections, une étape postérieure est nécessaire pour adapter les AL aux scénarios identifiés. Ce point n'est pas traité en détail dans cette thèse parce que c'est une problématique des techniques utilisées pour le calcul des AL, qui fait partie de l'approche introduite. Dans les travaux futurs, une analyse complète dans ce sens doit être réalisée pour des applications concrètes.

Le besoin de réponses en temps réel est une des limitations les plus contraignantes dans le contexte des CPS. L'approche développée a montré son potentiel concernant ce point. Une analyse en profondeur de cet aspect sera nécessaire avant le passage à une vraie implémentation dans un système réel.

Les vecteurs résultants de l'évaluation de la qualité ainsi que les détections ne sont pas évidents pour des opérateurs qui ne sont pas familiarisés avec l'approche. Des interfaces et des logiciels peuvent faciliter leur interprétation. Ainsi des techniques d'agrégation peuvent être utiles pour traiter cet enjeu.

La méthodologie et l'approche considèrent individuellement les sous-systèmes pour l'évaluation de leur qualité. Par contre, ils font partie des réseaux interconnectés où ils interagissent entre eux. Avec ces communications, la qualité se transfère ainsi entre les sous-systèmes. Les travaux réalisés n'ont pas abordé la propagation de la qualité pour éviter les difficultés que les systèmes complexes présentent. Par exemple, quand nous mesurons la cohérence d'une information, les autres informations utilisées sont considérées comme parfaites.

D'après la comparaison avec l'état de l'art de la qualité, notre méthodologie se distingue parce qu'elle parcourt les quatre niveaux de la pyramide DIKW. L'analyse des quatre niveaux permet d'étendre l'évaluation de la qualité à la partie supérieure de la pyramide. En ce qui concerne les CPS, cela permet d'évaluer l'intelligence, où se trouvent les algorithmes et les procédures, et les connaissances, qui groupent les résultats et réactions déterminés.

La définition du contexte est une question difficile à résoudre [Pet10]. Dans le cas d'étude, nous avons défini des contextes d'utilisation pour certaines informations. Par exemple, des conditions ou des intervalles ont été définis pour certains AL. Seulement quelques contextes ont été pris en compte afin de montrer leur intérêt avec un exemple simplifié. Une analyse plus exhaustive permettrait d'améliorer les résultats obtenus.

IV.4 Travaux futurs

Grâce à la diversité des problèmes étudiés, les perspectives de ce travail sont multiples et variées. Ainsi, l'identification et la construction des cas d'étude ont entraîné un effort considérable ayant un impact non négligeable sur l'approfondissement de certains sujets. Par conséquent, nous suggérons d'étudier davantage le modèle proposé et d'analyser les nouvelles problématiques identifiées.

Comme cela a été indiqué dans la discussion sur les résultats, l'identification des éléments pertinents de la qualité pour chaque sous-système ainsi que leur analyse pour fixer leurs ALs sont réalisées manuellement. L'automatisation de ces deux étapes doit être explorée. Pendant l'application de l'approche sur les cas d'étude, certaines relations ont été identifiées. Plusieurs éléments de la qualité se répètent en fonction du type de sous-système. Ainsi, en fonction du type de données et du sous-système, certaines techniques pour la détection d'anomalies paraissent plus adéquates. Ces connexions peuvent constituer des pistes de recherche.

Pour l'instant, l'approche a prouvé être utile et valide pour deux cas particuliers de CPS. Des pistes pour les travaux futurs comportent la validation avec un nombre plus important et varié de systèmes. De plus, même si les scénarios représentent des situations réelles, il s'agit d'expérimentations complètement contrôlées donc une implémentation sur des systèmes réels fera sans aucun doute apparaître de nombreux défis. Par exemple, dans la définition des ALs il faudra prendre en compte une multitude de contextes. Concernant le cas d'étude des cuves, l'influence de l'état de la mer sur la précision réelle devra être prise en compte.

L'interaction des sous-systèmes se traduit par un problème de propagation de la qualité. Une thèse a été dédiée à cette problématique dans les systèmes complexes d'information [TLKC15]. La validation et réponse complète à cette question est loin d'être résolue à cause de la propagation variable de la qualité. Dans la littérature, des premières approches de cette problématique existent considérant plusieurs points de vue comme la propagation de la qualité des données capteur vers les décisions du management [KL09a]. Ces travaux ciblent des problématiques particulières sans application générale. En conséquence, il y en a encore un manque de méthodologies pour son analyse.

Un exercice intéressant qui n'a pas été abordé pendant cette thèse est l'implémentation directe de l'approche sur les systèmes étudiés. Cela permettrait de construire un démonstrateur de l'approche et la réalisation de tests en direct. Cette implémentation pro-

met d'être particulièrement longue, car il faut dédier un temps de découverte des systèmes et leur fonctionnement. Dans le cas des cuves, une première étape d'ingénierie inverse a été déjà réalisée ainsi que l'installation d'un logiciel appelé ScadaBR [1] qui permet leur contrôle et l'automatisation du procès. Dans le cas des drones, plusieurs travaux ont été réalisés par des chercheurs pour faciliter l'utilisation du Parrot spider comme une plate-forme de recherche. Le Crazyflie est un drone de développement ouvert ce qui permettrait d'intégrer des scénarios et fonctionnalités dans le code original.

Une des limitations des CPS est la contrainte du temps réel. Cette limitation a guidé de forme transversale les travaux réalisés, mais du fait de son importance majeure, une étude plus approfondie doit être réalisée. Une implémentation en temps réel permettrait d'intégrer ce type d'approches sur des systèmes réels. Sans cela leur utilisation servira uniquement à une identification des anomalies *a posteriori*.

