

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique



Université A. Mira de Béjaïa  
Faculté des Sciences Exactes  
Département d'Informatique

## *MÉMOIRE DE MASTER RECHERCHE*

**En**

**Informatique**

**Option**

*Intelligence Artificielle*

**Thème**

**Fusion de Données dans l'Internet des Objets**

Présenté par :

**M. ASSOUL Nabil**

Devant le jury composé de :

Examineur	Mme ALOUI	Soraya	U. A/Mira Béjaïa.
Examineur	Melle DJEBBAR	Nacera	U. A/Mira Béjaïa.
Encadrant	M. ACHROUFENE	Achour	U. A/Mira Béjaïa.

Année universitaire 2020/2021

# Dédicace

“

*À ma famille, mes enfants, ma femme pour son soutien et  
de m'avoir supporté durant ces trois dernières années,*

*À mes parents, mon père qui n'a cessé de me soutenir pour  
continuer mes études ,*

*À mes ami(e)s, mes collègues de travail surtout Madjid  
merci d'avoir pris sur toi afin que je puisse étudier en  
toute tranquillité,*

*À tous mes professeurs de l'université*

*Merci.*

”

**- Nabil**

# Remerciements

Tout d'abord, je remercie Allah le tout puissant de m'avoir donné le courage et la patience nécessaires à mener ce travail à son terme.

Je tiens à remercier tout particulièrement mon encadrant **M. ACHROUFENE Achour**, pour l'aide qu'il m'a apporté, pour sa patience et son encouragement. Son aide m'a été très précieuse pour structurer ce travail et pour améliorer la qualité des différents chapitres. Sans oublier les membres de jury **Mme ALOUI Soraya** et **Melle BJEBBAR Nacera** d'avoir accepté de juger ce modeste travail.

Je tiens aussi à adresser mes plus sincères remerciements à **M. BOUSSOUFA Abdelmadjid**, collègue de travail sans son sacrifice ce travail n'aurait jamais vu le jour.

Un très grand remerciement et une très grande reconnaissance sont destinés à **M. BENIDIR Ahmed, IHADDADEN Bilal** de m'avoir aidé.

Je remercie également tout les étudiants de ma spécialité, mes enseignants ainsi que tous les gens qui m'ont soutenus de près ou de loin.

Je désire remercier également **M. KHENOUSS Lachemi**, **Mme TAHAKOURTH**, **M. AMROUNE Kamel**, et **M. MAOULOUD Omar** pour leurs encouragements.

Je souhaite aussi remercier l'équipe pédagogique et administrative de la faculté des sciences exactes et du département Informatique pour leurs efforts dans le but de nous offrir une excellente formation.

Pour finir, je souhaite remercier toute personne ayant contribué de près ou de loin à la réalisation de ce travail.

---

# Résumé

Avec l'émergence de l'Internet des objets (IoT), les différents dispositifs intelligents installés sur un bon nombre d'objets manipulés au quotidien nous offrent l'opportunité de collecter, traiter et communiquer des données en temps réel, et ainsi comprendre des indicateurs environnementaux, économiques, ... pour pouvoir agir en fonction des décisions qui en résultent. L'un des défis majeurs de l'IoT est la transmission et le traitement des masses volumineuse de données générées par les objets connectés. Pour répondre à ce défi, la fusion de données est généralement la solution adoptée puisqu'elle permet de réduire la masse de données transmise et l'énergie consommée dans le réseau sans autant perdre en la qualité de la prise de décision. Ce travail consiste à étudier quelques méthodes de fusion de données et de proposer une solution qui permet de choisir la méthode la plus appropriée pour une application IoT donnée. Cette solution est basée sur un ensemble de paramètres, certains sont relatifs à l'application en question comme la nature des données à fusionner et le temps requis, et d'autres dépendent du réseau tels que la densité et la capacité des objets connectés.

---

**Mots clés :** Internet des Objets, Fusion de données, Théories des probabilités, Théorie de Dempster-Shafer, Théorie des sous-ensembles flous.

---

# Abstract

With the emergence of the Internet of Things (IoT), the various intelligent devices installed on a large number of objects used in everyday life offer us the opportunity to collect, process and communicate data in real time, and thus to understand environmental and economic indicators, ... in order to be able to act according to the resulting decisions. One of the major challenges of the IoT is the transmission and processing of the massive amounts of data generated by connected objects. To meet this challenge, data fusion is generally the solution adopted since it allows to reduce the mass of data transmitted and the energy consumed in the network without losing in the quality of the decision making. This work consists in studying some data fusion methods and proposing a solution that allows to choose the most appropriate method for a given IoT application. This solution is based on a set of parameters, some of which are related to the application in question such as the nature of the data to be fused and the time required, and others which depend on the IoT network such as the density and capacity of the connected objects.

---

**Keywords :** Internet Of Things, Data Fusion. Probability Theory, Demspter-Shafer Theory, Fuzzy sets Theory.

---

# Table des matières

Dédicace . . . . .	i
Remerciements . . . . .	ii
Introduction générale . . . . .	1
<b>1 État de l'art sur l'IoT . . . . .</b>	
1.1 Introduction . . . . .	4
1.2 Définition . . . . .	4
1.3 Caractéristiques d'un objet . . . . .	4
1.4 Classification des objets . . . . .	4
1.5 Evolution d'Internet et son impact dans le monde . . . . .	5
1.6 Applications de l'IoT . . . . .	6
1.7 Utilité de l'internet des objets . . . . .	8
1.8 Inconvénients de l'IoT . . . . .	9
1.8.1 L'alimentation des capteurs . . . . .	9
1.8.2 Confidentialité des utilisateurs et des données . . . . .	9
1.9 Iot et fusion de données . . . . .	9
1.10 Conclusion . . . . .	11
<b>2 Fusion de Données . . . . .</b>	
2.1 Introduction . . . . .	13
2.2 Quelques définitions . . . . .	13
2.3 Architecture de la fusion de données . . . . .	13
2.3.1 Modélisation . . . . .	14
2.3.2 Estimation . . . . .	14
2.3.3 Combinaison . . . . .	14
2.3.4 Décision . . . . .	14
2.4 Type de données fusionnées . . . . .	15
2.5 Imperfections de l'information . . . . .	15
2.5.1 Incertitude . . . . .	15
2.5.2 Imprécision . . . . .	16
2.5.3 Incomplétude . . . . .	16

## Table des matières

---

2.5.4	Redondance . . . . .	16
2.5.5	Complémentarité . . . . .	16
2.5.6	Ambiguïté . . . . .	16
2.5.7	Conflit . . . . .	17
2.6	Méthodes utilisées dans la fusion de données . . . . .	17
2.6.1	Approches probabilistes . . . . .	17
2.6.1.1	Petit rappel sur les probabilités . . . . .	17
2.6.1.2	Probabilité . . . . .	18
2.6.1.3	Quelques probabilités particulières . . . . .	19
2.6.1.4	Règle de BAYES . . . . .	19
2.6.1.5	Décision . . . . .	19
2.6.2	Théorie des fonctions de croyance . . . . .	20
2.6.2.1	Définition . . . . .	20
2.6.2.2	Fonction de masse . . . . .	20
2.6.2.3	Elément focal . . . . .	21
2.6.2.4	Fonction de crédibilité . . . . .	21
2.6.2.5	Fonction de plausibilité . . . . .	21
2.6.2.6	Fonction d'affaiblissement . . . . .	22
2.6.2.7	Combinaison . . . . .	23
2.6.2.8	Combinaison conjonctive . . . . .	23
2.6.2.9	Combinaison disjonctive . . . . .	24
2.6.2.10	Combinaison hybride . . . . .	24
2.6.2.11	Prise de décision . . . . .	24
2.6.3	Théorie des sous-ensembles flous . . . . .	25
2.6.3.1	Introduction . . . . .	25
2.6.3.2	Sous-ensembles flous . . . . .	25
2.6.3.3	Définition . . . . .	26
2.6.3.4	Description de sous-ensembles flous . . . . .	26
2.6.3.5	Opérations sur les ensembles flou . . . . .	27
2.6.3.6	Théorie des possibilités . . . . .	28
2.7	Défis de la fusion de données . . . . .	30
2.8	Quelques objectifs de la fusion de données dans l'IoT . . . . .	30
2.9	Conclusion . . . . .	31
<b>3</b>	<b>Etude comparative et sélection d'une méthode de fusion . . . . .</b>	
3.1	Introduction . . . . .	33
3.2	Méthodes probabiliste et bayésienne . . . . .	33
3.2.1	Quelques applications . . . . .	33
3.2.2	Limites . . . . .	33

## Table des matières

---

3.3	Théorie des croyances . . . . .	34
3.3.1	Quelques applications . . . . .	34
3.3.2	Limites . . . . .	34
3.4	Théorie des possibilités . . . . .	34
3.5	Etude comparative des méthodes de fusion . . . . .	35
3.6	Proposition de critères de sélection d'une méthode de fusion pour une application IoT . . . . .	37
3.7	Application d'aide au choix de la méthode de fusion . . . . .	39
3.8	Conclusion . . . . .	44
	<b>Conclusion générale et perspectives . . . . .</b>	<b>45</b>

# Table des figures

1.1	Mode connections des IoT ( <b>4</b> ). . . . .	5
1.2	L'évolution des Objets connectés ( <b>3</b> ). . . . .	5
1.3	Application des IoT ( <b>41</b> ). . . . .	6
1.4	L'IoT dans le domaine du transport ( <b>4</b> ). . . . .	7
1.5	L'IoT et le Corona virus ( <b>4</b> ). . . . .	7
1.6	L'IoT domotique ( <b>4</b> ). . . . .	8
1.7	L'IoT dans le domaine de l'agriculture ( <b>3</b> ). . . . .	8
1.8	Gestion des données pour les systèmes de santé sensibles au contexte ( <b>34</b> ). . . . .	10
2.1	Architecture d'une fusion de donnée ( <b>13</b> ). . . . .	14
2.2	Représentation des fonctions de croyance et de plausibilité ( <b>26</b> ). . . . .	22
2.3	Fonction caractéristique de l'ensemble A et fonction d'appartenance de l'ensemble B . . .	26
2.4	Caractéristiques d'un ensemble flou . . . . .	27
3.1	différent niveaux de fusion [34] . . . . .	39
3.2	Le rapport entre le nombre de noeuds et le temps de réponse pour les réseaux bayésiens [42] . . . . .	40
3.3	Le rapport entre le nombre de noeuds et le temps de réponse pour Dempster-Shafer [39]	40
3.4	Test de réception de données selon la distance entre l'émetteur et le récepteur [41] . . .	41
3.5	Page d'accueil de l'application de choix des méthodes. . . . .	42
3.6	Choix d'une méthode dans le cas de fusion locale. . . . .	42
3.7	Méthode de fusion proposée dans le cas de la classification. . . . .	43
3.8	Choix pour la densité du réseau . . . . .	43
3.9	Choix selon la capacité des capteurs . . . . .	44



# Liste des tableaux

2.1	Opérateurs de combinaison. . . . .	28
3.1	Analyse comparative des techniques de fusion [39]. . . . .	35
3.2	Comparaison entre les méthodes de fusion [40]. . . . .	36
3.3	Comparaison entre les mécanismes d'inférence RB et DS [40]. . . . .	37

# Introduction générale

## Contexte

Internet connaît aujourd'hui une extension inédite avec le développement des objets connectés. Jusqu'alors, internet se voyait comme la capacité des hommes de communiquer à tout moment et en tout lieu ; avec l'avènement des objets connectés, le monde physique peut désormais interagir, que ce soit pour des relations de hommes à hommes, d'hommes à objets ou d'objets à objets et c'est ce qui est appelé communément « l'internet des objets » (IoT, Internet of Things).

L'extension exponentielle du nombre d'objet connectés génère des masses importantes de données circulant sur le réseau de l'IoT. Dès lors, la collecte, le traitement et la transmission de ces données soulèvent de nouveaux défis et ouvrent des pistes de recherche pour répondre aux besoins de l'IoT. L'une des piste les plus explorée dans sens est l'approche de la fusion des données. En effet, cette approche permet de réduire la quantité de données à transmettre et la consommation de l'énergie. Cependant, dans la littérature plusieurs méthodes de fusion sont proposées, en conséquence comment trouver la méthodes de fusion la plus appropriée pour le domaine d'intérêt ?

Notre travail s'inscrit dans cette dynamique où nous allons étudier les différentes méthodes de fusion proposées et tenter d'élaborer un processus d'aide au choix de la méthode de fusion répondant au contraintes du réseau et de l'application en question.

## Problématique

Dans le contexte où plusieurs objets sont connectés collectent et transmettent des informations, nous trouvons devant le dilemme de choisir ou de prendre la bonne information ou la décision la plus adéquate selon les données reçues. La précision de l'information fusionnée dépend essentiellement de la méthode utilisée. Par contre, dans l'IoT, nous somme devant un nombre de contraintes à prendre en considération à savoir la gestion et le traitement du nombre important de données générées, la manière dont nous devons transmettre ces données, la gestion du nombre d'objets connectés, les temps de traitement que ce soit issus de la fusion ou celui impartit des application de l'IOT, et enfin prendre en compte

tout ce qui en relation avec les contraintes matériels telles que les types des réseaux, la capacité des objets connectés en traitement et en mémoire.

## Objectifs

L'objectif visé dans ce travail est de proposer une solution permettant de rendre les systèmes basés sur l'IoT plus performants en utilisant l'approche de fusion de données. Pour ce faire, nous devons d'abord comprendre comment les objets connectés interagissent, collectent et interprètent les données. Ensuite, choisir la méthode de fusion la plus adéquate selon la nature de l'information, le temps de fusion et la capacité des objet, ce qui va avoir un impact en gain de temps de traitement et avoir l'information la plus précise ou exacte possible après la fusion.

## Organisation du mémoire

Ce mémoire est organisé en trois chapitres :

Dans le premier chapitre "**État de l'art sur l'IOT**", nous allons présenter en détails la technologie IoT, le principe de son utilisation et surtout l'impact qu'elle a et va avoir sur notre mode de vie. De plus, nous exposerons quelques applications existantes et d'autre envisageables pour l'IoT, tout en évoquant les avantages et inconvénients auxquels est confrontée cette nouvelle technologie.

Le deuxième chapitre "**Fusion de données**" portera sur le processus de la fusion de données et les différentes techniques et théories utilisées dans la fusion. Nous allons présenter les plus importantes et les plus utilisées qui sont les méthodes basées sur les probabilités, sur la théorie des croyances et les sous-ensembles flous.