Dans les méthodologies de mesure de qualité, les humains sont classiquement considérés comme une entité spéciale qui crée et consomme des informations. Nous avons considéré les humains comme un autre sous-système qui produit et consomme des flux de données et d'information à travers des interfaces d'interaction. Cette généralisation a permis d'établir la nouvelle catégorisation des dimensions de l'information. À présent, l'humain a une certaine importance dans le contrôle et évaluation des systèmes, par exemple le pilotage des drones. L'automatisation et l'informatisation des systèmes font que l'humain a un rôle de plus en plus réduit. Plus de travail est donc nécessaire sur ce point afin de pouvoir « déshumaniser » les méthodes utilisées dans les CPS. Par exemple, certaines dimensions identifiées comme la « croyance » ne représentent plus qu'une « cohérence » entre l'humain et l'information évaluée donc elles deviennent redondantes. Une exploration en profondeur devra être réalisée pour évaluer et définir comment affronter ce défi.

1. ScadaBR est un logiciel libre qui permet le contrôle et supervision de systèmes industriels. Site officiel : `http://www.scadabr.com.br/`

Liste de publications

IV.5 Article de revue

— Pedro Merino Laso, David Brosset, et John Puentes. Dataset of anomalies and malicious acts in a cyber-physical subsystem. Data in Brief, 14 :186–191, 2017.

IV.6 Articles de conférence

— Pedro Merino Laso, David Brosset, et John Puentes. Monitoring approach of cyber-physical systems by quality measures. In International Conference on Sensor Systems and Software, pages 105–117. Springer, 2016.

— Pedro Merino Laso, David Brosset, et John Puentes. Analysis of quality measurements to categorize anomalies in sensor systems. In Proceedings of computing conference 2017, pages 1330–1338, 2017.

Bibliographie

[AA14] Andronicus A Akinyelu et Aderemi O Adewumi. Classification of phishing email using random forest machine learning technique. *Journal of Applied Mathematics*, 2014, 2014.

[AAR96] Andreas Arning, Rakesh Agrawal, et Prabhakar Raghavan. A linear method for deviation detection in large databases. In *KDD*, pages 164–169, 1996.

[AC89] Bovas Abraham et Alice Chuang. Outlier detection and time series modeling. *Technometrics*, 31(2) :241–248, 1989.

[Acc17] Accenture. Cost of cyber crime study : Insights on the security investments that make a difference. Technical report, Ponemon Institute, 2017.

[Agg05] Charu C Aggarwal. On abnormality detection in spuriously populated data streams. In *Proceedings of the 2005 SIAM International Conference on Data Mining*, pages 80–91. SIAM, 2005.

[Ard83] Sasan Ardalan. Kalman parametric and nonparametric system identification applied to echo cancellation. Technical report, Center for communications and signal processing. North Carolina State University, 1983.

[BC11] Bruce Barnett et Andrew Crapo. A semantic model for cyber security. In *Proceedings of the Fifth Grid-Interop Forum. Gridwise Architectural Council*, 2011.

[BG11] Radhakisan Baheti et Helen Gill. Cyber-physical systems. *The impact of control technology*, 12 :161–166, 2011.

[BG16] Anna L Buczak et Erhan Guven. A survey of data mining and machine learning methods for cyber security intrusion detection. *IEEE Communications Surveys & Tutorials*, 18(2) :1153–1176, 2016.

[BGS+13] Joel W Branch, Chris Giannella, Boleslaw Szymanski, Ran Wolff, et Hillol Kargupta. In-network outlier detection in wireless sensor networks. *Knowledge and information systems*, 34(1) :23–54, 2013.

[BH01] Richard J Bolton et David J et al. Hand. Unsupervised profiling methods for fraud detection. *Credit Scoring and Credit Control VII*, pages 235–255, 2001.

[BLH99] R Brause, T Langsdorf, et Michael Hepp. Neural data mining for credit card fraud detection. In *Tools with Artificial Intelligence, 1999. Proceedings. 11th IEEE International Conference on*, pages 103–106. IEEE, 1999.

[BP85] Donald P Ballou et Harold L Pazer. Modeling data and process quality in multi-input, multi-output information systems. *Management science*, 31(2) :150–162, 1985.

[BS06] Carlo Batini et Monica Scannapieco. *Data Quality : Concepts, Methodologies and Techniques (Data-Centric Systems and Applications)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

[BSP12] Rafael Ramos Regis Barbosa, Ramin Sadre, et Aiko Pras. Towards periodicity based anomaly detection in scada networks. In *Emerging Technologies & Factory Automation (ETFA), 2012 IEEE 17th Conference on*, pages 1–4. IEEE, 2012.

[BSP13] Rafael Ramos Regis Barbosa, Ramin Sadre, et Aiko Pras. Flow whitelisting in scada networks. *International journal of critical infrastructure protection*, 6(3) :150–158, 2013.

[BW03] David Bailey et Edwin Wright. *Practical SCADA for Industry*. Newnes, 2003.

[BYTE16] Sofia Belikovetsky, Mark Yampolskiy, Jinghui Toh, et Yuval Elovici. dr0wned-cyber-physical attack with additive manufacturing. *preprint arXiv :1609.00133*, 2016.

[CBK09] Varun Chandola, Arindam Banerjee, et Vipin Kumar. Anomaly detection : A survey. *ACM computing surveys (CSUR)*, 41(3) :15, 2009.

[CCBGA05] Frédéric Cuppens, Nora Cuppens-Boulahia, et Joaquin Garcia-Alfaro. Detection and removal of firewall misconfiguration. In *Proceedings of the 2005 IASTED International Conference on Communication, Network and Information Security*, volume 1, pages 154–162, 2005.

[Cen14] Center for Strategic and International Studies. Net losses : Estimating the global cost of cybercrime. economic impact of cybercrime ii. Technical report, McAfee, 2014.

[CEWB97] Kenneth C Cox, Stephen G Eick, Graham J Wills, et Ronald J Brachman. Brief application description ; visual data mining : Recognizing telephone calling fraud. *Data Mining and Knowledge Discovery*, 1(2) :225–231, 1997.

[Che01] Zhengxin Chen. *Data mining and uncertain reasoning : an integrated approach*. Wiley New York, 2001.

[CHMS09] Yuan Cao, Haibo He, Hong Man, et Xiaoping Shen. Integration of self-organizing map (som) and kernel density estimation (kde) for network intrusion detection. In *SPIE Europe Security+ Defence*, pages 74800N–74800N. International Society for Optics and Photonics, 2009.

[Coh95] William W Cohen. Fast effective rule induction. In *Proceedings of the twelfth international conference on machine learning*, pages 115–123, 1995.

[Com16] European Commission. Eu transport in figures. statistical pocketbook 2016. Technical report, European Union, 2016.

[CPGM06] Vasilis Chatzigiannakis, Symeon Papavassiliou, Mary Grammatikou, et B Maglaris. Hierarchical anomaly detection in distributed large-scale sensor networks. In *Computers and Communications, 2006. ISCC'06. Proceedings. 11th IEEE Symposium on*, pages 761–767. IEEE, 2006.

[CR13] A. M. Chandrasekhar et K. Raghuveer. Intrusion detection technique by using k-means, fuzzy neural network and svm classifiers. In *2013 International Conference on Computer Communication and Informatics*, pages 1–7, Jan 2013.

[CYB00] Kim B. Clark Carliss Young Baldwin. *Design Rules : The power of modularity*. MIT Press, 2000.

[CZ15] Li Cai et Yangyong Zhu. The challenges of data quality and data quality assessment in the big data era. *Data Science Journal*, 14, 2015.

[DA04] Batini De Amicis. A methodology for data quality assessment on financial data. *Studies in Communication Sciences*, 4 :115–136, 2004.

[DD16] Sumeet Dua et Xian Du. *Data mining and machine learning in cybersecurity*. CRC press, 2016.

[Dir06] Direction centrale de la sécurité des systèmes d'information. Menaces sur les systèmes informatiques guide num 650. Technical report, Secrétariat général de la défense nationale, 2006.

[Edg87] FY Edgeworth. Xli. on discordant observations. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 23(143) :364–375, 1887.

[Eng99] Larry P. English. *Improving Data Warehouse and Business Information Quality : Methods for Reducing Costs and Increasing Profits*. John Wiley & Sons, Inc., New York, NY, USA, 1999.

[Esk00] Eleazar Eskin. Anomaly detection over noisy data using learned probability distributions. In *In Proceedings of the International Conference on Machine Learning*. Citeseer, 2000.

[Fed11] Federal Ministery of Interior. Cyber security strategy for germany. Technical report, Germany government, 2011.

[FMC11] Nicolas Falliere, Liam O Murchu, et Eric Chien. W32. stuxnet dossier. *White paper, Symantec Corp., Security Response*, 5 :69, 2011.

[FZHH14] Wenying Feng, Qinglei Zhang, Gongzhu Hu, et Jimmy Xiangji Huang. Mining network data for intrusion detection through combining svms with ant colony networks. *Future Generation Computer Systems*, 37 :127–140, 2014.

[GA10] Mohinder S Grewal et Angus P Andrews. Applications of kalman filtering in aerospace 1960 to the present [historical perspectives]. *IEEE Control Systems*, 30(3) :69–78, 2010.

[GAJM16] Jonathan Goh, Sridhar Adepu, Khurum Nazir Junejo, et Aditya Mathur. A dataset to support research in the design of secure water treatment systems. In *The 11th International Conference on Critical Information Infrastructures Security*, 2016. https ://itrust.sutd.edu.sg/dataset/.

[GL15] Jianwen Guo et Feng Liu. Automatic data quality control of observations in wireless sensor network. *Geoscience and Remote Sensing Letters, IEEE*, 12(4) :716–720, April 2015.

[GSM+11] R. Gandhi, A. Sharma, W. Mahoney, W. Sousan, Qiuming Zhu, et P. Laplante. Dimensions of cyber-attacks : Cultural, social, economic, and political. *Technology and Society Magazine, IEEE*, 30(1) :28–38, 2011.

[GWCC08] Marco Guerriero, Peter Willett, Stefano Coraluppi, et Craig Carthel. Radar/ais data fusion and sar tasking for maritime surveillance. In *Information Fusion, 2008 11th International Conference on*, pages 1–5. IEEE, 2008.

[HB95] Eric Horvitz et Matthew Barry. Display of information for time-critical decision making. In *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*, pages 296–305. Morgan Kaufmann Publishers Inc., 1995.

[HBS16] Franz Hamilton, Tyrus Berry, et Timothy Sauer. Ensemble kalman filtering without a model. *Physical Review X*, 6(1) :011021, 2016.

[HLV03] Wenjie Hu, Yihua Liao, et V Rao Vemuri. Robust anomaly detection using support vector machines. In *Proceedings of the international conference on machine learning*, pages 282–289, 2003.

[HMA07] David J Hill, Barbara S Minsker, et Eyal Amir. Real-time bayesian anomaly detection for environmental sensor data. In *Proceedings of the Congress-International Association for Hydraulic Research*, volume 32, page 503. Citeseer, 2007.

[HOA+15] Nutan Farah Haq, Abdur Rahman Onik, Md Avishek, Khan Hridoy, Musharrat Rafni, Faisal Muhammad Shah, et Dewan Md Farid. Application of machine learning approaches in intrusion detection system : a survey. *International Journal of Advanced Research in Artificial Intelligence*, 2015.

[IKP95] K. Ilgun, R. A. Kemmerer, et P. A. Porras. State transition analysis : a rule-based intrusion detection approach. *IEEE Transactions on Software Engineering*, 21(3) :181–199, Mar 1995.

[INC07] INCOSE. *Systems engineering handbook. A guide for system life cycle processes and activities.* INCOSE, 2007.

[ISA] ISA. Ansi/isa-95.

[ISO05] ISO. Iec 27001 : 2005. information technology. security techniques. information security management systems. requirements., 2005.

[JRMB+15] Hossein Joudaki, Arash Rashidian, Behrouz Minaei-Bidgoli, Mahmood Mahmoodi, Bijan Geraili, Mahdi Nasiri, et Mohammad Arab. Using data mining to detect health care fraud and abuse : a review of literature. *Global journal of health science*, 7(1) :194, 2015.