Le troisième chapitre "**Comparaison et choix de méthode**", présente une étude comparative des méthodes de fusion détaillées dans le chapitre précédent, leurs limites et leurs avantages. Il se termine par la proposition d'une solution d'aide au choix d'une méthode de fusion et son implantation.

# Chapitre 1

## État de l'art sur l'IoT

### 1.1 Introduction

La technologie de l'Internet des Objets ou Internet of Things (IoT) est l'extension du réseau Internet à des choses/objets et à des lieux dans le monde physique. Les objets deviennent alors inter-connectés et mis en réseau, tel que les montres connectées, bracelets connectés ou encore équipements sportifs connectés...,etc

L'internet des objets est en pleine croissance bénéficiant de la création du Cloud Computing et de son autonomie, de ce fait il peut être appliqué dans divers domaines.

Dans ce chapitre nous allons voir c'est quoi un objet connecté, l'évolution des IoT, les types d'objets, leurs domaines d'applications et leurs inconvénients.

### 1.2 Définition

On peut dire que l'IoT est un procédé interconnectant des objets physiques où virtuels leurs permettant l'émission et la réception d'informations. En réalité sa définition peut être vu d'un ordre conceptuel et technique.

Du point de vue conceptuel, il caractérise des objets physiques connectés ayant leurs propre identité numérique et sont capables de communiquer les uns avec les autres.

Du point de vue technique, elle consiste à son identification numérique (adresse IP, protocoles,...) et ce via les différents réseaux de communication wifi, bluetooth,...etc

### 1.3 Caractéristiques d'un objet

Un objet connecté est caractérisé par par son identité numérique ie : son adresse sur le réseau où il s'est connecté, il est aussi caractérisé par par la nature de sa conception elle même, on veut dire par ça sa capacité de traitement d'informations, le type d'information dont il a la charge d'observer (sonore, visuelles,...), et son autonomie en consommation d'énergie.

### 1.4 Classification des objets

Nous commençons par la figure-1.1 qui explique la mise en service des IoT, qui ne sont ni plus ni moins que des équipements connectés à un réseau de télécommunications.

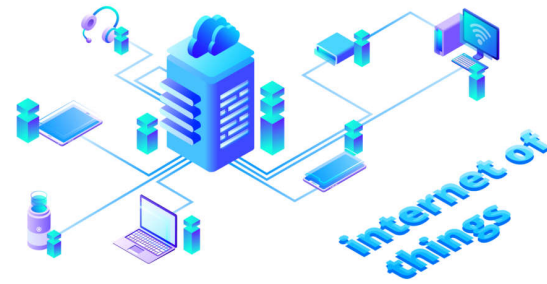


FIG. 1.1 : Mode connections des IoT (4).

### Type d'objets

- **Electroniques** : comme les véhicules connectés en 4G pour optimiser les performances.
- **Electriques** : Tout ce qui est de la domotique, allumage à distance etc...
- **Non électriques** : vêtements, équipements ...
- **capteurs environnementaux** : chaleur, proximité...

## 1.5 Evolution d'Internet et son impact dans le monde

En 2003, la population mondiale a frôlé les 6 milliards d'individus et un demi-milliard d'appareils connectés à Internet. L'idée de l'Internet des objets est apparue en 2009, booster par l'apparition des Smart-phones, le nombre d'appareils connectés à Internet a atteint 12,5 milliards en 2010, alors que la population mondiale était de 6,8 milliards [3].

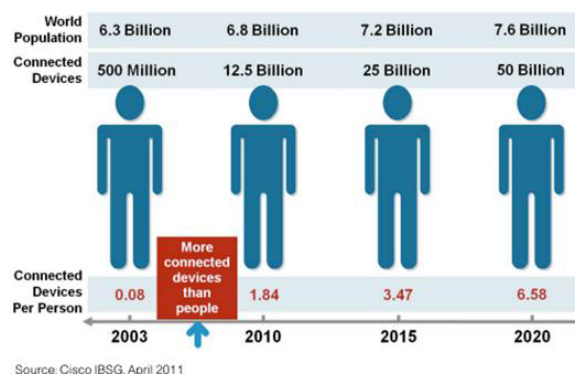


FIG. 1.2 : L'évolution des Objets connectés (3).

Aujourd'hui l'IoT prend de l'ampleur et en ce qui concerne l'avenir, Les experts estiment que 50 milliards d'appareils seront connectés d'ici fin 2021 [3], ces estimations ne prennent pas en considération l'évolution rapide d'Internet ni des avancées technologiques, mais uniquement les faits de l'heure actuelle.

Le nombre de capteurs connectés à Internet pourrait augmenter de plusieurs millions, voire de plusieurs milliards du fait que tout ce qui existe se connecte (Animaux, lampes, maisons, personnes, équipements, voitures,...).

### 1.6 Applications de l'IoT

L'IoT est utilisé dans divers domaines, de nos jours plusieurs secteurs utilisent des dispositifs et des capteurs dédiés aux collectes d'informations, ou à la gestion et le contrôle de bâtiments commerciaux où résidentiels de sécurité où encore d'éclairage, nous allons énumérer quelques types d'applications à titre d'exemple.



FIG. 1.3 : Application des IoT (41).

- **Les transports :** Le domaine des transports est parmi les leaders dans l'exploitation de la technologie de l'IoT, en effet depuis son apparition les transports ont beaucoup bénéficié de cette technologie, on observe une augmentation conséquente d'année en année du nombre d'objets intégrés à ce domaine que ce soit dans l'automobile ou la gestion de la circulation routière et aérienne.



FIG. 1.4 : L'IoT dans le domaine du transport (4).

- **La santé** :Par exemple, (Porteuse Digital Health) qui est le premier médicament connecté sur le marché grâce à un capteur directement intégré dans l'être humain qui permet après ça le suivi des patients à distance [2].



FIG. 1.5 : L'IoT et le Corona virus (4).

Plusieurs autres dispositifs médicaux sont disponibles, on peut citer par exemple le dernier utilisé lors de la pandémie du COVID qui a permis de localiser et de cibler toutes les personnes ayant été en contact avec une autre personne infectée via leurs smart-phone.

- **La domotique** :La domotique regroupe l'ensemble des technologies permettant l'automatisation des équipements d'un habitat. Elle vise à apporter un contrôle total sur tous les équipements ainsi optimiser leur gestion pour un gain en matière de consommation d'énergie et à aider à mieux sécuriser les bâtiments.





FIG. 1.6 : L'IoT domotique (4).

- **Agriculture** : L'agriculture intelligente a pour objet de renforcer la capacité des systèmes agricoles, elle contribue à la sécurité alimentaire en intégrant des nouvelles technologies, telles que l'imagerie satellite où les systèmes de positionnement par GPS qui vont s'occuper de récolter les informations utiles sur l'état du sol, taux d'humidité, taux des sels minéraux, etc, et envoyer ces informations au fermier pour prendre les mesures nécessaires garantissant la bonne production.



FIG. 1.7 : L'IoT dans le domaine de l'agriculture (3).

## 1.7 Utilité de l'internet des objets

L'explosion de l'IoT a fortement impacté les domaines technologique de la vie en général. C'est une technologie axée sur la connectivité des objets en vue de faciliter leurs utilisations. Elle permet d'effectuer un contrôle à distance à travers des réseaux sans fil. Selon l'union internationale des télécommunications, il s'agit d'une infrastructure d'ordre mondial qui permet de jouir de services technologiques évolués [6].

Cette technologie a impactée la quasi totalité des domaines de la vie quotidienne des êtres humains, le nombre d'objets connectés à Internet a dépassé celui des personnes, cela a offert une infinité d'opportunités en matière de création d'applications dans les domaines de l'automatisation, de la détection et de la communication.

### 1.8 Inconvénients de l'IoT

L'IoT facilite grandement le quotidien de nombreuses personnes, que ce soit dans un cadre particulier ou professionnel. Toutefois, à trop se reposer sur cette technologie révolutionnaire, si un problème de performance venait à l'impacter cela aurait pour conséquence une baisse sur l'activité et le chiffre d'affaires. Nous présentons ci-dessous quelques inconvénients de l'internet des objets.

#### 1.8.1 L'alimentation des capteurs

Pour que l'IoT puisse démontrer tout son potentiel, les capteurs devront être auto-suffisants. Sinon, on devait changer les batteries de milliards d'appareils déployés aux quatre coins de la planète et même dans l'espace, cela serait évidemment impossible, nous devons donc trouver un moyen pour générer de l'électricité en puisant dans l'environnement, par exemple en utilisant les vibrations, la lumière et les courants d'air. D'ailleurs, des scientifiques ont annoncé la création d'un nano générateur, il s'agit d'une puce flexible capable de générer de l'électricité à partir de mouvements corporels tels qu'un pincement de doigt.

#### 1.8.2 Confidentialité des utilisateurs et des données

Le piratage informatique sera un problème majeur, l'IoT devrait croître de plus de 12 milliards d'appareils en 2016 et 50 billions d'ici fin 2021 [4]. Chaque appareil est un point d'accès potentiel pour une attaque du réseau par les hackers. Dans une enquête faite par l'organisation Forester, partout dans le monde, 47% des organisations industrielles qui utilisent ou envisagent d'utiliser l'IoT avait précédemment connu des violations de la sécurité dans leurs applications industrielles [4].

### 1.9 Iot et fusion de données

Dans les environnements omniprésents de l'Internet des objets, un volume élevé de données hétérogènes est produit à partir de différents appareils dans un laps de temps rapide. Dans toutes les applications IoT, la qualité de l'information joue un rôle important dans la prise de décision, ces données peuvent être conflictuelles, incertaines ou redondantes, la fusion de ces données devient alors une nécessité absolue afin de minimiser le risque d'erreurs et la bonne prise de décision. Plusieurs études sont faites sur ce domaine afin de définir la ou les meilleures méthodes de fusion, nous citerons à titre d'exemple la recherche de Sahar Boulkaboul et Djamel Djenouri [33] qui consiste en l'amélioration de la méthode de Dempster-Shafer (que nous verrons dans le chapitre II) et ont proposé une nouvelle méthode intitulée **DFIOT** (Data Fusion Internet Of Things) qui améliore le gain d'énergie et gère au mieux la fusion de données

conflictuelles et ce en prenant compte de la densité du réseau le flux de données et améliore même les temps de réponses.

Dans le domaine médical il est bon de citer les travaux de Zartasha Baloutche et al. [34], leur concept consiste à donner une nouvelle perception de la manière dont on prépare les données pour être fusionnées et ceci en prenant compte d'autres aspects hormis les méthodes de fusions tel que Privation de capteur : en cas de défaillance d'un noeud, cette observation spécifique sera perdue. Pour cela ils ont proposé un schéma de fusion dédié aux applications de la santé tel que défini dans la figure 1.8.

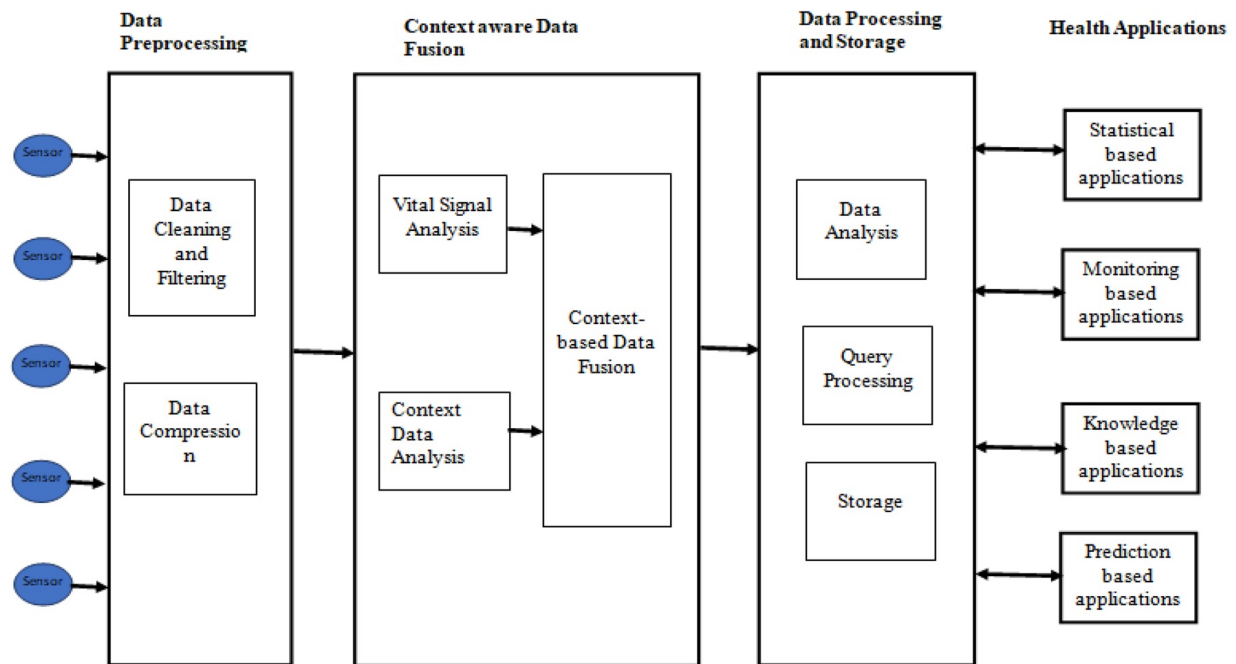


FIG. 1.8 : Gestion des données pour les systèmes de santé sensibles au contexte (34).

Une approche multi-couche est illustrée à la Fig.1.8, qui comprend le pré-traitement des données, la fusion de données contextuelles, ainsi que le traitement et le stockage des données. Les données sont collectées au niveau de la couche physique. Ensuite, au niveau de la couche de pré-traitement des données, le nettoyage et le filtrage nécessaires sont appliqués pour supprimer les valeurs aberrantes (irrégularités ou valeurs inattendues).

Une fois le processus de fusion terminé et le contexte créé, l'étape suivante consiste à traiter ces données. La phase de traitement et de stockage des données comprend trois sous-blocs, à savoir l'analyse des données, le traitement des requêtes et le stockage.