[Kar14] Argyro P Karanasiou. The changing face of protests in the digital age : on occupying cyberspace and distributed-denial-of-services (ddos) attacks. *International Review of Law, Computers & Technology*, 28(1) :98–113, 2014.

[KBD+15] Nizar Kheir, Gregory Blanc, Hervé Debar, Joaquin Garcia-Alfaro, et Dingqi Yang. Automated classification of c&c connections through malware url clustering. In *IFIP International Information Security Conference*, pages 252–266. Springer, 2015.

[KL08] Florian Knorn et Douglas J Leith. Adaptive kalman filtering for anomaly detection in software appliances. In *INFOCOM Workshops 2008, IEEE*, pages 1–6. IEEE, 2008.

[KL09a] Anja Klein et Wolfgang Lehner. How to optimize the quality of sensor data streams. In *Computing in the Global Information Technology, 2009. ICCGI'09. Fourth International Multi-Conference on*, pages 13–19. IEEE, 2009.

[KL09b] Anja Klein et Wolfgang Lehner. Representing data quality in sensor data streaming environments. *Journal of Data and Information Quality (JDIQ)*, 1(2) :10, 2009.

[Kni02] John C Knight. Safety critical systems : challenges and directions. In *Software Engineering, 2002. ICSE 2002. Proceedings of the 24rd International Conference on*, pages 547–550. IEEE, 2002.

[KNN16] Gunupudi Rajesh Kumar, Mangathayaru Nimmala, et Gugulothu Narasimha. An approach for intrusion detection using novel gaussian based kernel function. *Journal of Universal Computer Science*, 22(4) :589–604, 2016.

[Kru05] Ronald L Krutz. *Securing SCADA systems*. John Wiley & Sons, 2005.

[KS07] Takahisa Kobayashi et Donald L Simon. Hybrid kalman filter approach for aircraft engine in-flight diagnostics : Sensor fault detection case. *Journal of engineering for gas turbines and power*, 129(3) :746–754, 2007.

[KS16] Ron S Kenett et Galit Shmueli. *Information Quality : The Potential of Data and Analytics to Generate Knowledge*. John Wiley & Sons, 2016.

[LBK15] Jay Lee, Behrad Bagheri, et Hung-An Kao. A cyber-physical systems architecture for industry 4.0-based manufacturing systems. *Manufacturing Letters*, 3 :18–23, 2015.

[LC09] Simon Liu et Bruce Cheng. Cyberattacks : Why, what, who, and how. *IT professional*, 11(3) :14–21, 2009.

[LCD05] Anukool Lakhina, Mark Crovella, et Christophe Diot. Mining anomalies using traffic feature distributions. In *ACM SIGCOMM Computer Communication Review*, volume 35, pages 217–228. ACM, 2005.

[LG12] John Zhong Lei et Ali A Ghorbani. Improved competitive learning neural networks for network intrusion and fraud detection. *Neurocomputing*, 75(1) :135–145, 2012.

[LKT15] Wei-Chao Lin, Shih-Wen Ke, et Chih-Fong Tsai. Cann : An intrusion detection system based on combining cluster centers and nearest neighbors. *Knowledge-based systems*, 78 :13–21, 2015.

[Los01] David Loshin. *Enterprise Knowledge Management. The Data Quality Approach*. Academic Press, 2001.

[LSKW02] Yang W. Lee, Diane M. Strong, Beverly K. Kahn, et Richard Y. Wang. Aimq : A methodology for information quality assessment. *Inf. Manage.*, 40(2) :133–146, 12 2002.

[LV02] Yihua Liao et V Rao Vemuri. Use of k-nearest neighbor classifier for intrusion detection. *Computers & security*, 21(5) :439–448, 2002.

[LX01] Wenke Lee et Dong Xiang. Information-theoretic measures for anomaly detection. In *Security and Privacy, 2001. S&P 2001. Proceedings. 2001 IEEE Symposium on*, pages 130–143. IEEE, 2001.

[Min96] J.C. Mingers. An evaluation of theories of information with regard to the semantic and pragmatic aspects of information systems. *Systems practice*, 9(3) :187–209, 1996.

[MLBP16] Pedro Merino Laso, David Brosset, et John Puentes. Monitoring approach of cyber-physical systems by quality measures. In *International Conference on Sensor Systems and Software*, pages 105–117. Springer, 2016.

[MLBP17a] Pedro Merino Laso, David Brosset, et John Puentes. Analysis of quality measurements to categorize anomalies in sensor systems. In *Proceedings of computing conference 2017*, pages 1330–1338, 2017.

[MLBP17b] Pedro Merino Laso, David Brosset, et John Puentes. Dataset of anomalies and malicious acts in a cyber-physical subsystem. *Data in Brief Journal*, 14 :186–191, 2017.

[MOD96] MODICON, Inc. *Modicon. Modbus Protocol. Reference Guide*, 1996.

[MPS+03] David Moore, Vern Paxson, Stefan Savage, Colleen Shannon, Stuart Staniford, et Nicholas Weaver. Inside the slammer worm. *IEEE Security & Privacy*, 1(4) :33–39, 2003.

[MS96] Amihai Motro et Philippe Smets. *Uncertainty management in information systems : from needs to solutions*. Springer Science & Business Media, 1996.

[MSJ10] Min Seok Mok, So Young Sohn, et Yong Han Ju. Random effects logistic regression model for anomaly detection. *Expert Systems with Applications*, 37(10) :7162–7166, 2010.

[MTT15] Thomas H Morris, Zach Thornton, et Ian Turnipseed. Industrial control system simulation and data logging for intrusion detection system research. In *7th Annual Southeastern Cyber Security Summit. Huntsvile, AL*, 2015. https ://sites.google.com/a/uah.edu/tommy-morris-uah/ics-data-sets.

[MVA+17] Fabio Mazzarella, Michele Vespe, Alfredo Alessandrini, Dario Tarchi, Giuseppe Aulicino, et Antonio Vollero. A novel anomaly detection approach to

identify intentional ais on-off switching. *Expert Systems with Applications*, 78 :110–123, 2017.