### 1.10 Conclusion

Dans ce chapitre nous avons vu c'est quoi l'IoT, ses domaines d'application, ses avantages et ses inconvénients. Cette technologie a démontrée tout le potentiel qu'elle apporte à la vie humaine et qu'elle peut apporter dans un futur non loin. Nous avons aussi constaté que l'IoT génère beaucoup de données et qu'il faut traiter ces données ce qui nous amène à la fusion de données. Nous avons vu certain travaux qui traitent de cet aspect (méthodes de fusion). Dans le chapitre suivant nous allons présenter différentes méthodes de fusion de données dans l'IoT,

# Chapitre 2

## Fusion de Données

### 2.1 Introduction

Nous avons vu dans le premier chapitre que les objets connectés génèrent un nombre important de données dans le contexte de l'IoT. La gestion de ces données, spécialement la transmission et le traitement ainsi que la fusion, devient une nécessité pour une bonne prise de décision sur le domaine d'intérêt. Pour cela, nous allons présenter, dans ce chapitre, le processus de fusion de données, les différents types d'informations et leurs imperfections, quelques méthodes de fusion utilisées dans l'IoT en l'occurrence les probabilités, les fonctions de croyance et les sous-ensembles flous.

### 2.2 Quelques définitions

La fusion est un processus multi-niveaux et à facettes multiples ayant pour objet la détection automatique, l'association, la corrélation, l'estimation et la combinaison d'informations de sources singulières et plurielles [15].

Selon le groupe de travail européen FUSION (de 1996 à 1999) [16], la fusion de données constitue un cadre formel dans lequel s'expriment les moyens et techniques permettant l'alliance des données provenant de sources diverses. Elle vise à l'obtention d'information de plus grande qualité.

I. Bloch [13] propose une définition plus spécifique, la fusion d'informations consiste à regrouper des informations issues de plusieurs sources d'informations et à exploiter l'information regroupée.

Pour nous, la fusion de données consiste à combiner des informations issues de plusieurs sources afin d'améliorer la prise de décision.

Historiquement, les premières techniques de fusion de données ont été développées dans le cadre d'applications militaires [13]. Dans ce domaine, nous pouvons citer la détection et le suivi d'intrus, la surveillance de champs de bataille où la détection de mines. Depuis lors, l'usage de la fusion de données s'est largement répandu dans d'autres domaines, citons la télédétection, l'aide aux diagnostics médicaux, la robotique où encore l'assistance aux personnes et aux opérateurs [14]. De manière générale, la fusion de données s'est développée pour répondre à un besoin de synthèse et de fiabilité de l'information, car dans nombre d'applications, la quantité des données à manipuler devient de plus en plus importante et difficile à appréhender par des opérateurs humains.

### 2.3 Architecture de la fusion de données

La figure 2.1 montre une architecture typique de la fusion de données. Le processus de fusion peut être résumé en 4 étapes primordiales qui sont les suivantes [13] :

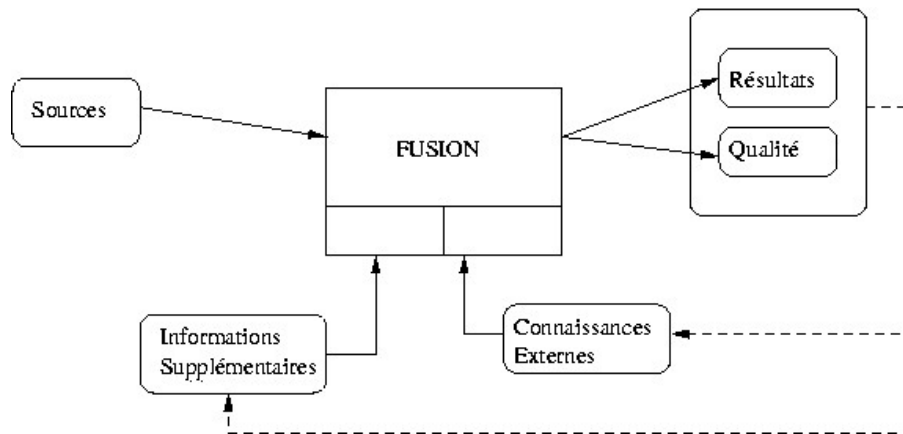


FIG. 2.1 : Architecture d'une fusion de donnée (13).

### 2.3.1 Modélisation

La modélisation est la première étape ; elle consiste au choix du formalisme (i.e. la représentation des informations à fusionner). Cette étape est donc particulièrement déterminante. Elle repose essentiellement sur les connaissances de l'expert, ainsi que, dans une moindre mesure, sur les informations supplémentaires disponibles tel que la fiabilité des informations à fusionner.

### 2.3.2 Estimation

L'estimation est dépendante de la modélisation, elle n'est pas une étape systématique, mais souvent nécessaire pour la plupart des formalismes. Il s'agit par exemple de l'estimation des distributions, si c'est le formalisme choisi. Pour cette étape les informations supplémentaires peuvent aider.

### 2.3.3 Combinaison

L'étape de combinaison est la phase de regroupement des informations. C'est le cœur de la fusion. Il s'agit de choisir une méthode adaptée au formalisme de la modélisation. Ici aussi des informations supplémentaires peuvent guider ce choix. Les méthodes de fusion fournissent un résultat de même nature que les informations combinées. L'information ainsi fournie s'interprète de la même façon que les informations de départ. Les propriétés intéressantes des opérateurs pour la fusion sont l'associativité, la commutativité, ainsi que l'adaptabilité [17].

### 2.3.4 Décision

La dernière étape de la fusion de données est la décision. En effet, d'après les définitions de la fusion, la combinaison est réalisée en vue d'améliorer la prise de décision. Le choix du critère se fait en fonction

du formalisme (i.e. du choix de la modélisation et de la combinaison). Les connaissances externes peuvent ici apporter une aide importante au choix du critère. Classiquement, il s'agit de la minimisation ou de la maximisation d'une fonction issue de la combinaison. C'est cette même fonction, ou une autre qui peut servir au calcul d'un indice de qualité. Cette étape doit donc fournir la "meilleure" décision.

## 2.4 Type de données fusionnées

Il est possible de caractériser les données en deux types voire trois [13]. Le premier type concerne l'information numérique (i.e. qui s'exprime sous forme de nombre). Le deuxième type de données est le symbolique. C'est toutes informations pouvant s'exprimer sous forme de symboles, où de propositions (ex. ce qui est grand n'est pas petit).

I. Bloch [13] propose un type de données hybride (i.e. une information symbolique codée par des nombres). Par exemple, le résultat de différentes classifications peut être modélisé par des nombres correspondant aux classes déterminées, Ces nombres peuvent ensuite être fusionnés pour améliorer la classification. Cependant ce dernier type est discutable, car peut être considéré comme un type numérique.

## 2.5 Imperfections de l'information

Les approches de fusion d'informations reposent sur une bonne modélisation des imperfections de l'information pour pouvoir en tenir compte le mieux possible. En effet, à partir du moment où les informations sont imparfaites trois actions s'offrent à nous :

- **Soit on supprime complètement les informations imparfaites**
- **Soit on tolère ces informations avec leurs imperfections**
- **Soit on modélise les informations en tenant compte des imperfections.**

C'est cette dernière qui est retenue par les approches de fusion de données. Les imperfections des données à fusionner sont nombreuses. Les principales imperfections considérées sont l'incertitude et l'imprécision [18]. A ces deux imperfections peuvent être ajouté l'incomplétude [20], la redondance et la complémentarité [13], le conflit [19], et enfin l'ambiguïté [13].

### 2.5.1 Incertitude

L'incertitude caractérise le degré de conformité à la réalité d'une information. Une information incertaine décrit donc une connaissance partielle de la réalité, le fait que la réalité ne soit pas connue, ne nous permet pas d'affirmer l'information. Cette information est donc incertaine [21].



### 2.5.2 Imprécision

L'imprécision d'une information est caractérisée par le contenu de l'information. Nous distinguons l'imprécision sans erreur (par exemple : il pleut peu) et l'imprécision avec erreur (par exemple : il pleut 10 mm<sup>3</sup>/h alors qu'il y en a 15). La réalité doit donc être connue ou estimée. L'imprécision est donc rencontrée dans de nombreux problèmes ce qui a entraîné de nombreux développements de modèles de l'imprécision [22].

### 2.5.3 Incomplétude

L'incomplétude est le manque d'information apporté par la source. Une des premières mesures de la quantité d'information a été introduite par C.E. Shannon [23], connue sous le nom d'entropie de Shannon. Nous pouvons distinguer deux cas d'incomplétude de l'information :

Soit c'est un défaut d'information (par exemple une source ne fournit pas d'informations alors qu'elle est sensée en donner (cas d'un défaut de transmission)).

Soit l'incomplétude est issue d'un défaut de modélisation de la source ou du problème (par exemple un radar ne permet pas de fournir une image des sous-marins immergés).

### 2.5.4 Redondance

La redondance d'informations est le fait d'avoir plusieurs fois la même information. exemple l'observation d'un même objet par différentes sources peut permettre de situer l'objet avec précision.

### 2.5.5 Complémentarité

On dit que des informations sont complémentaires si elles sont de nature différente et permettent de préciser la réponse du système. La complémentarité des informations d'une source par rapport à une autre permet d'avoir une information plus précise sur l'objet observé.

### 2.5.6 Ambiguïté

L'ambiguïté d'une information est le fait que celle-ci engendre plusieurs interprétations.

### 2.5.7 Conflit

Deux informations ou plus sont en conflit si leurs observations sont contradictoires, les origines du conflit proviennent essentiellement de trois situations [24]

- **Les sources ne sont pas fiables. L'information est erronée et peut conduire à une ambiguïté.**
- **Le cadre de discernement est non exhaustif. L'hypothèse d'un monde fermé est alors fausse**
- **Les sources observent des phénomènes différents, dans ce cas il ne faut pas combiner les informations**

## 2.6 Méthodes utilisées dans la fusion de données

Plusieurs méthodes sont utilisées dans la fusion de données, parmi les méthodes les plus utilisées on peut citer les méthodes probabilistes, les méthodes basées sur la croyance, et les sous-ensembles flous.

### 2.6.1 Approches probabilistes

L'approche bayésienne et plus généralement les approches probabilistes reposent sur un cadre mathématique rigoureux : la théorie des probabilités. En fait, l'approche bayésienne fait plus référence à l'étape de combinaison et en moindre mesure à l'étape d'estimation. Dans le cadre probabiliste, les imperfections de l'information et les informations elles même sont modélisées à partir de distributions de probabilités ou de mesures statistiques à estimer telle que la quantité d'informations de ce fait, cette approche ne permet de modéliser que l'incertitude de l'information.

#### 2.6.1.1 Petit rappel sur les probabilités

L'objet de cette théorie concerne une expérience aléatoire elle est décrite dans un espace dont les éléments sont les résultats possibles de cette expérience :

- **Expérience aléatoire( $\epsilon$ )** : expérience dont le résultat ne peut être prédit ou déterminé avec certitude a priori parmi l'ensemble des possibilités.
- **Univers( $\Omega$ )** : ensemble de tous les résultats possibles de  $\epsilon$
- **Résultat élémentaire ( $\varphi$ )** : résultat possible de  $\epsilon$ , est un élément de  $\Omega$ .

### Exemple

1.  $\epsilon$  : Tirer une boule numérotée.
2.  $\Omega = \{1,2,3,4,5,6,7,8,9\}$
3.  $\varphi = 6$  est une possibilité.

### Evenements

Soit  $P(\Omega)$  l'ensemble des parties (ou sous ensembles) de  $\Omega$  :

- **(A) et un évènement aléatoire** : une partie de  $\Omega$  (sous ensemble de  $\Omega$  ), donc  $P(\Omega)$  est l'ensemble de tous les évènements.
- **( $\varphi$ ) est un évènement élémentaire** : un élément de  $\Omega$ .
- **Réalisation d'un évènement** : A est réalisé lorsque le résultat de l'expérience aléatoire est  $\varphi$  et que  $\varphi \in A$  : A se réalise  $\Leftrightarrow \varphi \in A$

### 2.6.1.2 Probabilité

Une probabilité est une fonction qui permet de calculer la chance que se réalise un évènement de  $P(\Omega)$ .

Soit  $(\Omega, P(\Omega))$  un espace probablisable. Une probabilité sur  $(\Omega, A)$  est une application :  $P(\Omega) \rightarrow [0,1]$  satisfaisant les axiomes fondamentaux suivants :

- Axiome 1:  $0 \leq P(A) \leq 1$ :  $A \in P(\Omega)$
- Axiome 2:  $P(\Omega) = 1$ .
- Axiome 3:  $P(\bigcup_{i \in N} A_i) = \sum_{i \in N} P(A_i)$ ,  $\forall (A_i)_{i \in N}$

On peut dire alors que  $(\Omega, P(\Omega), P)$  est un espace probablisé.

Supposons que l'ensemble des évènements possibles est fini ou dénombrable. On note

$\Omega = \{ \varphi_1, \varphi_2, \dots, \varphi_i, \dots, \varphi_n \}$  l'ensemble des résultats possibles.

On définit la probabilité  $P_i$  de chaque résultat élémentaire  $\varphi_i$  on a alors une suite de probabilités  $(P_1, P_2, P_3, \dots, P_i, \dots, P_n)$  tels que :

$$\begin{cases} 0 \leq p_i \leq 1 \\ \sum_{i=1}^n p_i = 1 \end{cases}$$

La probabilité d'un évènement quelconque  $A$  est donné par

$$P(A) = \sum_{\varphi_i \in A} p_i$$

### 2.6.1.3 Quelques probabilités particulières

#### Probabilité conjointe

Une probabilité conjointe est une probabilité d'une assignation de toutes les variables.

#### Probabilité Marginale

Une probabilité marginale est une probabilité sur un sous-ensemble des variables.

#### Probabilité d'une disjonction

Probabilités de disjonction (« OU ») d'événements revient à calculer la probabilité qu'un événement se réalise ou qu'un autre événement se réalise.

On peut traduire cela par une formule générale suivante pour deux événements A et B :

$$P(A \text{ ou } B) = P(A) + P(B) - P(A \text{ et } B) \quad (2.1)$$

#### Probabilité conditionnelle

Une probabilité conditionnelle est la probabilité qu'un événement A se réalise à condition que l'événement B se réalise. Sa formule générale peut s'écrire de la sorte :

$$P(A | B) = \frac{P(A \cap B)}{P(B)} \quad \text{avec } (P(B) \neq 0) \quad (2.2)$$

### 2.6.1.4 Règle de BAYES

La règle de BAYES se base sur le concept (cause|effet) on l'écrit comme suit :

$$P(\text{Cause} | \text{Effet}) = \frac{P(\text{Effet} | \text{Cause}) P(\text{Cause})}{P(\text{Effet})}$$

On peut déduire que pour deux événements A et B la règle s'écrit de la sorte :

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)}$$

### 2.6.1.5 Décision

Pour la dernière étape de la fusion, la règle la plus employée pour la prise de décision est le maximum à posteriori.

$$p(d_k | S_1, \dots, S_m) = \max_{i \in \{1, \dots, n\}} p(d_i | S_1, \dots, S_m) \quad (2.3)$$

Il est signalé que pour les approches probabilistes, les réseaux bayésiens sont les plus utilisés.