[MWLZ09] Stuart E Madnick, Richard Y Wang, Yang W Lee, et Hongwei Zhu. Overview and framework for data and information quality research. *Journal of Data and Information Quality (JDIQ)*, 1(1) :2, 2009.

[Nat04] National Marine Electronics Association and others. Nmea 2000 standard for serial-data networking of marine electronic devices. Technical report, 2004.

[Nau01] Felix Naumann. From databases to information systems-information quality makes the difference. In *IQ*, pages 244–260, 2001.

[NC03] Caleb C Noble et Diane J Cook. Graph-based anomaly detection. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 631–636. ACM, 2003.

[NP16] Jeyasingam Nivethan et Mauricio Papa. On the use of open-source firewalls in ics/scada systems. *Information Security Journal : A Global Perspective*, 25(1-3) :83–93, 2016.

[OCM12] Leo Obrst, Penny Chase, et Richard Markeloff. Developing an ontology of the cyber security domain. In *STIDS*, pages 49–56, 2012.

[OCWM14] Alessandro Oltramari, Lorrie Faith Cranor, Robert J Walls, et Patrick D McDaniel. Building an ontology of cyber security. In *STIDS*, pages 54–61, 2014.

[Pet10] Mathieu Petit. *Approche spatiale pour la caractérisation du contexte d'exécution d'un système d'information ubiquitaire.* PhD thesis, Arts et Métiers ParisTech. IRENav - École Navale, 2010.

[Pit11] Roland E Pittman. *Taxonomy of learning. A Brief Introduction.* PediaPress, 2011.

[PLW02] Leo L. Pipino, Yang W. Lee, et Richard Y. Wang. Data quality assessment. *Commun. ACM*, 45(4) :211–218, 4 2002.

[PMLL13] John Puentes, Julien Montagner, Laurent Lecornu, et Jaakko Lähteenmäki. Quality analysis of sensors data for personal health records on mobile devices.

In Rajeev Bali, Indrit Troshani, Steve Goldberg, et Nilmini Wickramasinghe, editors, *Pervasive Health Knowledge Management*, Healthcare Delivery in the Information Age, pages 103–133. Springer New York, 2013.

[PS14] J. Petit et S.E. Shladover. Potential cyberattacks on automated vehicles. *Intelligent Transportation Systems, IEEE Transactions on*, PP(99) :1–11, 2014.

[PWC15] PWC. The global state of information security® survey 2015 - managing cyber risks in an interconnected world. Technical report, pwc, 2015.

[PWC16] PWC. The global state of information security® survey 2016 - turnaround and transformation in cybersecurity. Technical report, pwc, 2016.

[PWC17] PWC. The global state of information security® survey 2017 - how businesses are addressing iot risks. Technical report, pwc, 2017.

[RB10] Galina L Rogova et Eloi Bosse. Information quality in information fusion. In *Information Fusion (FUSION), 2010 13th Conference on*, pages 1–8. IEEE, 2010.

[RC08] Julian Rrushi et Roy Campbell. Detecting cyber attacks on nuclear power plants. In *Critical Infrastructure Protection II*, pages 41–54. Springer, 2008.

[RH08] Jennifer E Rowley et Richard J Hartley. *Organizing knowledge : an introduction to managing access to information*. Ashgate Publishing, Ltd., 2008.

[RHDCGA16] José Rubio-Hernán, Luca De Cicco, et Joaquín García-Alfaro. Revisiting a watermark-based detection scheme to handle cyber-physical attacks. In *Availability, Reliability and Security (ARES), 2016 11th International Conference on*, pages 21–28. IEEE, 2016.

[RLPB06] Sutharshan Rajasegarar, Christopher Leckie, Marimuthu Palaniswami, et James C Bezdek. Distributed anomaly detection in wireless sensor networks. In *Communication systems, 2006. ICCS 2006. 10th IEEE Singapore International Conference on*, pages 1–5. IEEE, 2006.

[RN04] Galina L Rogova et Vincent Nimier. Reliability in information fusion : literature survey. In *Proceedings of the seventh international conference on information fusion*, volume 2, pages 1158–1165, 2004.

[SCSC03] Mei-Ling Shyu, Shu-Ching Chen, Kanoksri Sarinnapakorn, et LiWu Chang. A novel anomaly detection scheme based on principal component classifier. Technical report, DTIC Document, 2003.

[SGLW08] L. Sha, S. Gopalakrishnan, X. Liu, et Q. Wang. Cyber-physical systems : A new frontier. In *2008 IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (sutc 2008)*, pages 1–9, June 2008.

[Sha48] C. E. Shannon. A mathematical theory of communication. *SIGMOBILE Mob. Comput. Commun. Rev.*, 5(1) :3–55, January 1948.

[Sha06] William T Shaw. *Cybersecurity for SCADA systems*. Pennwell books, 2006.

[SK14] S. C. Stubberud et K. A. Kramer. Threat assessment for gps navigation. In *2014 IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA) Proceedings*, pages 287–292, June 2014.

[Sme97] Philippe Smets. Imperfect information : Imprecision and uncertainty. In *Uncertainty management in information systems*, pages 225–254. Springer, 1997.

[ST12] Bhavin Shah et Bhushan H Trivedi. Artificial neural network based intrusion detection system : A survey. *International Journal of Computer Applications*, 39(6) :13–18, 2012.

[SVM+04] Monica Scannapieco, Antonino Virgillito, Carlo Marchetti, Massimo Mecella, et Roberto Baldoni. The daquincis architecture : A platform for exchanging and improving data quality in cooperative information systems. *Inf. Syst.*, 29(7) :551–582, 9 2004.

[SXZF07] Jimeng Sun, Yinglian Xie, Hui Zhang, et Christos Faloutsos. Less is more : Compact matrix decomposition for large sparse graphs. In *Proceedings of the 2007 SIAM International Conference on Data Mining*, pages 366–377. SIAM, 2007.

[Sym17] Symantec. Istr. internet security threat report. volume 22. Technical report, Symantec, 2017.