## 2.6.2 Théorie des fonctions de croyance

### 2.6.2.1 Définition

La théorie des fonctions de croyance est issue des travaux de Dempster en 1967 [26], repris par Shafer [27] sous le nom de la théorie de l'évidence, elle porte également le nom de théorie de Dempster-Shafer. Cette méthode repose sur la modélisation de la croyance d'un événement. La modélisation se réalise à partir de fonctions de masse permettant une bonne représentation des connaissances, elle est souvent présentée d'un point de vue probabiliste, mais peut être vue comme un modèle formel de degrés de connaissance. Cette théorie permet particulièrement de bien modéliser l'incertitude, mais aussi l'imprécision. Elle tient compte également des ambiguïtés et des conflits entre les sources.

### 2.6.2.2 Fonction de masse

Considérons l'ensemble appelé cadre de discernement  $\Theta = \{d_1, d_2, \dots, d_n\}$ , ensemble de toutes les décisions possibles - par exemple l'ensemble de toutes les classes envisageables dans une problématique de classification. L'espace des fonctions de masse  $m$  est donné par l'ensemble de toutes les décisions possibles  $d_i$  noté :

$$2^\theta = \{\emptyset, \{d_1\}, \{d_2\}, \{d_1, d_2\}, \{d_3\}, \{d_1, d_3\}, \{d_2, d_3\}, \{d_1, d_2, d_3\}, \dots, \theta\}$$

Une fonction de masse est définie sur  $2^\theta$  à valeurs dans  $[0; 1]$ . Pour une source  $S_j$ , la fonction de masse  $m_j$  vérifie en général par construction :

$$\sum_{A \in 2^\theta} m_j(A) = 1 \quad (2.4)$$

La fonction de masse  $m_j(A)$  caractérise le degré de croyance en la proposition A donnée par la source  $S_j$ .

Une hypothèse souvent faite par commodité est celle d'un monde clos, ou monde fermé, c'est-à-dire que toutes les décisions possibles sont représentées dans  $\theta$ , et dans ce cas :

$$m_j(\emptyset) = 0 \quad (2.5)$$

Sinon si nous envisageons le fait qu'une décision hors de  $\theta$  est possible, c'est-à-dire le monde ouvert, alors :

$$m_j(\emptyset) > 0, \quad (2.6)$$

### 2.6.2.3 Élément focal

Un élément focal est un élément  $A$  de  $2^\theta$  tel que  $m_j(A) > 0$ . La réunion des éléments focaux est appelé le noyau.

### 2.6.2.4 Fonction de crédibilité

La fonction de crédibilité ou de croyance  $\text{Bel}$  (pour belief en anglais) est définie pour tout  $A \in 2^\theta$  comme la somme des masses des hypothèses incluses dans  $A$ , c'est-à-dire qui induisent obligatoirement  $A$ , comme suit :

$$\forall A \in 2^\theta, \text{Bel}(A) = \sum_{B \subseteq A, B \neq \emptyset} m(B) \quad (2.7)$$

Cette fonction mesure toute la croyance sur  $A$  i.e. l'intensité que les informations fournies par la source  $S_j$  soutiennent la proposition  $A$ .

### 2.6.2.5 Fonction de plausibilité

La fonction de plausibilité  $\text{Pls}$  est définie pour tout  $A \in 2^\theta$  comme la somme des masses des hypothèses dont l'intersection avec  $A$  n'est pas nulle, c'est-à-dire qui pourraient induire  $A$  :

$$\forall A \in 2^\theta, \text{Pls}(A) = \sum_{B \cap A \neq \emptyset} m(B) = 1 - \text{Bel}(\bar{A}) \quad (2.8)$$

Cette fonction mesure l'intensité avec laquelle nous trouvons  $A$ , i.e. avec laquelle nous ne doutons pas en  $A$ .

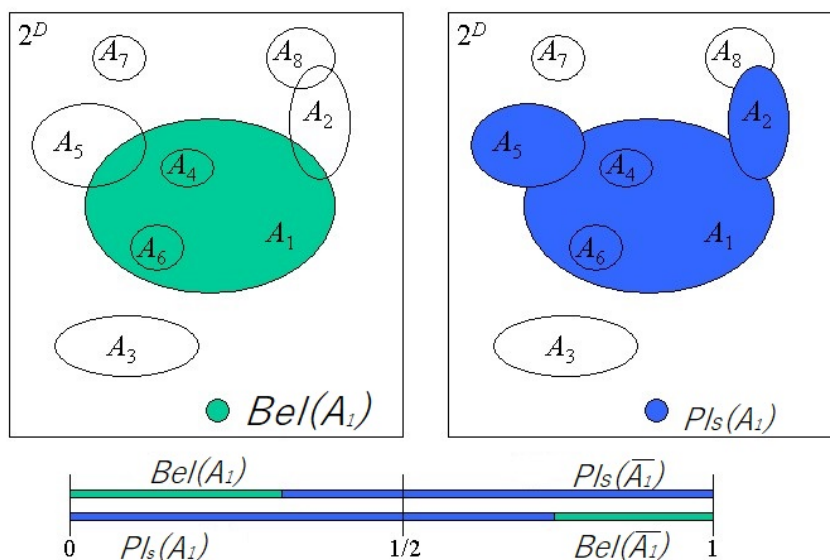


FIG. 2.2 : Représentation des fonctions de croyance et de plausibilité (26).

A partir des définitions des fonctions de croyance et de plausibilité, pour  $\forall A \in 2^\Theta$  : on déduit les propriétés suivantes :

$$\text{Pls}(\Theta) = 1 \quad (2.9)$$

$$\text{Bel}(\Theta) = 1 \quad (2.10)$$

$$\text{Pls}(A) \geq \text{Bel}(A) \quad (2.11)$$

$$\text{Pls}(A) = 1 - \text{Bel}(\bar{A}) \quad (2.12)$$

$$\text{Bel}(A) = 1 - \text{Pls}(\bar{A}) \quad (2.13)$$

### 2.6.2.6 Fonction d'affaiblissement

L'affaiblissement des fonctions de masse, permet de modéliser la fiabilité des sources  $S_j$  en introduisant un coefficient  $\alpha_j$ . Ainsi, nous redéfinissons la fonction de masse pour tout  $A \in 2^\theta, A \neq \theta$  par :

$$\begin{cases} m'_j(A) = \alpha_j m_j(A) \\ m'_j(\theta) = 1 - \alpha_j (1 - m_j(\theta)) \end{cases} \quad (2.14)$$

$\alpha_j \in [0, 1]$  est le coefficient d'affaiblissement de la source  $S_j$ . Dans le cas où  $\alpha_j = 0$ , la source  $S_j$  n'est pas du tout fiable, et dans ce cas toute la masse est affectée à  $\theta$  :  $m_j(\theta) = 1$ , ce qui représente l'ignorance totale.

### 2.6.2.7 Combinaison

Plusieurs modes de combinaison ont été développés dans le cadre de la théorie des croyances. Il existe principalement deux types de combinaison : la combinaison conjonctive et la combinaison disjonctive, qui ont été déclinés en un grand nombre d'opérateurs de combinaison dont des combinaisons hybrides.

### 2.6.2.8 Combinaison conjonctive

Elle combine les fonctions de masse en considérant les intersections des éléments de  $2^\theta$ . Les règles les plus connues sont Dempster [26] et reprise par Shafer [27], la règle de Smets [28] et la règle de Yager [29].

**Règle de Dempster :** Connue également sous le nom de Somme orthogonale [67], cette règle est proposée par Shafer et s'utilise dans le cas de monde fermé. Pour deux fonctions de masse de croyance  $m_1$  et  $m_2$ , et  $B_1, B_2 \in 2^\theta$ , elle s'écrit ( $m_1 \oplus m_2$ ) :

$$\forall A \in 2^\theta \setminus \{\emptyset\}, m_{1 \oplus 2}(A) = m(A) = \frac{1}{1 - \kappa} \sum_{B_1 \cap B_2 = A} m_1(B_1) m_2(B_2) \quad (2.15)$$

$$\kappa = \sum_{B_1 \cap B_2 = \emptyset} m_1(B_1) m_2(B_2) \quad (2.16)$$

Avec  $k$  qui représente la masse de l'ensemble vide ( $\kappa = m(\emptyset)$ ) et donc la masse de  $\emptyset$  devient  $m(\emptyset) = 0$

$k$  représente aussi le degré de conflit entre les fonctions de masse  $m_1$  et  $m_2$ , appelé aussi facteur de normalisation.

**Règle de Smets :** La règle de Smets est utilisée lorsque les sources à combiner sont fiables. Elle attribue la masse conflictuelle  $k$  à l'ensemble vide  $\emptyset$  en supposant que la solution peut ne pas se trouver dans l'ensemble  $\emptyset$  (monde ouvert). Dans ce cas le conflit peut provenir de la non prise en compte d'une ou plusieurs hypothèses dans le cadre de discernement, elle se représente sous la forme :

$$\forall A \in 2^\theta \setminus \{\emptyset\}, m_{Conj}(A) = \sum_{B_1 \cap B_2 = A} m_1(B_1) m_2(B_2) \quad (2.17)$$

En combinant les deux équations (2.15) et (2.17), une autre écriture de la règle de Dempster peut se présenter sous la forme :

$$\forall A \in 2^\theta \setminus \{\emptyset\}, m(A) = \frac{1}{1 - \kappa} m_{Conj}(A) \quad (2.18)$$

**Règle de Yager :** Pour deux fonctions de masse de croyance  $m_1$  et  $m_2$ , et  $B_1, B_2 \in 2^\theta$ , la règle de Yager [29] est définie comme suit :



$$\begin{cases} m_{Yag}(A) = \sum_{B_1 \cap B_2 = A} m_1(B_1) m_2(B_2), & \forall A \in 2^\Theta \setminus \{\emptyset, \Theta\} \\ m_{Yag}(\Theta) = \sum_{B_1 \cap B_2 = \Theta} m_1(B_1) m_2(B_2) + \kappa \end{cases} \quad (2.19)$$

Cette règle de combinaison considère qu'au moins une des sources d'informations à combiner est fiable, donc une solution au problème appartient obligatoirement à l'ensemble  $\Theta$ . En cas de conflit, la masse conflictuelle  $\kappa$  est affecté à  $\Theta$

### 2.6.2.9 Combinaison disjonctive

D'autres approches de combinaison ont été proposées telle que la combinaison disjonctive [30] Elle est donnée non pas en considérant les intersections, mais les unions. Les éléments focaux de la combinaison s'obtiennent alors par une combinaison des unions des fonctions de masse, elle est donnée comme suit pour deux masses de croyance  $m_1$  et  $m_2$ .

$$\forall A \in 2^\Theta, m_{Disj}(A) = \sum_{B_1 \cup B_2 = A} m_1(B_1) m_2(B_2) \quad (2.20)$$

### 2.6.2.10 Combinaison hybride

Pour chercher à conserver les avantages de la combinaison conjonctive et disjonctive, Dubois et Prade [31] ont proposé un compromis : une combinaison mixte. Cette combinaison est donnée pour  $m_1$  et  $m_2$  et  $B_1, B_2 \in 2^\Theta$  par :

$$m_{DP}(A) = \sum_{B_1 \cap B_2 = A} m_1(B_1) m_2(B_2) + \sum_{B_1 \cup B_2 = A, B_1 \cap B_2 = \emptyset} m_1(B_1) m_2(B_2) \quad \forall A \in 2^\Theta \setminus \{\emptyset\} \quad (2.21)$$

### 2.6.2.11 Prise de décision

La prise de décision consiste à sélectionner une hypothèse (simple ou composée) parmi l'ensemble des hypothèses de départ (cadre de discernement), en se basant sur la fonction de masse obtenue lors de la combinaison, on cite les plus connues :

#### Maximum de crédibilité

Le maximum de crédibilité consiste à choisir l'hypothèse avec la plus grande valeur de crédibilité. Pour une décision  $d_k$  qui correspond à l'hypothèse  $H_k$  de  $\Theta$  :

$$\forall H_i \in \Theta, Bel(H_k) = \max Bel(H_i) \quad (2.22)$$

### Maximum de plausibilité

Le maximum de plausibilité est un mode optimiste qui consiste à prendre la valeur maximale de la plausibilité pour les singletons ainsi pour une décision  $d_k$  qui correspond à l'hypothèse  $H_k$  de  $\Theta$  :

$$\forall H_i \in \Theta, \text{Pls}(H_k) = \max \text{Pls}(H_i) \quad (2.23)$$

## 2.6.3 Théorie des sous-ensembles flous

### 2.6.3.1 Introduction

La théorie des ensembles flous est développée par le Professeur Zadeh au milieu des années 60 . Elle permet une modélisation très souple des informations imprécises. La théorie des possibilités introduite par Zadeh en 1978, puis développée en France par Dubois et Prade [32] est dérivée de la théorie des sous-ensembles flous. Cette approche permet de représenter à la fois les imprécisions et les incertitudes de l'information.

### 2.6.3.2 Sous-ensembles flous

Un ensemble classique est composé d'éléments qui satisfont des propriétés précises. Ainsi les éléments qui ne satisfont pas ces propriétés n'appartiennent pas à cet ensemble. Notons  $X$  un ensemble net et  $x, y$  ses éléments. Un sous-ensemble  $A$  de  $X$  peut s'écrire à partir de sa fonction caractéristique :

$$\chi_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A \end{cases} \quad (2.24)$$

**Exemple :** nous voulons vérifier si des éléments d'un sous-ensemble  $A$  appartiennent à l'ensemble  $X$  caractérisé pas par l'âge d'une population à risque comprise entre 40 et 65 ans, la fonction caractéristique est définie comme suit :

$$\chi_A(x) = \begin{cases} 1 & \text{si } 40 \leq x \leq 65 \\ 0 & \text{sinon.} \end{cases} \quad (2.25)$$

Maintenant nous voulons caractériser l'ensemble  $B$  des personnes avoisinant les 50 ans, la proposition avoisinant les 50 ans n'est pas précise.  $B$  pourrait être représenté par une fonction caractéristique, mais faut-il choisir les âges compris entre 40 et 65 ans ou 45 et 55 ans comme étant les âges avoisinant 50 ? La fonction d'appartenance permet de généraliser la fonction caractéristique en tenant compte des âges pas trop éloignées de 50 et sans être trop proches.

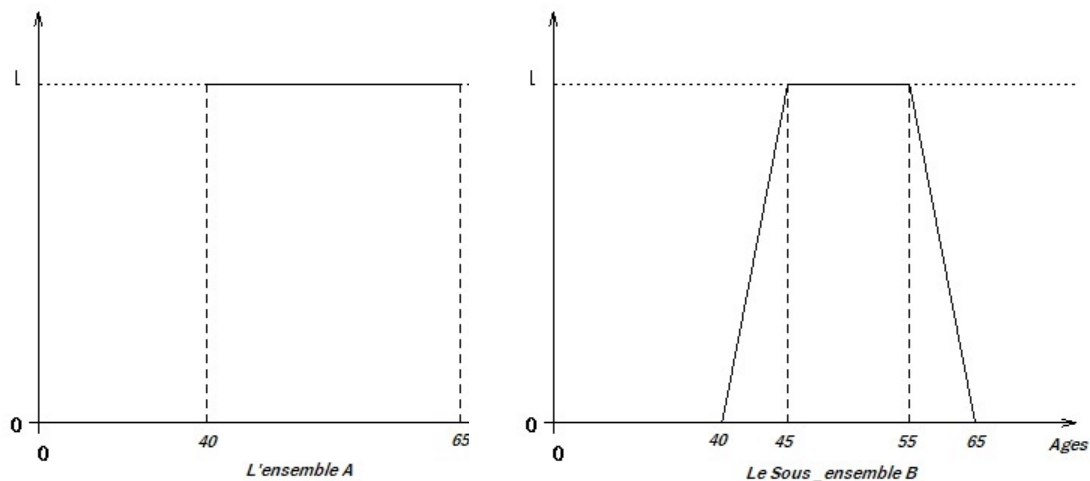


FIG. 2.3 : Fonction caractéristique de l'ensemble A et fonction d'appartenance de l'ensemble B

### 2.6.3.3 Définition

Un sous-ensemble flou F de X est défini par une fonction d'appartenance  $\mu_F$  qui à tout x de X associe une valeur réelle  $\mu_F(x)$  dans  $[0; 1]$  représentant le degré d'appartenance de x au sous-ensemble flou F.

### 2.6.3.4 Description de sous-ensembles flous

**Noyau :** le noyau d'un sous-ensemble flou A de X, noté  $\text{Noy}(A)$  est l'ensemble des éléments appartenant totalement à A :

$$\text{Noy}(A) = \{x \in A : \mu_A(x) = 1\}. \tag{2.26}$$

**Support :** le support d'un sous-ensemble flou A de X, noté  $\text{Supp}(A)$ , est l'ensemble des éléments appartenant au moins un peu à A :

$$\text{Supp}(A) = \{x \in A : \mu_A(x) > 0\} \tag{2.27}$$

**Hauteur :** la hauteur d'un sous-ensemble flou A, notée  $h(A)$  est la valeur maximale atteinte sur le support de A :

$$h(A) = \sup_{x \in A} \mu_A(x) \tag{2.28}$$

**Cardinal :** Le cardinal d'un sous-ensemble flou A de X fini, est le nombre d'éléments qui appartiennent à A pondéré par leur degré d'appartenance.

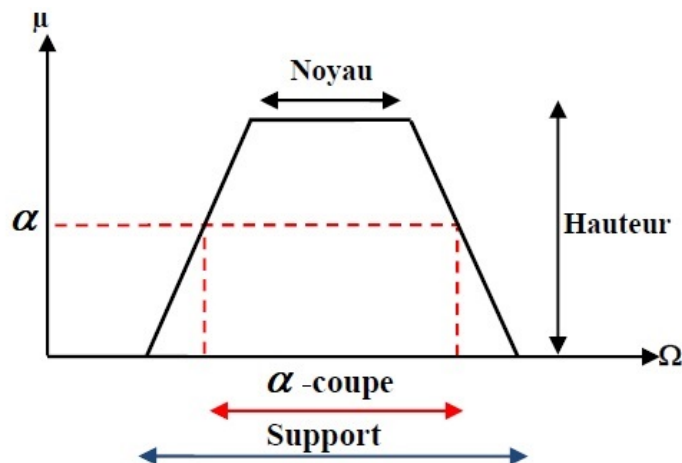


FIG. 2.4 : Caractéristiques d'un ensemble flou

$$|A| = \sum_{x \in \mathcal{X}} \mu_A(x) \quad (2.29)$$

**α-coupe :** Une α coupe de F a une valeur  $\alpha \in [0,1]$  et le sous-ensemble  $F_\alpha$  sont les éléments dont leurs degrés d'appartenance est supérieur ou égal à  $\alpha$ .

$$F_\alpha = \{x \in X / \mu_F(x) \geq \alpha\} \quad (2.30)$$

### 2.6.3.5 Opérations sur les ensembles flou

Les opérations sur les sous-ensembles nets peuvent être généralisées aux sous-ensembles flous.

**Egalité** Deux sous-ensembles flous A et B de X sont dits égaux s'ils ont des fonctions d'appartenance égales en tout élément de X :

$$A = B \iff \forall x \in \mathcal{X} \mu_A(x) = \mu_B(x) \quad (2.31)$$

**Inclusion** Soit deux sous-ensembles flous A et B de X. Si tout élément x de X appartient toujours moins à A qu'à B, alors A est inclus dans B ( $A \subseteq B$ )

$$A \subseteq B \iff \forall x \in \mathcal{X} \mu_A(x) \leq \mu_B(x) \quad (2.32)$$

**Complément** Le complément d'un sous-ensemble flou A de X, noté  $\bar{A}$  est défini par la fonction d'appartenance donnée pour tout x de X par :

$$\mu_{\bar{A}}(x) = 1 - \mu_A(x) \quad (2.33)$$

### Intersection

L'intersection de deux sous-ensembles flous A et B de X est le sous-ensemble flou formé des éléments de X affectés du plus petit des degrés d'appartenance de A et de B. La fonction d'appartenance de  $A \cap B$  est donnée pour tout x de X par :

$$\mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x)) \quad (2.34)$$

### Opérateurs de combinaisons

Les opérateurs de combinaisons les plus utilisés sont les t-norme et t-conorme tel que décrit dans le tableau 2.1 ci-dessous

Dénomination	Intersection ET (t-norme)	Réunion OU (t-conorme)	Complément NON
Opérateurs de Zadeh min/max	$\mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x))$	$\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x))$	$\mu_{\bar{A}}(x) = 1 - \mu_A(x)$
Probabiliste prod/prob	$\mu_{A \cap B}(x) = \mu_A(x)\mu_B(x)$	$\mu_{A \cup B}(x) = \mu_A(x) + \mu_B(x) - \mu_A(x)\mu_B(x)$	$\mu_{\bar{A}}(x) = 1 - \mu_A(x)$

TAB. 2.1 : Opérateurs de combinaison.

### 2.6.3.6 Théorie des possibilités

#### Mesure de possibilité

Une mesure de possibilité  $\Pi$  est une fonction qui attribue à chaque évènement A défini sur X, un coefficient dans  $[0, 1]$  évaluant à quel point cet évènement est possible :  $\Pi : A \rightarrow [0, 1]$  telle que :

$$\left\{ \begin{array}{l} \Pi(X) = 1 \\ \Pi(\emptyset) = 0 \\ \forall I \subset \mathbb{N}, \forall A_i \subseteq X, i \in I, \quad \Pi(\cup_{i \in I} A_i) = \sup_{i \in I} \Pi(A_i) \end{array} \right. \quad (2.35)$$

La mesure de possibilité satisfait la relation

$$\max(\Pi(A), \Pi(\bar{A})) = 1 \quad (2.36)$$

### Mesure de nécessité

Une mesure de nécessité  $N$  est définie pour chaque évènement  $A$  de  $X : \Pi : A \rightarrow [0, 1]$  telle que :

$$\begin{cases} N(X) = 1 \\ N(\emptyset) = 0 \\ \forall I \subset \mathbb{N}, \forall A_i \subseteq X, \quad N(\cap_{i \in I} A_i) = \inf N(A_i) \end{cases} \quad (2.37)$$

La relation reliant les mesures de nécessité et de possibilité est donnée par :

$$\forall A \subseteq X, N(A) = 1 - \Pi(\bar{A}) \quad (2.38)$$

La valeur prise par la mesure de nécessité pour un évènement  $A$  indique le degré de certitude avec lequel cet évènement sera réalisé. Plus ce degré est proche de 1, plus il est certain que cet évènement sera réalisé.

### Distributions de possibilité

Une distribution de possibilité est une fonction  $\pi$  de  $\mathcal{X}$  dans  $[0, 1]$  vérifiant la propriété de normalisation :

$$\sup_{x \in \mathcal{X}} \pi(x) = 1 \quad (2.39)$$

Cette condition signifie qu'au moins un élément de  $X$  est totalement possible et correspond à une hypothèse de monde fermé. A partir d'une distribution de possibilité  $\pi$ , les mesures de possibilité et de nécessité sont définies respectivement par :

$$\forall A \subseteq X, \Pi(A) = \sup_{x \in A} \pi(x) \quad (2.40)$$

$$\forall A \subseteq X, N(A) = 1 - \sup_{x \notin A} \pi(x) = \inf_{x \in A} (1 - \pi(x)) \quad (2.41)$$

### Opérateurs de combinaison pour les conflits et les fiabilités

#### Minimum ou maximum pondérés

Introduits par Dubois et Prade dans le cadre de la théorie des possibilités [32], ces opérateurs permettent de tenir compte d'une modélisation de la fiabilité des sources. Ainsi, nous attribuons un degré de fiabilité  $\lambda_j \in [0, 1]$  aux sources  $S_j$  (qui vaut 1 si la source est complètement fiable et 0 si elle n'est pas du tout fiable).

L'opérateur de combinaison maximum pondéré est défini par :

$$\pi(x) = \max_{j=1, \dots, m} \min(\pi^j(x), \lambda_j) \quad (2.42)$$

L'opérateur de combinaison minimum pondéré est définie par :

$$\pi(x) = \min_{j=1, \dots, m} \max(\pi^j(x), 1 - \lambda_j) \quad (2.43)$$

Une fois les informations issues des différentes sources combinées, la décision se fait à partir du maximum des degrés d'appartenance. Ainsi, nous choisirons une décision  $d_k$  de la manière suivante :

$$\mu_k(x) = \max \{\mu_i(x), i = 1, \dots, n\} \quad (2.44)$$

## 2.7 Défis de la fusion de données

La fusion de données a de nombreux défis à relever, qui sont explorés dans diverses publications [33], [36], [40]. Certains d'entre eux sont énumérés ci-dessous :

- Imperfection des données : les données du capteur sont parfois imprécises, ce comportement est tristement célèbre dans les réseaux de capteurs sans fil. L'imperfection doit être traitée efficacement avec l'utilisation d'algorithmes de fusion de données.
- Caractéristiques triviales : dans l'environnement IoT, les applications peuvent être constituées de plusieurs centaines et milliers de capteurs captant différents paramètres. Ces valeurs détectées dans de grandes configurations telles que les villes intelligentes et les usines industrielles se composent de données triviales et non triviales. Le traitement de données triviales peut affecter la précision de la fusion des données. Ainsi, les fonctionnalités les plus pertinentes doivent être sélectionnées avant la fusion des données.
- Processus dynamiquement itératif : la fusion de données n'est pas un processus statique par nature ; cependant, l'itération dynamique nécessite un raffinement régulier des estimations dans un environnement de fusion. Pas d'algorithme magique : avec le temps, les recherches dans le domaine de la fusion de données ont progressé et les algorithmes hautes performances sont maintenant là. Cependant, il est encore difficile de dire qu'un algorithme de fusion de données parfait existe.

## 2.8 Quelques objectifs de la fusion de données dans l'IoT

La fusion de données dans le paradigme de l'Internet des objets peut jouer un rôle majeur dans son succès pour les raisons suivantes :

- La fusion de données rend les informations plus intelligentes, décisives, sensibles et précises qui proviennent de plusieurs capteurs et sources. Les informations de chaque capteur en soi peuvent ne pas avoir beaucoup de sens.
- Dans l'IoT, un grand défi consiste à fabriquer des capteurs à très faible consommation qui n'ont pas besoin de remplacement de batterie au cours de leur durée de vie. Il est un fait établi que les

capteurs de haute précision peuvent entraîner une consommation d'énergie élevée. Pour gérer ce problème, un ensemble de capteurs à très faible consommation d'énergie peut être utilisé avec une faible précision. En utilisant la fusion de données, des informations très précises seront créées [25].

- La fusion de données peut être utile pour gérer les problèmes de big data de l'IoT, car nous fusionnons les données de nombreux capteurs en informations plus précises et exactes.
- Un autre avantage de la fusion de données est qu'il aide à masquer les informations critiques ou la sémantique qui sont responsables des résultats fusionnés, des exemples en sont les applications militaires, certaines zones médicales critiques et les bâtiments du renseignement.

## 2.9 Conclusion

Dans ce chapitre nous avons définis les types d'informations à fusionner et leurs imperfections et exposé quelques théories de fusion de donnée. Comme nous l'avons constaté toutes ces dernières utilisent le même principe de fusion à savoir l'étape de modélisation ensuite l'estimation, la combinaison et enfin la prise de décision. Nous avons aussi constaté que ces méthodes traitent chacune de certains domaines d'imperfection de l'information et non toutes, ce qui pose un défi au domaine de la fusion de donnée afin de traiter toutes les imperfections des données collectées. Dans le chapitre, suivant nous exposerons quelques résultats obtenus suite à une étude comparative sur certaines méthodes de fusion de données et nous finirons par proposer une solution d'aide au choix d'une méthode de fusion.



## Chapitre 3

# Etude comparative et sélection d'une méthode de fusion

### 3.1 Introduction

Au cours de notre travail nous avons vu que les IoT sont caractérisées par certains éléments dont la fusion de données doit prendre en considération où essayer d'en améliorer les performances tels que la capacité des noeuds ou la densité des réseaux. Dans ce chapitre, nous allons référencer quelques applications des différentes méthodes de fusion afin de tirer leurs limites, ceci va nous permettre de faire une synthèse sur la ou les méthodes de fusion à choisir pour l'IoT. Nous exposerons aussi notre proposition pour le choix des méthodes et nous finirons par donner l'application conçue pour cet objectif.