[SZ15] Kewei Sha et Sherali Zeadally. Data quality challenges in cyber-physical systems. *Journal of Data and Information Quality (JDIQ)*, 6(2-3) :8, 2015.

[TLKC15] Ion-George Todoran, Laurent Lecornu, Ali Khenchaf, et Jean-Marc Le Caillec. A methodology to evaluate important dimensions of information quality in systems. *J. Data and Information Quality*, 6(2-3) :11 :1–11 :23, June 2015.

[TLKLC13] I.-G. Todoran, L. Lecornu, A. Khenchaf, et J.-M. Le Caillec. Information quality evaluation in fusion systems. In *Information Fusion (FUSION), 2013 16th International Conference on*, pages 906–913, 7 2013.

[Tod14] Ion-George Todoran. *Étude dynamique de la qualité de l'information et des données d'un système d'information complexe.* PhD thesis, Télécom Bretagne, 2014.

[Uni16] United nations conference on trade and development. Review of maritime transport 2016. Technical report, United Nations, 2016.

[Wan98] Richard Y Wang. A product perspective on total data quality management. *Communications of the ACM*, 41(2) :58–65, 1998.

[Wan05] Yun Wang. A multinomial logistic regression modeling approach for anomaly intrusion detection. *Computers & Security*, 24(8) :662–674, 2005.

[WFA13] Yun Wang, Weihuang Fu, et Dharma P Agrawal. Gaussian versus uniform distribution for intrusion detection in wireless sensor networks. *IEEE Transactions on Parallel and Distributed systems*, 24(2) :342–355, 2013.

[WFP99] Christina Warrender, Stephanie Forrest, et Barak Pearlmutter. Detecting intrusions using system calls : Alternative data models. In *Security and Privacy, 1999. Proceedings of the 1999 IEEE Symposium on*, pages 133–145. IEEE, 1999.

[WP10] Philip Woodall et Ajith Kumar Parlikad. A hybrid approach to assessing data quality. *Proceedings of the 2010 international conference on information quality*, 1 :1–15, 2010.

[WS96] Richard Y. Wang et Diane M. Strong. Beyond accuracy : What data quality means to data consumers. *J. Manage. Inf. Syst.*, 12(4) :5–33, 3 1996.

[WYL+13] Jiafu Wan, Hehua Yan, Di Li, Keliang Zhou, et Lu Zeng. Cyber-physical systems for optimal energy management scheme of autonomous electric vehicle. *The Computer Journal*, 56(8) :947–956, 2013.

[Xia09] Jing Xiao. *Gestion des incertitudes dans le processus de développement de systèmes complexes*. PhD thesis, Institut National Polytechnique de Toulouse, 2009.

[YC02] Dit-Yan Yeung et Calvin Chow. Parzen-window network intrusion detectors. In *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, volume 4, pages 385–388. IEEE, 2002.

[YTWM00] Kenji Yamanishi, Jun-Ichi Takeuchi, Graham Williams, et Peter Milne. Online unsupervised outlier detection using finite mixtures with discounting learning algorithms. In *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 320–324. ACM, 2000.

[Zel05] Milan Zeleny. *Human Systems Management : Integrating Knowledge, Management and Systems*. World Scientific Publishing Co. Pte. Ltd., 2005.

[Zin07] Chaim Zins. Conceptual approaches for defining data, information, and knowledge. *Journal of the American Society for Information Science and Technology*, 58(4) :479–493, 2007.

[ZRM+16] Amrapali Zaveri, Anisa Rula, Andrea Maurino, Ricardo Pietrobon, Jens Lehmann, et Sören Auer. Quality assessment for linked data : A survey. *Semantic Web*, 7(1) :63–93, 2016.

Annexes

Annexe A

Dataset pour le cas d'étude du système des cuves

Dans cette annexe, une description détaillée du dataset utilisé pour l'évaluation de la méthodologie est présentée. Une description générale du système précède une analyse plus détaillée de chaque sous-système et du *dataset* créé pour étudier les anomalies, sabotages et dysfonctionnements qu'ils peuvent subir.

A.1 Détails de la plate-forme

A.1.1 Composants

La plate-forme est composée de plusieurs sous-systèmes. Leurs modèles et marques sont affichés sur la Table A.1 pour pouvoir bien les identifier. Ainsi, les valeurs de leurs états possibles sont représentées.

Table A.1: Composants installés sur le sous-système.

Composant	**Modèle**	**Marque**	**États**
PLC	TWDLCAE40DRF	Schneider Electric	———
Module extension	TM2AMI2HT	Schneider Electric	———
Capteur à ultrason	XX918A3F1M12	Telemecanique	0 - 10000
Pompes	Générique	Générique	ON/OFF
Flotteurs	Générique	Générique	ON/OFF
Alarme	Générique	Schneider Electric	ON/OFF

TABLE A.2: Registres du PLC.

Registre	**Bit**							
	0	1	2	3	4	5	6	7
2					IN_3	IN_2	IN_1	IN_0
3	P_2	P_1		AL				
4	Sonde à ultrason (8-bit int)							

Tous ces composants sont interconnectés grâce au PLC qui permet le contrôle et le monitoring. Le PLC utilise différents registres pour chacun d'eux. Leurs correspondances sont représentées sur le Tableau A.2.

A.2 Dataset

Pour pouvoir travailler *offline* sur les données générées par cette plate-forme, un *dataset* a été créé. Une explication du protocole, ainsi qu'une description des cas étudiés sont présentées dans la suite.

A.2.1 Protocole

Pour réaliser les logs qui forment le *dataset*, une machine a été connectée au réseau de "contrôle et monitoring" avec le seul objectif d'enregistrer ces fichiers. Pour récupérer les valeurs qui forment ces logs, cette machine de surveillance envoie de commandes aux PLC toutes les 0,1 seconde pour récupérer les valeurs de chaque registre de la mémoire. Toutes ces données sont enregistrées sur différents fichiers CSV (Comma-separated values).