### 3.2 Méthodes probabiliste et bayésienne

#### 3.2.1 Quelques applications

Nous recensons dans cette partie quelques travaux de manière non exhaustive ayant employé l'approche bayésienne pour la fusion d'informations.

Chauvin [35] étudie les performances du modèle bayésien de fusion pour la segmentation d'images satellites (en 3 classes) avec pour opérateur de combinaison le produit et pour règle de décision le maximum a posteriori. Il est montré que les performances sont en fonction de l'apprentissage (choix de la densité de probabilité conditionnelle et des probabilités a priori) et de la qualité des images, mesurée par un critère fondé sur l'inertie intra-classes de Fisher.

L. Xu et al [36] comparent le classifieur bayésien à la théorie de Dempster- Shafer pour une application en reconnaissance manuscrite. Il est montré que si les matrices de confusion sont bien apprises, l'approche bayésienne est la meilleure. Cependant cette approche est peut fiable et ses performances sont vite dégradées si l'apprentissage est mauvais.

L'approche bayésienne est aussi utilisée dans le traitement et la fusion d'images, le domaine médical n'échappe pas à cette méthode, en effet elle beaucoup utilisé dans ce domaine tels que l'aide a la décision dans la spécification d'une maladie suite à un certain nombres de causes, la détection de maladies sur une image scanner, étude des effets indésirables de médicaments [37].

#### 3.2.2 Limites

Le problème majeur des probabilités est qu'elles représentent essentiellement l'incertitude et très mal l'imprécision, ceci entraîne souvent une confusion des deux notions.

Dans la modélisation probabiliste nous raisonnons sur des singletons qui représentent les différentes décisions. Ces décisions sont exhaustives et exclusives. Ceci entraîne que le monde doit être fermé, ce qui ne correspond pas toujours à la réalité.

Elle constitue aussi un défi majeur à la fusion vu sa consommation élevée en matière de mémoire donc peu pratique sur des capteurs de faible en capacité mémoire.

### 3.3 Théorie des croyances

#### 3.3.1 Quelques applications

La théorie des croyances peut être appliquée dans un grand nombre de situations ce qui a poussé à l'introduction de fonction de croyance dans toute sorte de domaines.

En traitement d'images, la théorie de Dempster-Shafer a souvent été abordée, surtout dans le cas d'images médicales.

La localisation est un problème souvent abordé en fusion d'informations en général. La théorie des croyances a été employée et adaptée pour la localisation de robots, à partir d'images à ultrason, mais aussi pour l'exploitation d'un réseau autoroutier [38].

Reconnaissance de cibles via les données fréquentielles de SONAR et satellite ou à base d'images Radar.

Appliquée dans le domaine météorologique, dans la fusion de données chimiques et bien d'autres domaines.

#### 3.3.2 Limites

Le principal défaut de cette approche est sa complexité à croissance exponentielle avec la taille du cadre de discernement ce qui entraîne une augmentation considérable du temps de réponse et donc de prise de décision.

### 3.4 Théorie des possibilités

A l'instar des autres méthodes, la théorie des possibilités avec les sous-ensembles flous est beaucoup utilisée. On peut citer quelques domaines d'application tel la robotique, le diagnostic médical, la fouille de données et surtout dans le domaine de classification. En effet elle est utilisée beaucoup dans le domaine de classification d'objet sous-marins ou classification d'informations. Le principal hic de cette méthode est le temps de calcul par conséquent le temps de réponse.

### 3.5 Etude comparative des méthodes de fusion

Avant de se prononcer sur le choix des méthodes de fusion, il est judicieux de relater quelques études comparatives entre ces dernières. Beaucoup de recherches sont faites pour déterminer les avantages et les inconvénients de chacune des méthodes de fusion, nous citons quelques unes ci-dessous.

Le travail effectué par M. MAMBE [39] l'a conduit à une proposition de méthode (ABC-Iot) basée sur la théorie de Dempster Cette dernière améliore la prise en charge de données plus conflictuelles et améliore les temps de réponses. Une étude comparative de différentes méthodes l'a conduit aux résultats donnés dans le tableau 3.1.

Techniques de fusion	Avantages	Limites
Réseaux bayésiens	Fonctionne bien lorsque des informations multi-sources sont disponibles. Peut combiner plusieurs sources de données.	Haute exigence de mémoire. Consomme plus de temps avant de produire un résultat.
La logique floue	Méthodes basées sur la décision. Réponse rapide aux changements rapides.	Plus d'efforts de calcul.
La théorie de Dempster Shafer	Méthode fondée sur la décision. Traite des problèmes statistiques	La complexité augmente lorsque les capteurs augmentent.

TAB. 3.1 : Analyse comparative des techniques de fusion [39].

Autre recherche qu'on peut citer est celle de S. Gîte et H. Agrawal [40] qui consiste à faire une étude sur la fusion de données dans un système multicapteurs. Leur étude se focalise beaucoup plus sur les Réseaux bayésiens et la théorie de Dempster-Shafer, néanmoins une étude comparative globale des différentes méthodes a été accomplie et se résume dans le tableau 3.2.

Techniques de fusion	Avantages	Limites
Réseaux bayésiens	Fonctionne bien lorsque des informations multi-sources sont disponibles. Peut combiner plusieurs sources de données. Méthode appropriée pour l'inférence de haut niveau	Besoin en mémoire élevé. Consomme plus de temps avant de produire un résultat.
La logique floue	Méthodes basées sur la décision. Réponse rapide aux changements rapides. Utile pour la représentation contextuelle.	Plus d'efforts de calcul. Les résultats peuvent être douteux
La théorie de Dempster-Shafer	Méthode fondée sur la décision. Traite des problèmes statistiques. Modélise l'incertitude. Peut tirer une bonne inférence avec moins de données de capteurs disponibles. Plus flexible que les approches probabilistes	La complexité augmente lorsque les capteurs augmentent. La croyance et la probabilité peuvent parfois différer

TAB. 3.2 : Comparaison entre les méthodes de fusion [40].

Aussi lors de cette étude, une comparaison entre la théorie de Dempster-Shafer et les réseaux bayésiens a été faite quant à leurs capacités d'inférer avec d'autres mécanismes et algorithmes lors de la fusion afin d'en tirer profit que ce soit en temps de fusion ou de rapidité de fusion. Leur comparaison se résume dans le tableau 3.3.

Paramètres	Réseaux bayésiens	Théorie Dempster-Shafer
Haut niveau de la fusion	Oui	oui
La prise de décision	Possible	Plus correct
Gestion de l'incertitude	Extension limitée	Effectivement fait
Tolérance de l'imprécision	Extension limitée	Possible
Disponibilité probabiliste d'informations	Fonctionne mieux quand l'intégralité de l'information est disponible	Fonctionne mieux que toutes autres méthodes
Combinaison avec d'autres algorithmes	Pas très utile	Produit des résultats plus précis et robustes

TAB. 3.3 : Comparaison entre les mécanismes d'inférence RB et DS [40].

### 3.6 Proposition de critères de sélection d'une méthode de fusion pour une application IoT

Après une petite synthèse de ces travaux de comparaisons, nous pouvons conclure que la sélection d'une méthode de fusion de données pour une application de IoT peut se faire selon plusieurs critères. Dans ce travail, les critères considérés et que nous jugeons essentiels sont :

- **La nature des données à fusionner.**
- **La densité du réseau IoT.**
- **La capacité de noeuds fusion.**
- **Le temps nécessaire pour la fusion.**

### Nature des données à fusionner

Dans le cas où la nature des informations à fusionner se présente sous la forme de données de probabilités et que toutes les données sont disponibles, la méthode probabiliste via les réseaux bayésiens est la plus adéquate sans omettre que la méthode des fonctions de croyances est aussi valable mais le temps de réponse dans ce cas est plus favorable pour la méthode probabiliste.

Autre cas qui s'offre à nous et le fait que les données sont de natures incertaines, les deux méthodes des sous-ensembles flous et les fonctions de croyance sont les plus appropriées.

Le cas où les données sont de nature conflictuelle, un sérieux avantage est donné à la méthode des fonctions de croyance, nous avons vu dans cette dernière comment le conflit est pris en charge explicitement dans les opérateurs de combinaison contrairement aux méthodes.

La méthode des sous-ensembles flous prend un sérieux avantage dans le cas où la nature des données sont de type à classer en ensemble ou démontrer l'appartenance d'une donnée à un ensemble bien déterminé.

### Densité du réseau IoT

La densité d'un réseau IoT reste un défi majeur pour le choix de la méthode de fusion. En effet quand le nombre de noeuds augmente la difficulté de fusion ainsi que les temps de réponse augmentent en conséquence. Il reste ainsi très difficile d'opter pour une méthode précise, le passage à une fusion distribuée [41] est pour le moment la meilleure solution possible ainsi on pourra faire intervenir plusieurs méthodes de fusion sur différents niveaux selon les données et selon les temps de réponses requis.

La figure 3.1 montre les différents niveaux de fusion. Le premier est fusion niveau bas où toutes les caractéristiques des données reçues sont fusionnées afin de prendre une décision unique. Le deuxième est fusion niveau objet où les données reçues sont d'abord traitées localement pour former les objets qui sont ensuite fusionnés dans un seul centre de fusion et de décision. Le dernier est fusion niveau décision avec la prise de décision localement sur les caractéristiques pour chaque noeud et fusionner ensuite les décisions locales pour avoir une décision finale.

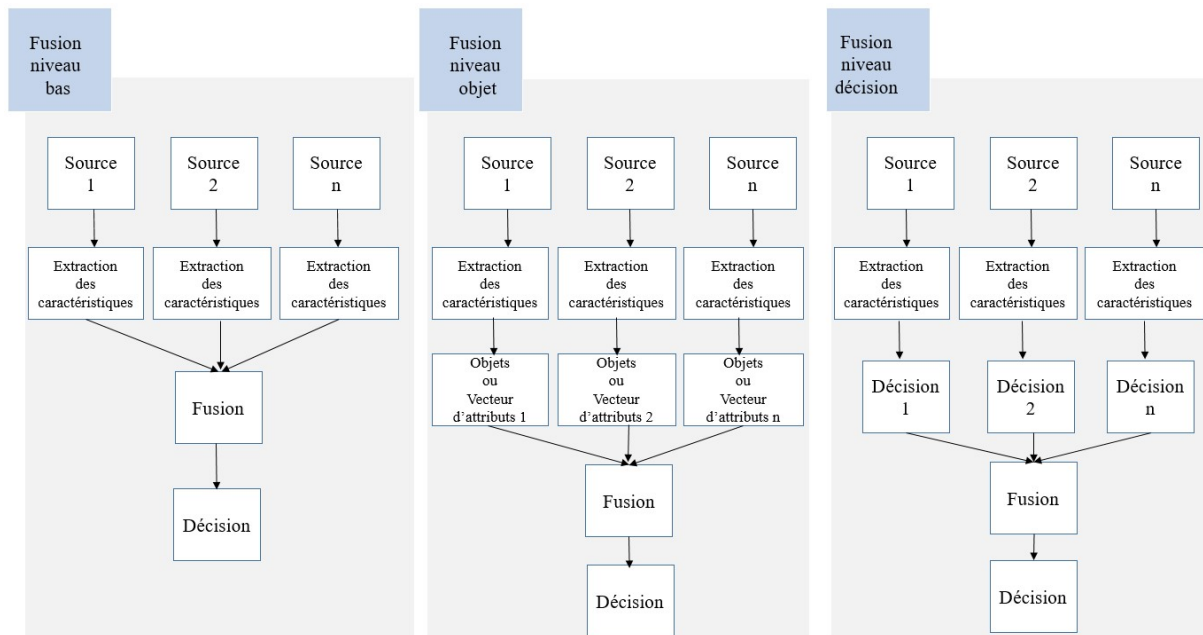


FIG. 3.1 : différent niveaux de fusion [34]

### Capacité des noeuds fusion

Un paramètre à prendre avec beaucoup rigueur est la capacité des noeuds. En effet, comme nous l'avons vu les réseaux bayésiens et les logiques floues nécessitent beaucoup de mémoire pour le calcul, ainsi c'est des méthodes à exclure si on a à fusionner des données avec des noeuds faibles en mémoire et CPU. Les méthodes basées sur les fonctions de croyance sont donc plus appropriées si le cadre de discernement est réduit. Les sous-ensembles flous sont également recommandés avec l'utilisation des opérateur de combinaison t-norme et t-conorme.

### Temps nécessaire pour la fusion

Pour ce paramètre toutes les méthodes se valent, et le choix de la méthode idoine dépendra de plusieurs autres paramètres tel que la nature des informations à fusionner ou bien le nombre de noeuds impliqués dans la fusion sans oublier le temps requis par l'application de l'IoT a ne pas dépasser.

## 3.7 Application d'aide au choix de la méthode de fusion

Afin de voir l'influence du nombre de noeuds sur le temps de la fusion, nous présentons les figures 3.2 [42] et 3.3 [39]. On constate que pour Dempster-Shafer et les réseaux bayésiens plus le nombre de capteurs augmente plus le temps de traitement augmente.