Le *dataset* est composé par 15 fichiers CSV. La Table A.3 présente ces fichiers à mode de catalogue. Un ensemble d'enregistrements avec un état de normalité a été créé pour pouvoir entraîner ou paramétrer les algorithmes de détection et pouvoir détecter les anomalies causées. Ce log est assez long pour que les bruits et anomalies environnementales puissent être pris en compte. Les autres fichiers correspondent à différents scénarios qui présentent un risque pour la plate-forme.

TABLE A.3: Fichiers qui composent le *dataset* des cuves.

Scénario	**Sous-système affecté**	**Type d'événement**	**Durée (hh :mm :ss)**	**Taille**
Normal	Aucun	Normal	02 :01 :47	7.3 MB
Sac en plastique	Capteur à ultrason	Accident / Sabotage	00 :33 :20	4.2 MB
Mesure bloquée 1	Capteur à ultrason	Panne / Sabotage	00 :00 :25	74 KB
Mesure bloquée 2	Capteur à ultrason	Panne / Sabotage	00 :00 :17	48 KB
Objets flottants - cuve principale (2 objets)	Capteur à ultrason	Accident / Sabotage	00 :01 :35	272 KB
Objets flottants - cuve principale (7 objets)	Capteur à ultrason	Accident / Sabotage	00 :01 :22	234 KB
Humidité	Capteur à ultrason	Panne	00 :00 :18	52 KB
Panne du flotteur 1	Flotteur 1	Panne	00 :13 :55	1.8 MB
Panne du flotteur 2	Flotteur 2	Panne	00 :03 :40	610 KB
Attaque de déni de service (DoS)	Réseau	Cyberattaque	00 :01 :37	102 KB
Spoofing	Réseau	Cyberattaque	00 :34 :33	3.2 MB
Mauvaise connexion	Réseau	Panne / Sabotage	00 :15 :33	1.7 MB
Coups sur les cuves (intensité basse)	Tout le sous-système	Sabotage	00 :00 :39	112 KB
Coups sur les cuves (intensité moyenne)	Tout le sous-système	Sabotage	00 :00 :32	91 KB
Coups sur les cuves (intensité haute)	Tout le sous-système	Sabotage	00 :00 :33	95 KB

A.2.2 Structure

Une ligne d'un fichier du *dataset* est composée par : un *time stamp*, le numéro de registre et la valeur contenue dans cette espace de la mémoire. Le *Time Stamp* a le format suivant : $dd/mm/yyyy\ hh : mm : ss.sss$. Dans chaque registre différentes données peuvent être enregistrées. Dans la Table A.2, les sous-systèmes sont identifiés sur l'espace mémoire. Les flotteurs sont représentés par IN_i, les pompes par P_j et l'alarme par AL.

Pour simplifier l'utilisation du *dataset*, un script de lecture a été inclus. Ce script est écrit en python et permet de sélectionner un fichier CSV du *dataset*, créer des vecteurs avec les données de chaque capteur et les visualiser avec un plot. Cela permet d'utiliser les données sans avoir à maîtriser la structure des fichiers de log.

Annexe

B Acronymes

ADM : *Anomaly Detection Module.*

AI : *Artificial Intelligence.*

AIS : *Automatic Identification System.*

ANSSI : Agence Nationale de la Sécurité des Systèmes d'Information.

API : Automate Programmable Industriel.

CAD : *Computer Aided Design.*

CNUCED : Conférence des Nations Unies sur le Commerce Et le Développement.

CPS : *Cyber-Physical System.*

CRC : Contrôle de Redondance Cyclique.

DCS : *Distributed Control System.*

DDoS : *Distributed Denial of Service.*

DoS : *Denial of Service.*

DQV : *Data Quality Vector.*

FIR : *Finite Impulse Response.*

GIGO : *Garbage In Garbage Out.*

HIDS : *Host Intrusion Detection System.*

IARPA : *Intelligence Advanced Research Projects Activity.*

IDS : *Intrusion Detection System.*

IHM : Interface Homme-Machine.

IoT : *Internet of Things.*

IQV : *Information Quality Vector.*

IP : *Internet Protocol.*

IPS : *Intrusion Prevention System.*

ISID : *Industrial Security Incidents Database.*

knn : *k-nearest neighbor.*

KQV : *Knowlege Quality Vector.*

NIDS : *Network Intrusion Detection System.*

nmap : *Network Mapper.*

NSA : *National Security Agency*

PCA : Principal Component Analysis.

PIB : Produit Intérieur Brut.

PLC : *Programmable Logic Controller.*

PUP : *Potentially Unwanted Programs.*

PwC : PricewaterhouseCoopers.

QoS : Quality of Service.

R&D : Recherche et Développement.

RAT : *Remote Access Trojan.*

RISI : *Repository of Industrial Security Incidents.*

SA : Système d'Automates.

SC : Système de Contrôle.

SCADA : *Supervisory Control and Data Acquisition.*

SQL : *Structured Query Language.*

SI : Système d'Information.

SIEM : *Security Information and Event Management.*

SVM : *Support Vector Machine.*

TEU : *Twenty-foot Equivalent Unit.*

To : Téraoctet.

VDR : *Voyage Data Recorder.*

WQV : *Wisdom Quality Vector.*

WWW : *World Wide Web.*

-

Résumé

Les systèmes navals représentent une infrastructure stratégique pour le commerce international et les activités militaires. Ces systèmes sont de plus en plus informatisés afin de réaliser une navigation optimale et sécurisée. Pour atteindre cet objectif, une grande variété de systèmes embarqués génèrent différentes informations sur la navigation et l'état des composants, ce qui permet le contrôle et le monitoring à distance. Du fait de leur importance et de leur informatisation, les systèmes navals sont devenus une cible privilégiée des pirates informatiques. Par ailleurs, la mer est un environnement rude et incertain qui peut produire des dysfonctionnements. En conséquence, la prise de décisions basée sur des fausses informations à cause des anomalies, peut être à l'origine de répercussions potentiellement catastrophiques.

Du fait des caractéristiques particulières de ces systèmes, les méthodologies classiques de détection d'anomalies ne peuvent pas être appliquées tel que conçues originalement. Dans cette thèse nous proposons les mesures de qualité comme une potentielle alternative. Une méthodologie adaptée aux systèmes cyber-physiques a été définie pour évaluer la qualité des flux de données générés par les composants de ces systèmes. À partir de ces mesures, une nouvelle approche pour l'analyse de scénarios fonctionnels a été développée. Des niveaux d'acceptation bornent les états de normalité et détectent des mesures aberrantes. Les anomalies examinées par composant permettent de catégoriser les détections et de les associer aux catégories définies par le modèle proposé. L'application des travaux à 13 scénarios créés pour une plate-forme composée par deux cuves et à 11 scénarios pour deux drones aériens a servi à démontrer la pertinence et l'intérêt de ces travaux.

Mots-clés : Qualité de données et de l'information ; Monitoring ; Réseau multi-source ; Système cyber-physique ; Système Naval ; Pyramide DIKW ; Détection d'anomalies ; Catégorisation d'anomalies.

Abstract

Naval systems represent a strategic infrastructure for international commerce and military activity. Their protection is thus an issue of major importance. Naval systems are increasingly computerized in order to perform an optimal and secure navigation. To attain this objective, on board vessel sensor systems provide navigation information to be monitored and controlled from distant computers. Because of their importance and computerization, naval systems have become a target for hackers. Maritime vessels also work in a harsh and uncertain operational environment that produces failures. Navigation decision-making based on wrongly understood anomalies could be potentially catastrophic.

Due to the particular characteristics of naval systems, the existing detection methodologies can't be applied. We propose quality evaluation and analysis as an alternative. The novelty of quality applications on cyber-physical systems shows the need for a general methodology, which is conceived and examined in this dissertation, to evaluate the quality of generated data streams. Identified quality elements allow introducing an original approach to detect malicious acts and failures. It consists of two processing stages : first an evaluation of quality ; followed by the determination of agreement limits, compliant with normal states to identify and categorize anomalies. The study cases of 13 scenarios for a simulator training platform of fuel tanks and 11 scenarios for two aerial drones illustrate the interest and relevance of the obtained results.

Key-words : Data and information quality, Monitoring ; Multi-source network ; Cyber-physical system ; Naval systems ; DIKW pyramid ; Anomaly Detection ; Anomaly categorization.

Résumé

Les systèmes navals représentent une infrastructure stratégique pour le commerce international et les activités militaires. Ces systèmes sont de plus en plus informatisés afin de réaliser une navigation optimale et sécurisée. Pour atteindre cet objectif, une grande variété de systèmes embarqués génèrent différentes informations sur la navigation et l'état des composants, ce qui permet le contrôle et le monitoring à distance. Du fait de leur importance et de leur informatisation, les systèmes navals sont devenus une cible privilégiée des pirates informatiques. Par ailleurs, la mer est un environnement rude et incertain qui peut produire des dysfonctionnements. En conséquence, la prise de décisions basée sur des fausses informations à cause des anomalies, peut être à l'origine de répercussions potentiellement catastrophiques.

Du fait des caractéristiques particulières de ces systèmes, les méthodologies classiques de détection d'anomalies ne peuvent pas être appliquées tel que conçues originalement. Dans cette thèse nous proposons les mesures de qualité comme une potentielle alternative. Une méthodologie adaptée aux systèmes cyber-physiques a été définie pour évaluer la qualité des flux de données générés par les composants de ces systèmes. À partir de ces mesures, une nouvelle approche pour l'analyse de scénarios fonctionnels a été développée. Des niveaux d'acceptation bornent les états de normalité et détectent des mesures aberrantes. Les anomalies examinées par composant permettent de catégoriser les détections et de les associer aux catégories définies par le modèle proposé. L'application des travaux à 13 scénarios créés pour une plate-forme composée par deux cuves et à 11 scénarios pour deux drones aériens a servi à démontrer la pertinence et l'intérêt de ces travaux.

Mots-clés : Qualité de données et de l'information, Monitoring, Réseau multi-source, Système cyberphysique, Système naval, Pyramide DIKW, Détection d'anomalies, Catégorisation d'anomalies

Abstract

Naval systems represent a strategic infrastructure for international commerce and military activity. Their protection is thus an issue of major importance. Naval systems are increasingly computerized in order to perform an optimal and secure navigation. To attain this objective, on board vessel sensor systems provide navigation information to be monitored and controlled from distant computers. Because of their importance and computerization, naval systems have become a target for hackers. Maritime vessels also work in a harsh and uncertain operational environment that produces failures. Navigation decision-making based on wrongly understood anomalies could be potentially catastrophic.

Due to the particular characteristics of naval systems, the existing detection methodologies can't be applied. We propose quality evaluation and analysis as an alternative. The novelty of quality applications on cyber-physical systems shows the need for a general methodology, which is conceived and examined in this dissertation, to evaluate the quality of generated data streams. Identified quality elements allow introducing an original approach to detect malicious acts and failures. It consists of two processing stages: first an evaluation of quality; followed by the determination of agreement limits, compliant with normal states to identify and categorize anomalies. The study cases of 13 scenarios for a simulator training platform of fuel tanks and 11 scenarios for two aerial drones illustrate the interest and relevance of the obtained results.

Keywords: Data and information quality, Monitoring, Multi-source network, Cyberphysical system, Naval systems, DIKW pyramid, Anomaly detection, Anomaly categorization

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

N° d'ordre : 2017IMTA0056

IMT Atlantique Bretagne-Pays de la Loire - www.imt-atlantique.fr

Campus de Brest
Technopôle Brest-Iroise
CS 83818
29238 Brest Cedex 03
T +33 (0)2 29 00 11 11
F +33 (0)2 29 00 10 00

Campus de Nantes
4, rue Alfred Kastler - La Chantrerie
CS 20722
44307 Nantes Cedex 3
T +33 (0)2 51 85 81 00
F +33 (0)2 51 85 81 99

Campus de Rennes
2, rue de la Châtaigneraie
CS 17607
35576 Cesson Sévigné Cedex
T +33 (0)2 99 12 70 00
F +33 (0)2 99 12 70 08