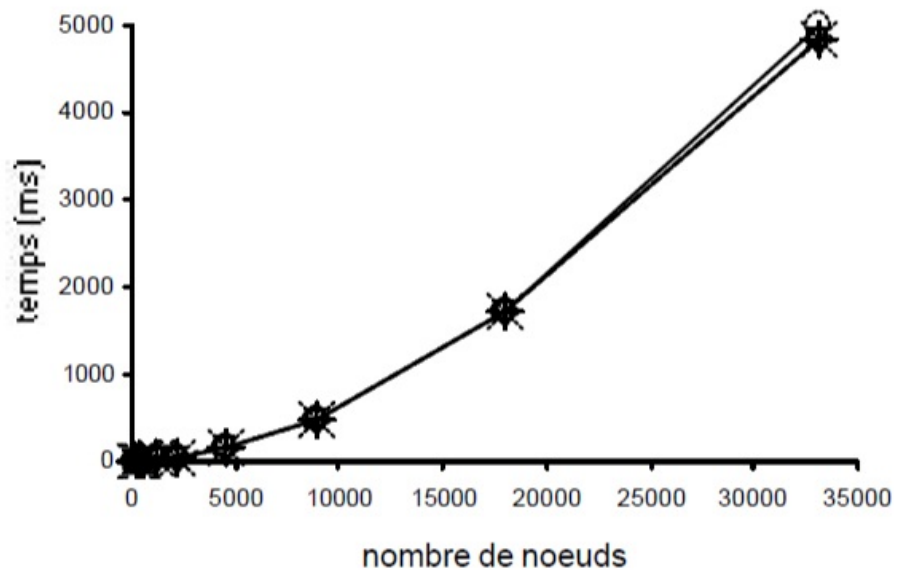


FIG. 3.2 : Le rapport entre le nombre de noeuds et le temps de réponse pour les réseaux bayésiens [42]

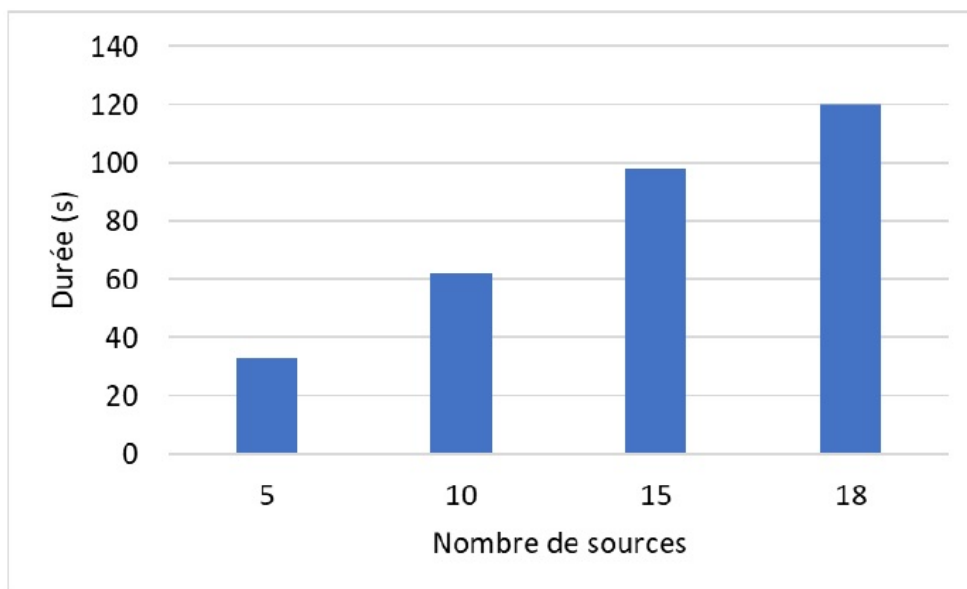


FIG. 3.3 : Le rapport entre le nombre de noeuds et le temps de réponse pour Dempster-Shafer [39]

On se penche maintenant sur le paramètre capacité des noeuds en termes de la puissance de transmis-

### Chapitre 3. Etude comparative et sélection d'une méthode de fusion

sion. Pour cela, un test a été établi dans [41] sur un réseau avec des émetteurs et récepteurs déterminés à l'avance ayant une puissance de transmission connues. La figure 3.4 démontre que plus les capteurs sont loin du récepteur plus le nombre de données reçus sont faibles. En conséquence pour un réseau de ce genre la distance maximal est de 63 m, au-delà de cette distance aucune données n'est exploitable, d'où l'obligation de passé a une fusion multi-couches où distribuée.

k	5,7E+14 Hz <sup>2</sup> .m <sup>2</sup>	d (m)	Pr (nW)
f	868 MHz	2	4728,413
Pe	25 mW	10	189,137
		20	47,284
		50	7,565
		100	1,891
		150	0,841
		200	0,473
		500	0,076
		1000	0,019
		1500	0,008
		2000	0,005

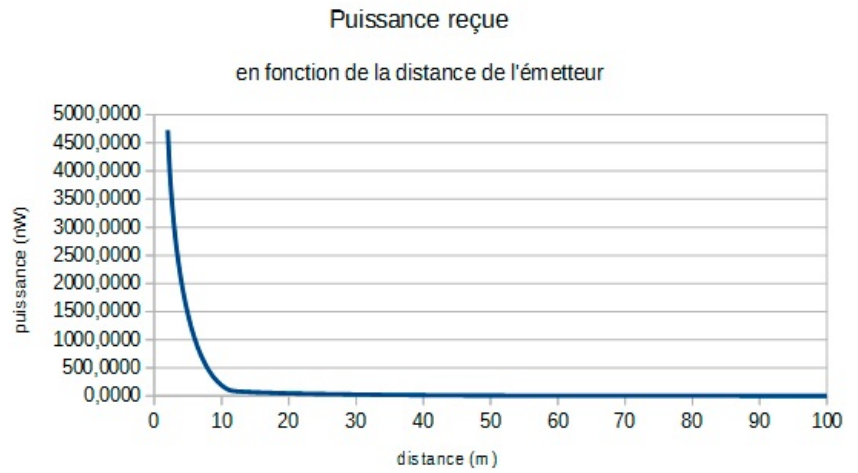


FIG. 3.4 : Test de réception de données selon la distance entre l'émetteur et le récepteur [41]

Donc, en fonction des conclusions tirées de ces travaux, nous avons opté pour une solution qui s'appuie sur la nature des données, la capacité du réseau et des capteurs telles que la distance maximale de transmission et la capacité mémoire à utiliser dans l'implantation. Au-delà d'une distance seuil, une fusion multi-couches est proposée.

La figure 3.5 montre l'écran d'accueil où nous allons paramétrer notre réseau selon lequel dépendra le choix de la méthode de fusion.

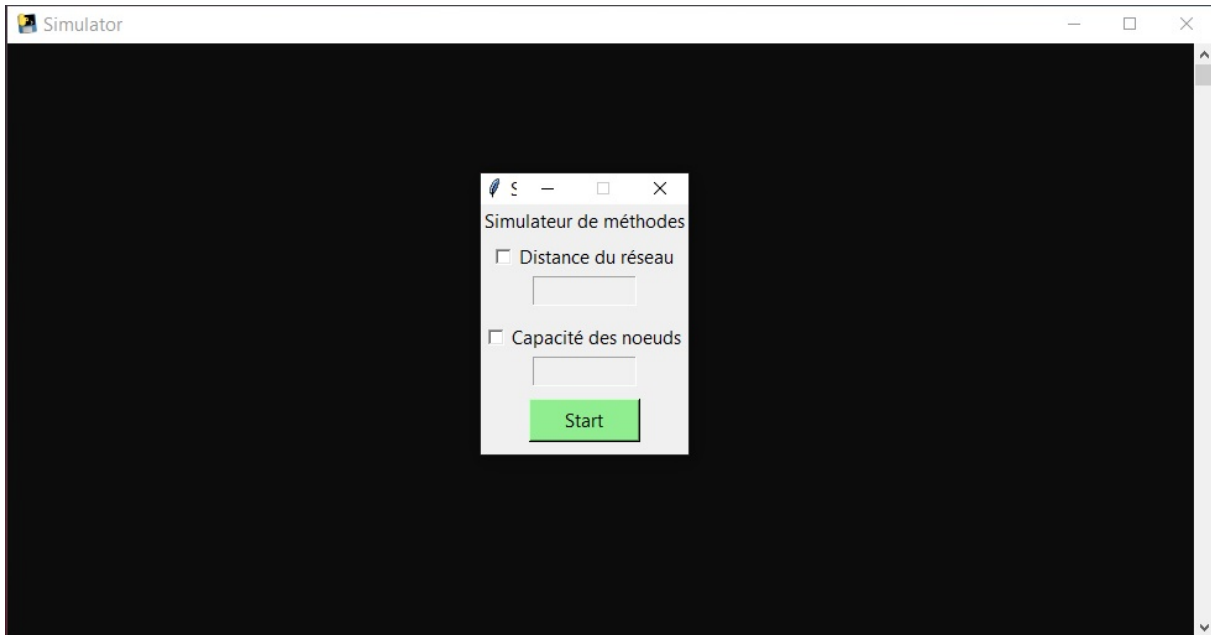


FIG. 3.5 : Page d'accueil de l'application de choix des méthodes.

Prenons l'exemple dans le cas où la fusion est de type locale (la distance du réseau est inférieure au seuil), l'application nous demande de spécifier la nature de l'information à fusionner comme c'est montré dans la figure 3.6, le choix de la méthode de fusion dépendra alors de la nature des informations. Dans notre proposition nous avons pris exemple sur trois types : des informations conflictuelles, des informations de type à classer et des informations probabilistes.

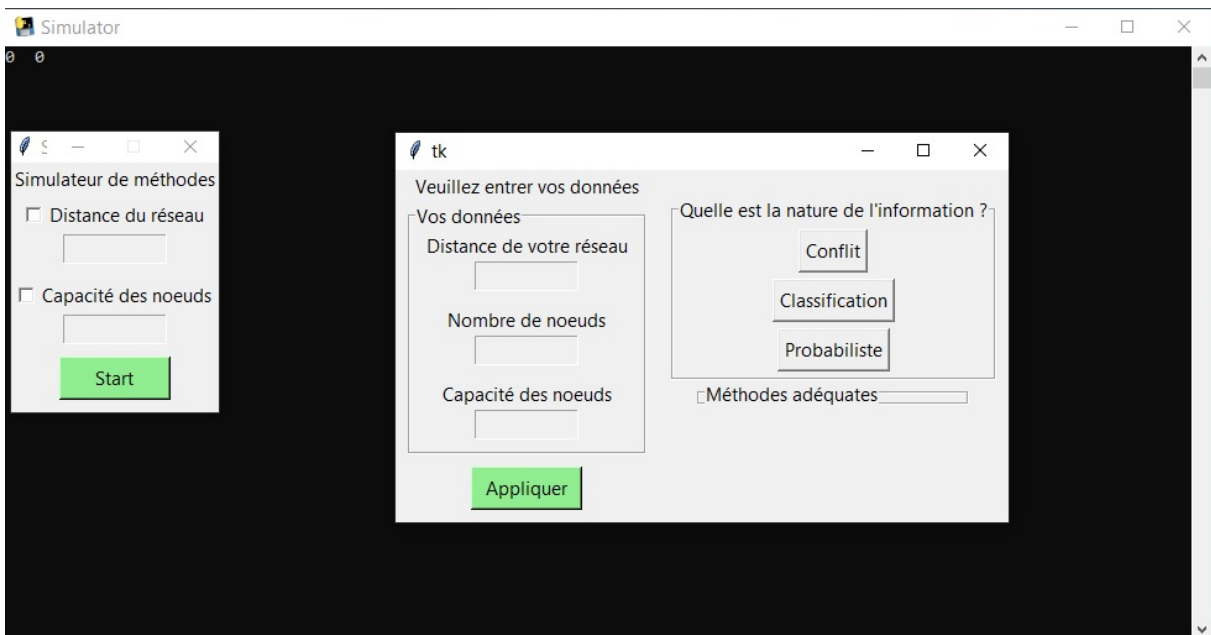


FIG. 3.6 : Choix d'une méthode dans le cas de fusion locale.

### Chapitre 3. Etude comparative et sélection d'une méthode de fusion

La figure 3.7 montre l'ordre des choix donné par l'application pour des informations de type à classifier.

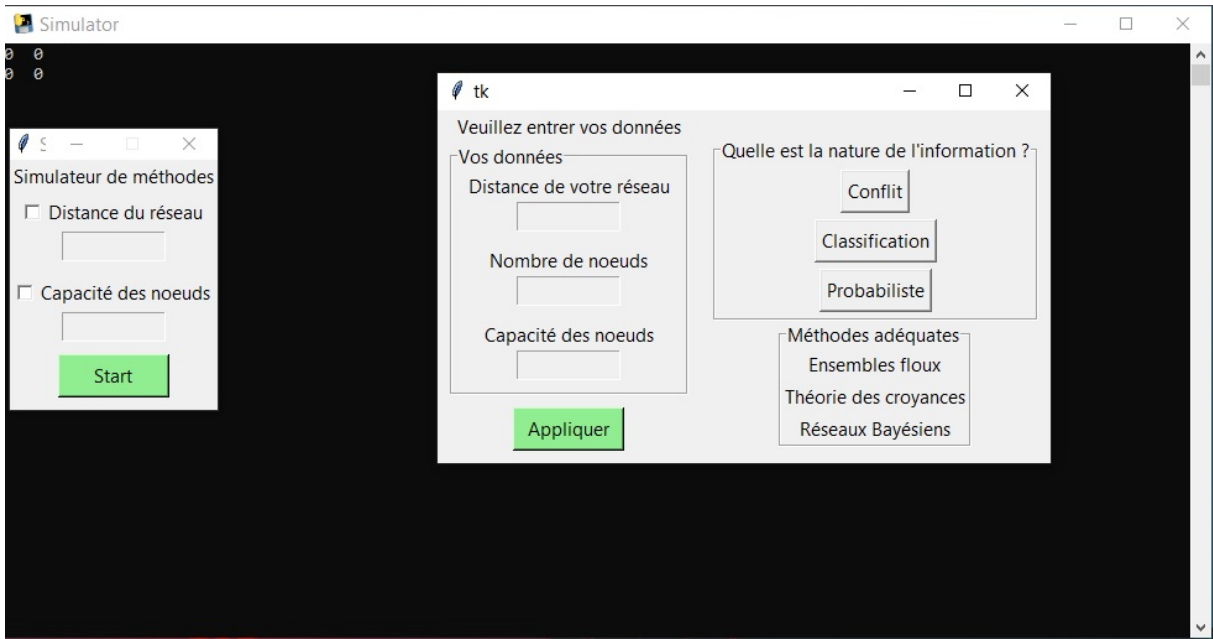


FIG. 3.7 : Méthode de fusion proposée dans le cas de la classification.

Pour le choix de méthode de fusion selon la densité du réseau, la figure 3.8 illustre un cas de figure.

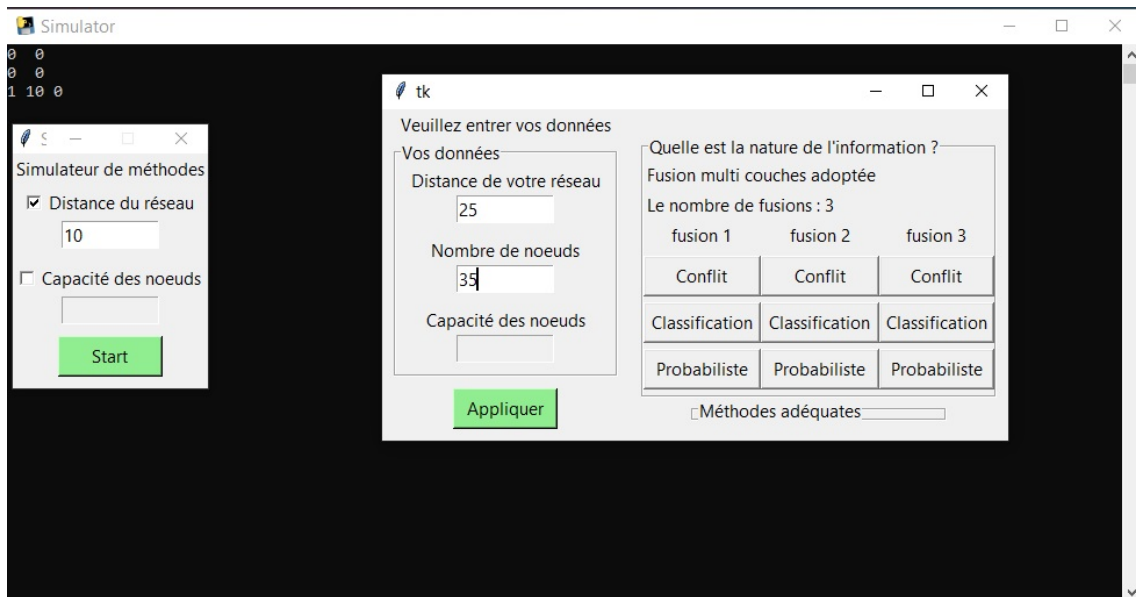


FIG. 3.8 : Choix pour la densité du réseau .

En ce qui concerne la capacité des noeuds, nous avons introduit un paramètre selon lequel l'application va décider d'exclure les méthodes qui nécessitent une capacité de mémoire supérieure à la norme établie.

Au cours de notre travail, nous avons constaté que les méthodes bayésiennes et des sous-ensembles flous nécessitent beaucoup de mémoire pour le calcul alors si les capteurs ont une capacité supérieure à la norme établie l'application va proposer une fusion local et demander le type d'information à fusionner, sinon la méthode des fonctions de croyance est la plus appropriée dans ce dernier cas de figure.

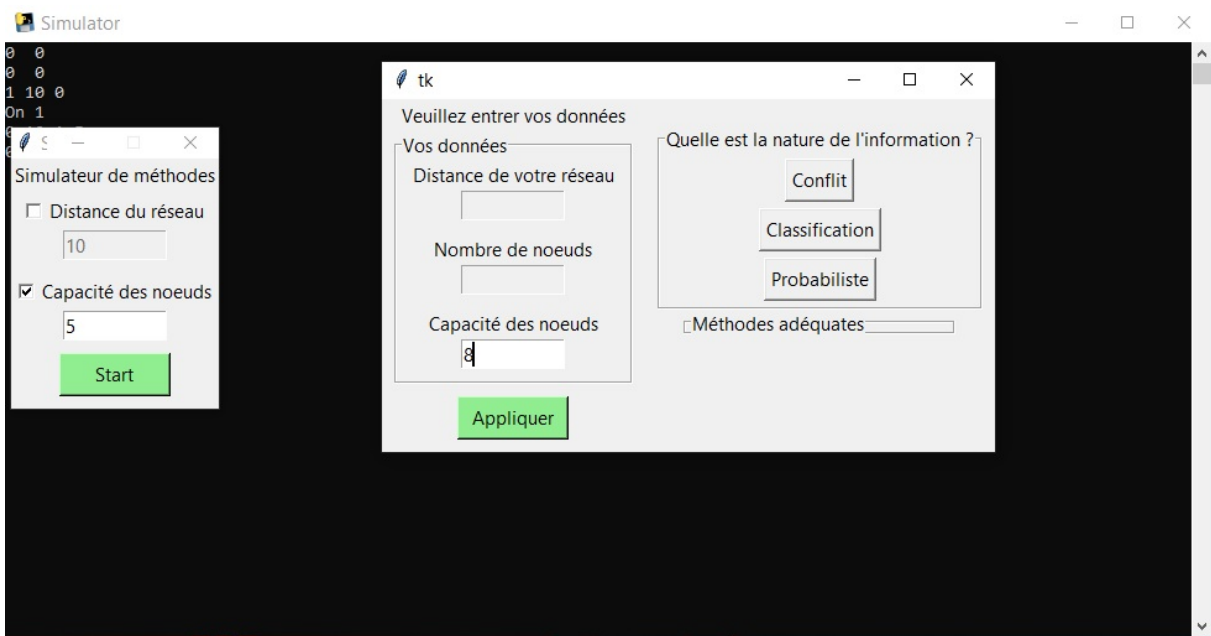


FIG. 3.9 : Choix selon la capacité des capteurs .

## 3.8 Conclusion

Dans ce dernier chapitre nous avons vu quelques études qui ont abouties à des conclusions sur les avantages et les limites des différentes méthodes de fusion. Nous nous somme basés sur ces études pour proposer une solution de sélection en intégrant des paramètres qui vont nous aider pour le choix de la méthode de fusion la plus appropriée pour une application IoT.

# Conclusion générale et perspectives

Dans le cadre de ce travail, nous avons proposé un système d'aide à la décision sur le choix de la méthode de fusion adéquate pour une application IoT. Ce choix se base sur des paramètres propres aux données à fusionner, les caractéristiques du réseau et des objets connectés.

Nous avons restreint notre travail à quelques méthodes de fusion que nous considérons les plus utilisées et les paramètres à ceux que nous jugeons les plus percutants.

Ainsi, nous avons présenté les méthodes probabilistes, les fonctions de croyance et les sous-ensemble flous et quelques études comparatives de ces méthodes. Quant aux paramètres considérés sont la nature de l'information, le durée de la fusion, la densité du réseau IoT et les capacités des objets connectés.

En perspectives, nous attachons de l'importance à étendre ce système aux autres méthodes de fusion, et en ce qui concerne l'automatisation du système. Nous projetons à le rendre plus autonome en le laissant définir lui même la nature des informations à fusionner, la densité du réseau ou même la capacité des noeuds et ceci selon les données reçues en amont par les objets. Quand au temps de réponse toutes les méthodes se valent à un point près, mais pour rendre la fusion plus robuste et le système plus fiable, nous envisageons dans un avenir que nous espérons proche combiner ce système a un réseau de neurones afin de lui permettre un apprentissage plus approfondie et une comparaison des résultats a la fin de chaque fusion, ceci lui permettra d'améliorer ses performances après chaque fusion effectuée.

Enfin nous espérons avoir contribué même de peu à tout les autres travaux qui visent à améliorer la fusion de données dans l'Internet des objets.

# Bibliographie

- [2] H. Chaouchi. "Objets connectés : Définition, Exemples, Enjeux patients et professionnels de santé. Perspectives européennes". p. 21.
- [5] A.H L'impact de l'internet des objets sur le futur de la logistique et du transport : cas du transport routier". fr. p. 129.
- [7] S. Ladjouzi et al. "Proposition d'un réseau de communication entre les objets dans une maison intelligente à base d'Internet des Objets". fr. p. 72.
- [8] L. Djakhdjakha. "Utilisation des Ontologies pour l'Intégration d'Internet des Objets dans la gestion des processus métier". fr. p. 58.
- [9] R. Marques. "SYSTÈME DE PRODUITS ET SERVICES BASÉS SUR L'IN-TERNET DES OBJETS : CONCEPTION ET IMPLANTATION PILOTE DANS UNE STATION-SERVICE". fr. p. 97.
- [10] L. de La Raudière. "La mission d'information sur ... est composée de : Mme Corinne Erhel ". fr. p. 129.
- [11] I. Saleh. "Les enjeux et les défis de l'Internet des Objets (IdO)". fr. In :IdO17.1 (avr. 2017).issn: 25148273.doi:10.21494/ISTE.OP.2017.0133.
- [12] R. Tafazolli. "TECHNOLOGIES FOR THE WIRELESS FUTURE". In :(2006).
- [13] I. Bloch : Fusion d'informations en traitement du signal et des images. Hermes Science Publication, Janvier 2003.
- [14] D. Hall and J. Llinas. HandBook of multisensor data fusion. CRC Press LLC, 2001.
- [15] Data Fusion Lexicon, Data Fusion Subpanel of the Joint Directos of Laboratories Technical Panel for C3. Rapport technique, F.E. White, Code 42402, NOSC, San Diego, Californie, 1991.
- [16] L. Wald : Data Fusion - Définitions and architectures - Fusion of images of diferent spacial resolutions. Les Presses de l'École des Mines, 2002.
- [17] F. Delmotte : Un nouvel opérateur de fusion adaptatif. Traitement du Signal, 17(4):299-311, 2000.
- [18] D. Dubois et H. Prade : La fusion d'informations imprécises. Traitement du Signal, 11(6):447-458, 1994.
- [19] F. Delmotte et P. Borne : Modeling of Reliability with Possibility Theory. IEEE Transactions on Systems, Man, and Cybernetics - Part A : Systems and Humans, 28(1):7888, Janvier 1998.
- [20] M. Delplanque, A.M. Desolt-Jolly et J. Jamin : Fusion dissymétrique d'informations incomplètes pour la classification d'objets sous-marins. Traitement du Signal, 14(5):511-522, 1997.
- [21] A. Clémentin, L. Delahoche, E. Brassart et C. Drocourt : Une architecture de propagation de l'incertitude pour le problème de la localisation. Revue des Sciences et Technique de l'Information - Technique et Sciences Informatiques ,22(7-8):903-933
- [22] P. Walley : Imprecise probabilities. In C. B. Read, D. L. Banks et S. Kotz,éditeurs : The Encyclopedia of Statistical Sciences. Wiley, New York (NY), 1997.
- [23] C.E. Shannon : A mathematical theory of communication. Bell System Technical Journal, 27:379-423, Octobre 1948.

- [24] A. Appriou : Décision et Reconnaissance des formes en signal, chapitre Discrimination multisérial par la théorie de l'évidence, pages 219-258. Hermes Science Publication, 2002.
- [25] S. Vural et E. Ekici, "On multihop distances in wireless sensor network with random node locations," IEEE Trans. Informatique mobile, vol. 9, non. 4, p. 540-552, avril 2010.
- [26] A.P. Dempster : Upper and Lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325-339, 1967.
- [27] G. Shafer : A mathematical theory of evidence. Princeton University Press, 1976.
- [28] Ph. Smets : The Combination of Evidence in the Transferable Belief Model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(5):447-458, 1990.
- [29] R.R. Yager : On the Dempster-Shafer Framework and New Combination Rules. *Information Sciences*, 41:93-137, 1987.
- [30] Ph. Smets : Belief Functions : the Disjunctive Rule of Combination and the Generalized Bayesian Theorem. *International Journal of Approximate Reasoning*, 9:1-35, 1993.
- [31] D. Dubois et H. Prade : Representation and combination of uncertainty with belief functions and possibility measures. *Computational Intelligence*, 4:244-264, 1988.
- [32] D. Dubois et H. Prade : Théorie des possibilités. Masson, Novembre 1987. Défié
- [33] S. Boukaboul et D. Djenouri. "Fusion de données pour l'Internet des objets". *Journal de la gestion des réseaux et des systèmes* (2020). doi: 10.1007/s10922-020-09519-y
- [34] Z. Baloutche et al. "Une approche de fusion de données contextuelle pour la santé-IoT". *Institut des applications informatiques et de la gestion de Bharati Vidyapeeth* (2018). *Int. j. inf. technique.* (Septembre 2018) 10(3) :241-245 doi :10.1007/s41870-018-0116-1
- [35] S. Chauvin : évaluation des performances du modèle Bayésien de fusion appliqué à l'imagerie satellitaire. In GRETSI, Juan-les-Pins, France, Septembre 1995.
- [36] L. Xu, A. Krzyzak et C.Y. Suen : Methods of Combining Multiple Classifiers and Their Application to Handwriting Recognition. *IEEE Transactions on Systems, Man Cybernetics*, 22(3):418-435, Mai 1992.
- [38] A. Gattein et P. Vannoorenberghe : A Comparative Analysis Of Two Approches Using The Road Network For Tracking Ground Targets. In *International Conference on Information Fusion*, Stockholm, Sweden, Juin 2004.
- [39] M. MAMBE Digraï Moïse. Modélisation de la qualité de l'information dans l'internet des objets. Thèse de doctorat. INSTITUT NATIONAL POLYTECHNIQUE FÉLIX HOUPHOUËT-BOIGNY. Cote d'Ivoire 2019.
- [40] S. Gîte, H. Agrawal. Sur la connaissance du contexte pour la fusion de données multicapteurs dans l'IoT. Département CS/IT, SIT, Symbiosis International University, Pune, Inde. 85-93. 2016
- [41] Eduscol : Internet des objets, ministère de l'enseignement Français, 02-15, 2019
- [42] C.Séka Kotchi et al : Etude de la performance des algorithmes d'inférence dans les réseaux bayésiens : Laboratoire d'Automatique, de Mécanique et d'Informatique industrielles et Humaines, Université de Valenciennes et du Hainaut Cambrésis, France



# Webographie

[1] 10 facons de tirer avantage de l'Internet of Things. fr-fr. Déc. 015. url:<https://www.objetconnecte.com/10-facons-tirer-avantage-linternet-of-things-0812/>(visité le 31/05/2021).

[3] F. Heroufosse. Les capteurs connectés, incontournables pour l'industrie alimentaire?fr.url:<https://blog.wagralim.be/les-capteurs-connectes-incontournables-pour-lindustrie-alimentaire?fr>(visité le 31/05/2021).

[4] IoT et Cybersécurité : une nécessaire intégration. fr-FR. Section : Blog. Juil. 2020.url:  
<https://wixalia.com/iot-et-cybersecurite-une-necessaire-integration/>(visité le 31/05/2021).

[6] L'Internet des Objets, l'IoT, la technologie au service des Supply Chain. fr-FR. Fév.2020. url:<https://www.sitescommerciaux.com/linternet-des-objets-liot-la-technologie-au-service-des-supply-chain/>(visité le 31/05/2021).

[37] Cairn.info La méthode Bayésienne pour aider à évaluer l'efficacité des thérapies personnalisées

<https://www.cairn.info/revue-hegel-2017-2-page-113.htm> (visité le 15/08/2021)